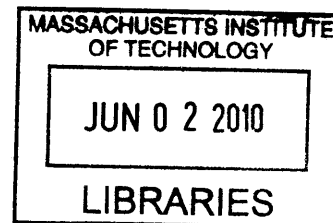# Protein MAS NMR Methodology and

Structural Analysis of Protein Assemblies

by

Marvin J. Bayro

B.A., Chemistry and Physics (2003)
Rutgers, The State University of New Jersey

Submitted to the Department of Chemistry
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Chemistry

at the

Massachusetts Institute of Technology

June 2010

Signature of Author _____
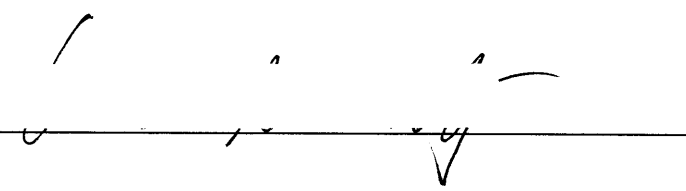
Department of Chemistry
May 18, 2010

Certified by_____

Robert G. Griffin
Professor of Chemistry
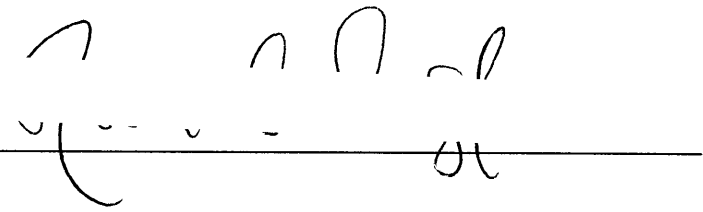Thesis Supervisor

Accepted by_____

Robert W. Field
Chairman, Departmental Committee on Graduate Students

This Doctoral thesis has been examined by a Committee of the Department of Chemistry as follows:
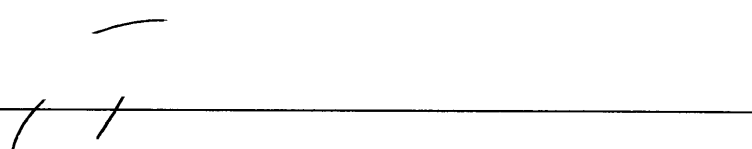
Professor Andrei Tokmakoff_____
Chairman

Professor Robert G. Griffin_____
Thesis Supervisor

Professor Troy Van Voorhis_____

Protein MAS NMR Methodology and Structural Analysis of Protein Assemblies

by

Marvin J. Bayro

Submitted to the Department of Chemistry
on May 18, 2010 in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Chemistry

Abstract

Methodological developments and applications of solid-state magic-angle spinning nuclear magnetic resonance (MAS NMR) spectroscopy, with particular emphasis on the analysis of protein structure, are described in this thesis. MAS NMR studies of biomolecules ranging from model peptides and proteins in crystalline form to amyloid fibrils and whole bacterial organelles are reported.

The methods presented include novel pulse sequences and optimized pulse sequence elements, experimental approaches designed for multiple-spin systems, a protocol for efficient sequential resonance assignment of proteins in the solid state, and techniques to determine the inter-molecular organization of amyloid fibrils formed by moderately sized proteins. Notably, an efficient dipolar recoupling technique, band-selective radio frequency-driven recoupling (BASE RFDR), is introduced and combined with alternating ^{13}C-^{12}C labeling to yield highly sensitive ^{13}C-^{13}C correlation spectra between distant nuclei in proteins. Various applications of the BASE RFDR scheme are presented, including protein resonance assignment, determination of tertiary structure of amyloid fibrils, and variable-temperature studies of protein dynamics.

The main biological systems analyzed are amyloid fibrils formed by the SH3 domain of PI3 kinase (PI3-SH3) and intact gas vesicles from *anabaena flos-aquae*, for which atomic-level structural information was previously unavailable. PI3-SH3 (86 residues) is a system thoroughly studied as a model of protein misfolding and amyloid formation by a natively globular protein. Gas vesicles are bacterial buoyancy organelles, with walls composed almost entirely by a single protein (GvpA, 70 residues), whose formation and structure constitute a highly intriguing biophysical problem. Nearly complete ^{13}C and ^{15}N resonance assignments and the molecular conformations of the polypeptide backbones of both PI3-SH3 and GvpA have been obtained via MAS NMR spectroscopy, enabling the proposal of models for the structure of these two protein assembly systems. In addition, the tertiary structure of PI3-SH3 amyloid fibrils has been elucidated by the application of novel methodology introduced in this thesis. Finally, investigations regarding the effects of temperature and protein dynamics on MAS NMR experiments and biomolecular dynamic nuclear polarization studies are presented.

Supervisor: Robert G. Griffin
Title: Professor of Chemistry and Director of the Francis Bitter Magnet Laboratory

Esta tesis está dedicada a mis padres, Elena y Juan Carlos,
y a mis abuelos, Tita, Bibi, papá Ernesto y papá Juan.

Contents

List of Figures

With brief descriptions

Figure 1-1
Simulations of polarization transfer between three coupled spins.

Figure 1-2
Single-crystallite numerical simulations of polarization exchange depicting the dipolar truncation effect.

Figure 2-1
RFDR pulse sequence highlighting different decoupling levels during ^{13}C pulses and during the windows between them.

Figure 2-2
Experimental ^{13}C magnetization profiles showing the effect of the decoupling field strength during the π pulses of RFDR at various spinning frequencies.

Figure 2-3
Experimental ^{13}C magnetization monitoring profiles at 30 kHz MAS frequency illustrating the complexity of the depolarization conditions encountered at high spinning frequencies, during both the ^{13}C pulses and the windows between them.

Figure 2-4
Contour plots of simulated ^{13}C magnetization remaining after a 1.6 ms RFDR period as a function of the decoupling fields, for increasing values of spinning frequency and ^{13}C π pulse strength.

Figure 2-5
Comparison of different RFDR pulse bandwidths at 30 kHz MAS frequency and 750 MHz ^{1}H Larmor frequency.

Figure 2-6
Long-range RFDR correlation spectra of [U-^{13}C, ^{15}N] N-f-MLF-OH recorded at 30 kHz MAS frequency and 750 MHz ^{1}H Larmor frequency with high-power decoupling (a) and without decoupling irradiation (b) throughout the mixing period.

Figure 2-7
RFDR correlation spectrum of protein [U-^{13}C, ^{15}N] GB1 recorded at 30 kHz MAS frequency and 750 MHz ^{1}H Larmor frequency without ^{1}H decoupling during the mixing period.

Figure 3.1-1
CMAR pulse sequence (a). No ^{1}H decoupling r.f. field is applied during the mixing time. (b) ^{13}C r.f. phase modulation applied during the recoupling period.

Figure 3.1-2
2D correlation spectrum of [U-^{13}C, ^{15}N] N-f-MLF-OH recorded using CMAR recoupling at 28.6 kHz MAS frequency without ^{1}H decoupling.

Figure 3.2-1
Pulse schemes for MOIST (a) and PATCHED (b) experiments

Figure 3.2-2
Profiles of recoupling efficiency versus radio frequency field offset for DCP, MOIST, and PATCHED schemes.

Figure 3.2-3
Profiles of ^{13}Cα-^{13}CO transfer efficiency versus rf field mismatch for CMRR and PAMORE pulse schemes.

Figure 3.2-4
Robustness of phase-shifted XY-32 against deviations in ^{13}C rf field for windowless low-power RFDR.

Figure 4-1
Labeling schemes employed in the model tripeptides used to illustrate the effect of dipolar truncation.

Figure 4-2
Simulated build-up curves of magnetization in model tripeptides using various recoupling pulse sequences.

Figure 4-3
Cross-sections of 2D RFDR correlation spectra demonstrating the attenuation of polarization transfer due to the dipolar truncation effect.

Figure 4-4
Experimental RFDR buildup curves for model tripeptides illustrating dipolar truncation and corresponding numerical simulations.

Figure 5-1
Band-selective radio frequency-driven recoupling (BASE RFDR) pulse sequence.

Figure 5-2
Long-range aliphatic correlation spectrum of 2-PI3-SH3 amyloid fibrils obtained with 17.92 ms of longitudinal mixing via BASE RFDR.

Figure 5-3
Section of BASE RFDR spectrum showing resonance assignments for multiple cross-peaks.

Figure 5-4
Comparison of long-range aliphatic correlations in 2-PI3-SH3 obtained with 17.92 ms of BASE RFDR and 500 ms of DARR mixing.

Figure 5-S1A
Numerical simulations of polarization transfer via RFDR-type longitudinal mixing at 10 kHz spinning frequency for various ^{13}C recoupling π-pulse nutation frequencies.

Figure 5-S1B
Numerical simulations comparing of effects of chemical shift offset and ^{1}H decoupling for various ^{13}C rf fields in a RFDR-type sequence at 30 kHz MAS frequency.

Figure 5-S1C
The BASE RFDR scheme. Recoupling is limited to spins with resonances within the bandwidth of low-power π ^{13}C pulses, effectively reducing the spin system.

Figure 5-S1D
Experimental demonstration of the BASE RFDR scheme. Transfer efficiency reaches 25% over a 4.9Å distance by diminishing the effect of interfering spins. In the broadband case polarization efficiency is lower (~7%).

Figure 5-S2
Numerical simulations of polarization transfer during BASE RFDR mediated by different dipolar interactions showing that direct ^{13}C-^{13}C couplings are dominant.

Figure 5-S3
Broadband RFDR spectrum of 2-PI3-SH3 fibrils acquired at 16 kHz spinning frequency and 16.4 T (700 MHz ^{1}H Larmor frequency).

Figure 5-S4
Overlay of BASE RFDR and DARR spectra of 2-PI3-SH3 fibrils.

Figure 5-S5
Simulations of transverse (a and b) and longitudinal (c and d) polarization transfer under a HORROR-type condition ($\omega_r/2\pi$ = 10 kHz, $\omega_1/2\pi$ = 5 kHz) using single-phase irradiation (a and c) and an XY phase sequence (b and d).

Figure 5-S6
^{13}C spectrum of [U-^{13}C, ^{15}N] N-a-VL recorded at 500 MHz ^{1}H Larmor frequency, showing chemical shift distribution in kHz, and the approximate band width of BASE RFDR pulses, down to the windowless limit.

Figure 5-S7
Windowless BASE RFDR spectrum of [U-^{13}C, ^{15}N] N-a-VL recorded with a 24 ms mixing time and the carrier frequency near the carbonyl region. Strong cross-peaks are observed between the three ^{13}C nuclei in the targeted region, as well as a number of off-resonance cross-peaks.

Figure 5-S8
Windowless BASE RFDR spectrum of 2-PI3-SH3 recorded at 700 MHz ^{1}H Larmor frequency, with $\omega_r/2\pi$ = 12 kHz, $\omega_1/2\pi$ = 6.0 kHz, and a mixing period of 24 ms. The carrier frequency was placed in the middle of the ^{13}Cα region.

Figure 6-1
Example of a labeling patterns in a three-residue segment found in alternating ^{13}C-^{12}C labeling with [2-^{13}C]glycerol and possible polarization transfer pathways.

Figure 6-2
Possible combinations of residue pairs for the three classes of alternating ^{13}C-^{12}C labeling spin systems, listed in decreasing frequency of occurrence.

Figure 6-3
Complete set of four 2D spectra used in our approach to assign the majority of the 2-GB1 backbone: TEDOR-1, TEDOR-2, RFDR-2, and BASE RFDR.

Figure 7-1
1D ^{13}C MAS NMR spectra of U-PI3-SH3 fibrils recorded at 750 MHz ^{1}H Larmor frequency and $\omega_r/2\pi$ = 16.67 kHz.

Figure 7-2
2D ^{13}C-^{13}C correlation spectra of U-PI3-SH3

Figure 7-3
Section of a 2D ^{13}C-^{13}C RFDR (1.76 ms) spectrum of U-PI3-SH3 illustrating aliphatic side-chain correlations.

Figure 7-4
Comparison of cross-peak information attainable in broadband Cα-C' correlation spectra of U-PI3-SH3 and 2-PI3-SH3.

Figure 7-5
Chemical shift analysis and secondary structure of PI3-SH3 in amyloid fibril form.

Figure 7-6
Secondary structure of PI3-SH3 amyloid fibrils in the context of previous structural and mechanistic studies.

Figure 7-7
Possible model for PI3-SH3 amyloid fibril architecture.

Figure 7-S1
Temperature dependence of ^{13}C cross-polarization MAS NMR spectra of U-PI3-SH3.

Figure 7-S2
Comparison of the spectral quality of ^{13}C CP spectra of PI3-SH3 samples with plain buffer and a glycerol/buffer mixture.

Figure 7-S3
2D ^{13}C-^{13}C correlation spectrum of U-PI3-SH3 recorded with DREAM mixing, showing various assigned resonances.

Figure 7-S4
Homonuclear and heteronuclear correlation spectra of 2-PI3-SH3 recorded PDSD (200 ms mixing time) and TEDOR (1.6 and 6.0 ms mixing time), respectively

Figure 7-S5
Full plot of the chemical shift analysis of PI3SH3 resonances

Figure 7-S6
Precision of the torsion angles predicted by TALOS analysis.

Figure 7-S7
Plot of the ψ vs. ϕ torsion angles in the backbone of PI3-SH3 in amyloid fibril form.

Figure 8-1
Subsection of a BASE RFDR spectrum of 2-GB1 showing various cross-peaks between nuclei that are distant in space and in sequence, including inter-molecular contacts.

Figure 8-2
Inter-nuclear distances in the crystal structure of GB1 corresponding to cross-peaks shown in Figure 8.1.

Figure 8-3
Inter-nuclear distances in parallel β-strands and resolvable ^{13}Cα-^{13}Cα correlations in three distinct strands and three identical in-register strands.

Figure 8-4
Section of a BASE RFDR spectrum of 2-PI3-SH3 fibrils presenting cross-peaks between ^{13}Cα nuclei separated by one and two residues (approx. 3.8 Å and 6.5 Å, respectively).

Figure 8-5
PDSD correlation spectrum of U-PI3-SH3 acquired with a short mixing time and 2-PI3-SH3 spectra acquired with short and long mixing times. In the latter, certain correlations are attributed to identical residues in adjacent parallel, in-register β-strands.

Figure 8-6
Inter-strand contacts between fractionally labeled ^{13}C sites account for the correlations highlighted in Figure 8.5.

18

Figure 8-7
Carbonyl region of a TEDOR 1D build-up curve on a sample with mixed labeling.

Figure 9-1
Electron micrograph of gas vesicles from *Anabaena flos-aquae*.

Figure 9-2
(a) Amino acid sequence of *A. flos-aquae* GvpA and corresponding PSIPRED secondary structure prediction. (b) BLAST results for residue conservation among GvpA sequences in different species.

Figure 9-3
Dependence of the strand tilt relative to the vesicle cylinder axis on the alignment of strands in the putative β-sheet.

Figure 9-4
Model of the β-sheet portion of GvpA with even-numbered turns in identical subunits and in-register arrangement of the conserved aromatic and charged residues at subunit interfaces.

Figure 9-5
Models of the β-sheet portion of GvpA with all the salt bridges on one side of the sheet, identical subunits, and in-register arrangement of conserved aromatic and charged residues at the subunit interfaces.

Figure 9-6
Models of the β-sheet portion of GvpA with all salt-bridges on one side of the sheet, identical subunits, and correctly tilted strands obtained by shifting subunits relative to each other.

Figure 9-7
Model of the β-sheet portion of GvpA with all salt-bridges on one side of the sheet, in-register alignment of conserved aromatic and charged residues at the subunit interfaces, and correctly tilted strands, obtained by allowing inequivalent subunits with odd-numbered turns centered alternately on G35 and V34.

Figure 9-8
Glycine region of an NCACX correlation spectrum showing assigned and duplicated cross-peaks.

Figure 9-9
Correlations between two distinct S49 amide ^{15}N signals and S49 and A48 ^{13}C nuclei obtained from NCACX and NCOCX spectra, respectively.

Figure 9-10
Correlations between two distinct Y53 amide ^{15}N signals and Y53 and T52 ^{13}C nuclei obtained from NCACX and NCOCX spectra, respectively.

Figure 9-11
Alanine CA-CB region of an RFDR spectrum in which at least 14 peaks are observed, although there are only 11 alanine residues in the sequence.

Figure 10-1
^{13}C MAS NMR spectra of GvpA acquired at 700 MHz ^1H Larmor frequency and 12.5 kHz spinning frequency.

Figure 10-2
^{13}C-^{13}C correlation spectrum acquired at 750 MHz ^1H Larmor frequency and 18 kHz MAS frequency with 3.52 ms of RFDR mixing. Cross-peaks corresponding to CA-CB correlations for several assigned residues are indicated.

Figure 10-3
^{15}N-$^{13}C\alpha$ region of an NCACX spectrum acquired at 900 MHz ^1H Larmor frequency and 20 kHz MAS frequency. The experiment used ^{15}N-^{13}C cross-polarization followed by ^{13}C-^{13}C RFDR mixing. Labels correspond to N-CA correlations, unless otherwise noted.

Figure 10-4
Secondary structure analysis of GvpA chemical shifts and comparison to PSIPRED results.

Figure 10-5
Ramachandran plot for GvpA using dihedral angles from TALOS.

Figure 10-6
Nearest-neighbor corrected ^{15}N secondary shifts, showing periodic patterns characteristic of curved helices.

Figure 10-7
^1H-^{13}C INEPT spectra showing highly dynamic ^{13}C sites in GvpA, recorded at two different temperatures.

Figure 10-8
Magic-angle spinning ^{13}C-^{13}C INEPT-TOBSY spectrum showing correlations between highly mobile ^{13}C nuclei.

Figure 10-9
NCA spectra of vesicles collapsed in ^2H_2O with (A) 0.5, and (B) 8.0 ms ^1H-^{15}N CP time. At short CP times, deuterium-exchanged ^{15}N nuclei are not polarized and their NCA cross-peaks are not visible, while at long CP times these ^{15}N nuclei are polarized by distant protons.

Figure 11-1
Sections of ^{13}C-^{13}C correlation spectra of PI3-SH3 fibrils corresponding to Lys side-chain resonances, recorded at ambient temperature with 300 ms PDSD mixing (top panels) and 18 ms BASE RFDR mixing (bottom panels).

Figure 11-2

Aliphatic region of ^{13}C cross-polarization spectra of PI3-SH3 fibrils recorded at decreasing temperatures. The temperature-dependent spectral changes are fully reversible.

Figure 11-3

Dynamic interference at low temperature observed by comparison of cross-polarization and Bloch decay (direct polarization) spectra.

Figure 11-4

Long-range correlation spectra of PI3-SH3 fibrils recorded with 16 ms BASE RFDR (a-c) and 300 ms PDSD (d-f) at various temperatures. At low temperature, slight cross-peak broadening is accompanied by increased volumes in BASE RFDR but not PDSD experiments.

Figure 11-5

Temperature-dependent enhancement of BASE RFDR polarization transfer efficiency in sequential Cα-Cα correlations.

Figure 12-1

(a) ^{13}C CP spectra of GNN[U-^{13}C,^{15}N QQN]Y fibrils obtained with DNP (top) and without DNP (bottom) at 400 MHz, $\omega_r/2\pi$ = 9 kHz, and 100 K, with a MW irradiation of 6.5 s and as the average of 16 scans. The enhancement factor fot the fibrils is 35, and it is the same for the glycerol resonances. (b) ^{13}C CP spectrum obtained with the same sample after the DNP experiments were performed, recorded at 750 MHz. (c) ^{13}C CP spectrum of fibril samples prepared without radical and glycerol, and never taken to low temperatures, recorded at 700 MHz. Experimental parameters for (b) and (c) are $\omega_r/2\pi$ = 9 kHz and 300 K, recorded as the average of 128 scans with recycle delay of 3 s.

Figure 12-2

^{13}C-^{13}C correlation spectra of GNN[U-^{13}C,^{15}N QQN]Y fibrils collected at (a) 750 MHz without DNP, using PDSD τ_{mix} = 10 ms, and at $\omega_r/2\pi$ = 9 kHz, 300 K, and (b) 400 MHz with DNP, using PDSD τ_{mix} = 5 ms, and at $\omega_r/2\pi$ = 9 kHz, 100 K.

Figure 12-3

^{15}N-^{13}C correlation spectra of GNN[U-^{13}C,^{15}N QQN]Y fibrils recorded with TEDOR mixing. (a) Spectrum recorded at 750 MHz, $\omega_r/2\pi$ = 9 kHz, 300 K and τ_{mix} = 1.8 ms. (b) and c) Spectra recorded with DNP at 400 MHz, $\omega_r/2\pi$ = 9 kHz, 100 K and τ_{mix} = 1.8 ms and 3.6 ms, respectively.

Figure 12-4

TEDOR build-up curves of GNN[U-^{13}C,^{15}N QQN]Y fibrils recorded (a) with DNP at 400 MHz ^{1}H Larmor frequency and a temperature of 100 K, and (b) without DNP at 750 MHz ^{1}H Larmor frequency and a temperature of 300 K.

Abbreviations

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
4D	Four-dimensional
AGG-2	Alanyl-[2-^{13}C]glycyl-[1-^{13}C]glycine
AGG-3	Alanyl-[1,2-^{13}C$_2$]glycyl-[1-^{13}C]glycine
AHT	Average Hamiltonian theory
BASE	Band selective
CMAR	Cosine-modulated adiabatic recoupling
CMRR	Cosine-modulated rotary resonance
CP	Cross polarization
CS	Chemical shift (isotropic)
CSA	Chemical shift anisotropy
CW	Continuous wave
DARR	Dipolar-assisted rotational resonance
DNP	Dynamic nuclear polarization
DRAMA	Dipolar recovery at the magic angle
DRAWS	Dipolar recoupling with a windowless sequence
DQ	Double quantum
EM	Electron microscopy
GB1	Beta-1 immunoglobulin-like domain of protein G
2-GB1	GB1 produced with [2-^{13}C] glycerol
GGV-3	[1-^{13}C]glycyl-[2-^{13}C]glycyl-[1-^{13}C]valine
GV	Gas vesicles
GvpA	Gas vesicle protein A
HORROR	Homonuclear rotary resonance
LP	Low power
MAS	Magic-angle spinning
MLF	N-formyl-methionyl-leucyl-phenylalanine

NMR	Nuclear magnetic resonance
PDSD	Proton-driven spin diffusion
PI3-SH3	SH3 domain of phosphatidylinositol 3-kinase
2-PI3-SH3	PI3-SH3 produced with $[2\text{-}^{13}C]$ glycerol
POST-C7	Permutationally offset stabilized C7
ppm	Parts per million
R^2	Rotational resonance
R^2T	Rotational resonance tickling
R^3	Rotary resonance recoupling
RAD	Radio-frequency assisted diffusion
REDOR	Rotational echo double resonance
RF	Radio frequency
RFDR	Radio frequency-driven recoupling
SH3	Src homology 3
SPC-5	Supercycled permutationally offset stabilized C5
SPECIFIC CP	Spectrally induced filtering in combination with CP
SSNMR	Solid-state nuclear magnetic resonance
TALOS	Torsion angle likelihood obtained from shifts and sequence similarity
TEDOR	Transferred echo double resonance
TEM	Transmission electron microscopy
TPPM	Two-pulse phase modulation
VL	Valyl-leucine
ZQ	Zero quantum

Acknowledgements

In a personal level, there are many people I want to thank for their friendship and support throughout my time at MIT. As it has been written before, "many a good time was had at the Muddy." Besides the occasional guests, the wings crew was Marc, Andy T., Ziad, Xander, Leo, Becky, Galia, Jason, Sean S., Tim, Bert, Bj, and yours truly. My closest friends from Rutgers, Paul, Mimi, and Mina, also have my enduring gratitude. Finally, I am indebted to my parents and my entire extended family in Perú for their everlasting encouragement, sincerity and love.

Chapter 1. Introduction

1.1 Preface

Much of the current understanding of biochemical and biophysical processes has been achieved by means of obtaining a detailed description of the molecular structure of proteins and nucleic acids. Such information is typically obtained using well-established methods, namely, X-ray crystallography and solution nuclear magnetic resonance (NMR) spectroscopy. Refined since the 1950's for the study of biological molecules, X-ray crystallography accounts for the vast majority of protein structures determined to date, particularly so in the case of very large proteins and protein complexes. Similarly, solution NMR spectroscopy allows the routine structure determination of moderately sized proteins and the study of dynamics, and has thus been an essential method in structural biology starting in the 1980's. Recent advances in solution NMR spectroscopy have facilitated the study of structure and dynamics of large proteins and intricate systems such as membrane proteins, protein complexes, and natively unfolded proteins. While solution NMR spectroscopy may be considered a conventional and largely routine method, it is a testament to the challenge of NMR spectroscopy applied to structural biology that most structure determination efforts continue to be carried out by highly specialized, expert laboratories.

Such challenge is most evident in the solid state. Solid-state NMR spectroscopy, despite having a rich history spanning several decades, has only recently begun to establish itself as an extremely versatile and powerful technique for the study of complex biological molecules. This development has been the result of improvements in all aspects of solid-state NMR spectroscopy, including instrumentation, experimental and theoretical development, and sample preparation. An inherent challenge in solid-state NMR spectroscopy is the anisotropic nature of nuclear spin interactions, for which approaches have been developed either via oriented systems or magic-angle spinning (MAS). The application of MAS in combination with the selective reintroduction of anisotropic couplings (recoupling), which contain distance and angular information at the atomic level, has enabled the elucidation of several peptide and protein structures in the solid state, with incrementing degrees of resolution in recent years. Indeed, the development of novel recoupling techniques may account for much of the success and promise of MAS NMR spectroscopy, along with resolution and sensitivity enhancement methods, isotopic enrichment schemes, among other advances. A more subtle but fundamental contribution has been the

vision, tenacity and determination of many researchers to help carry the field of solid-state NMR spectroscopy from its captivating beginnings to the forefront of physical chemistry and onto structural biology.

Solid-state NMR spectroscopy in general, and MAS NMR methodology in particular, offers the exciting capacity of providing atomic-level structural resolution in a large variety of biological systems, beyond what is typically tractable with X-ray crystallography and liquid-state NMR spectroscopy, which require diffraction-quality crystals and soluble samples, respectively. Instead, MAS NMR spectroscopy permits the study of proteins in their natural state, as in the case of protein aggregates and amyloid fibrils, or in conditions closely resembling it, as in the case of membrane proteins. Furthermore, since solid-state resonance line shapes do not depend on the molecular tumbling rate, arbitrarily large systems can be studied in principle, provided that sufficient resolution and sensitivity are available. While many of the key developments, including many contributed by this laboratory, have been based on small peptides and crystalline proteins, recent applications of nowadays-standard MAS NMR techniques include highly elaborate systems such protein fibrils, membrane proteins, and protein complexes. Therein lies the ongoing challenge for solid-state NMR spectroscopy. As the methodology continues to improve, the compromise between sample complexity and novel information attainable is diminishing progressively.

The work presented in this thesis is the summary of recent efforts by the author and co-workers in the Griffin laboratory towards the development of methodology for the study of proteins using MAS NMR spectroscopy and its application to characterize the *de novo* structure of two protein assembly systems. The original methodological contributions presented in this thesis can be categorized in the areas of pulse sequence design, resonance assignment strategies, and supra-molecular structure determination. The main biophysical applications undertaken in the course of this thesis have been (1) amyloid fibrils formed by the protein PI3-SH3, a model system for the mechanism of protein folding, unfolding, and fibril formation, and (2) gas vesicles, a bacterial organelle composed of protein GVpA and studied in intact form. The novel structural information obtained in our studies consists of the molecular conformations of PI3-SH3 in amyloid fibril form and GVpA within the gas vesicle assembly, and the inter-molecular architecture of PI3-SH3 fibrils. It must be noted that the complexity of the gas vesicle sample, which is harvested directly from bacteria, limited our approaches to standard MAS NMR

experiments, which nevertheless proved highly successful. On the other hand, PI3-SH3 samples provided sufficient sensitivity and could be prepared with various isotopic labeling schemes, so that we were able to use this system in the design of novel or improved experimental techniques. Thus PI3-SH3 amyloid fibrils have served as both an interesting application and as a realistic model system for MAS NMR methodology development.

1.2 Thesis Outline

This thesis is organized in three parts and twelve chapters. The remainder of this chapter contains first a description of the basic principles of NMR necessary for understanding both fundamental phenomena and key experiments, with an emphasis on solids and magic-angle spinning. Then a brief perspective on magic-angle spinning NMR spectroscopy is provided with which to understand the initial motivations behind much of the work described here.

Part I is entitled Development of Magic-Angle Spinning NMR Methodology for the Analysis of Protein Structure and includes various individual studies with the common objective of improving the efficiency of correlation spectroscopy in multiple-spin systems such as proteins, under different experimental conditions. The last chapter combines some of the concepts developed in previous chapters into a novel rapid backbone resonance assignment protocol for proteins in the solid state.

Chapter 2 describes the phenomenon of heteronuclear interference between the abundant ^{1}H bath and the observed ^{13}C spins, and how at increasing spinning frequencies it gives rise to magnetization losses due to depolarization in the course of a correlation experiment using radio frequency-driven recoupling (RFDR) mixing. This study was essentially an extension of early work in the group at moderate spinning frequencies, into an MAS frequency regime in which the complex interference patterns require careful experimental optimization and with the important conclusion that ^{1}H decoupling irradiation can be replaced by fast spinning frequencies (> 25 kHz) and strong ^{13}C pulses, even for fully protonated ^{13}C nuclei. One of the advantages of such experimental conditions is that eliminating the need for strong ^{1}H decoupling makes the ^{13}C-^{13}C transfer more efficient and reduces the amount of sample heating, an important concern in biological MAS NMR spectroscopy.

Chapter 3 presents various approaches for the compensation of experimental imperfections in dipolar recoupling pulse sequences and is divided into two sections. The first

section describes a novel dipolar recoupling scheme, cosine modulated adiabatic recoupling (CMAR), that implements an adiabatic sweep of the frequency of phase modulation through a resonant matching condition for efficient dipolar recoupling. CMAR was the first experimental demonstration of the concept of cosine modulated rotary resonance (CMRR) recoupling, a technique particularly well suited for fast MAS frequencies (20 kHz and above) due to its simultaneous recoupling of the irradiated nuclei (i.e., ^{13}C-^{13}C) and decoupling from other species (i.e., ^{1}H) with a single-channel radio frequency (r.f.) irradiation. The adiabatic sweep in CMAR makes is more robust with respect to phase errors than CMRR. The second section in this chapter presents various methods to reduce the effect of r.f. inhomogeneity in recoupling sequences. For that purpose, two techniques that utilize phase alternation to minimize the effects of r.f. inhomogeneity are introduced for the heteronuclear and homonuclear case, respectively. Finally, a novel 32-step phase sequence is demonstrated to compensate efficiently for r.f. field imperfection in RFDR-type sequences, resulting in improved experimental performance.

Chapter 4 explains and illustrates the dipolar truncation effect via theory, simulations, and experiments in model peptides specifically labeled to mimic three different spin systems, (1) a single spin pair with a long internuclear distance and the same spin pair in the case of (2) uniform ^{13}C isotopic labeling and (3) alternating ^{13}C-^{12}C labeling. The dipolar truncation effect is a multiple-spin phenomenon in which the recoupled weak dipolar interaction between a distant pair of nuclei is obscured by the simultaneous presence of a stronger dipolar coupling between one of the spins and a third one, due to the non-commutation of the interactions. This effect has deleterious consequences when highly efficient zero-order recoupling pulse sequences, such as the ones described in chapters 2 and 3, are applied to uniformly ^{13}C labeled peptides and proteins. Our results highlight the fact that while alternating labeling does not fully abolish dipolar truncation, it does facilitate the application of efficient zero-order recoupling schemes to spin systems presenting few one-bond dipolar couplings.

Chapter 5 introduces an experimental scheme through which to obtain highly efficient homonuclear correlations between distant spins in proteins. The approach utilizes band-selective (BASE) RFDR, a novel experimental regime for this well known pulse sequence in which the period of ^{13}C irradiation occupies the majority of the rotor period leading to efficient polarization transfer between closely spaced resonances at moderate spinning frequencies, in combination with alternating ^{13}C-^{12}C labeling and is demonstrated with a sample of PI3-SH3 amyloid fibrils.

The resulting spectra contain cross-peaks between distant aliphatic ^{13}C nuclei with unprecedented levels of sensitivity, surpassing robust spin diffusion techniques in many cases. This is the first example of such relative efficiency for distant spins, in the absence of dynamic interference. The chapter also includes additional examples of BASE RFDR in model systems, showing its performance under different experimental settings and applications to ^{15}N-^{15}N recoupling.

Chapter 6 describes a protocol for sequential assignment of protein resonances via a set of four two-dimensional MAS NMR dipolar recoupling experiments. Resonance assignment is a fundamental step in protein structure determination efforts and leads to site-specific conformation data derived from the analysis of ^{13}C and ^{15}N chemical shifts. The experiments employ in this protocol are aimed at exploiting the attenuation of dipolar truncation effects often found in ^{13}C spin systems of proteins prepared with ^{13}C-^{12}C alternating labeling, as determined by the analysis described in chapter 4. This advantageous property of alternating labeling allows the application of medium- and long-range mixing experiments to obtain correlations between distant ^{13}C-^{13}C and ^{15}N-^{13}C pairs in adjacent residues, and thus establish sequential connectivities that can aid in the resonance assignment process. A key experiment in this approach is BASE RFDR, described in chapter 5, which is used to obtain sequential Ca-Ca and Ca-Cb correlations. A broadband, two-bond RFDR experiment and TEDOR experiments for one- and two-bond ^{15}N-^{13}C correlations complete the set of spectra employ in this assignment protocol. The direct sequential correlations thus obtained in proteins prepared with [2-^{13}C]glycerol can be the analyzed in terms of the characteristic residue-pair patterns observed with this type of labeling, which leads to rapid assignment of the majority of the back-bone resonances for small to moderate size proteins.

Part II is entitled Applications to Amyloid Fibrils and Bacterial Organelles and summarizes our progress in characterizing the structure of proteins in two complex systems, an amyloid fibril formed by the 86-residue protein PI3-SH3, and gas vesicles, buoyancy organelles formed almost entirely of a single 70-residue protein. In both cases, nearly complete resonance assignments were obtained, from which the protein conformation could be derived. Additional MAS NMR experiments revealed further structural features in these highly intricate systems.

Chapter 7 reports the molecular conformation of the PI3-SH3 protein backbone in amyloid fibrils determined by MAS NMR spectroscopy. By application of the efficient resonance assignment protocol described in chapter 6 to a sample of PI3-SH3 produced with

alternating labeling via [2-^{13}C]glycerol, and additional experiments, resulted in almost complete assignment of ^{13}C and ^{15}N resonances of the entire polypeptide sequence. The secondary structure of the protein was obtained via chemical shift analysis and indicated regions of extensive β-sheet conformation, a few well-ordered random coil regions, and two short structurally disordered segments. This information, and the evaluation of dynamics, is consistent with a highly compact, rigid arrangement of the protein subunits that make up the fibril. The backbone conformation of PI3-SH3 in amyloid fibril form is found to be significantly different from the protein's native fold and agrees partially with algorithms that predict aggregation-prone regions in polypeptide chains. Comparison to previously published studies shows that the secondary structure of the protein in fibril form can help to explain the role of specific residues, and mutations thereof, in promoting or disrupting aggregation and fibril elongation. Finally, a simple structural model is proposed on the basis of spectral characteristics, secondary structure, and known fibril dimensions.

Chapter 8 outlines novel approaches to the elucidation of inter-molecular organization of protein assemblies and their application to determine the arrangement of protein subunits in PI3-SH3 amyloid fibrils. First we show that inter-molecular correlations in protein crystals can be obtained readily by applying the long-range homonuclear recoupling scheme introduced in chapter 5. The specific case of parallel, in-register inter-molecular arrangement is common in amyloid fibrils. In such architecture the β-sheets are formed between identical residue segments in neighboring molecules, which leads to degeneracy in the correlation spectra, i.e., cross-peaks between adjacent β-strands have the same chemical shifts as cross-peaks within each strand. Previous studies have addressed this issue with the incorporation of specific ^{13}C labels, but judicious use of that approach is limited to small peptides. A second option is to prepare samples with mixed ^{15}N and ^{13}C labels, which lead to spectra with compromised sensitivity. Instead, we show that the analysis of long-range correlation spectra of PI3-SH3 fibrils produced with alternating labeling can lead to the detection of a parallel, in-register intermolecular architecture through the identification of indirect and direct observations. Our conclusions are verified with independent measurements on a mixed ^{15}N/^{13}C sample, thus confirming the validity of our proposed techniques.

Chapter 9 presents the initial results obtained in the analysis of gas vesicles with MAS NMR methods. Little is known about how the proteinacious gas vesicle wall is formed, and

rough diffraction data have yet to be reconciled with a working model. Even though the gas vesicle wall is composed almost entirely of a single protein, GvpA, we observe multiple additional resonances in the NMR spectra, that appear to be duplicated peaks arising from parts of the protein that occupy two different conformations. Some of the duplicates are near regions expected to be flexible or predicted to be at the end of secondary structure elements, where turns may occur. We note that while several models of subunit arrangement may explain the basic features of the gas vesicle wall, the one that appears to be the most favorable in terms of molecular interactions is one in which the basic building block is a dimer with unequal subunits, of which, we suggest, the peak duplications may be evidence.

Chapter 10 describes the secondary structure of GvpA, the major component of wall of gas vesicles, derived from chemical shift analysis of resonances assignments obtained via MAS NMR spectroscopy. The secondary structure elements consist of an N-terminal α-helix, followed by two consecutive β-strands with a likely β-turn between then, and a C-terminal α-helix. This α-β-β-α structure is in stark contrast to previously proposed models based on X-ray fiber diffraction data, while it is compatible with the model we propose in chapter 9. We also present the results of experiments aimed at characterizing conformational flexibility in the protein and regions exposed to the solvent. This work is an excellent demonstration of the power and versatility of solid-state NMR spectroscopy to study highly complex systems such as these intact bacterial organelles and obtain molecular-level structural information.

Part III is entitled Manifestations of Protein Dynamics in MAS NMR Correlation Spectroscopy and describes studies aimed at characterizing how internal protein dynamics affect the quality and information content of solid-state NMR data. Protein dynamics in the solid state is an area of research that is emerging rapidly; however, the focus of this section is not the characterization of dynamics via MAS NMR spectroscopy, but rather to highlight how dynamic effects may be appear in solid-state NMR spectra, in particular as a function of temperature, including the cryogenic regime.

Chapter 11 presents variable-temperature MAS NMR studies on PI3-SH3 amyloid fibrils. Even though this system is shown to be highly rigid at room temperature, certain differences in correlation spectra of side-chain resonances may be understood from a dynamics perspective. Furthermore, as the temperature is decreased, line broadening is observed throughout the ^{13}C spectrum. We present an approach to identify dynamic interference effects as the source of line

broadening, and distinguish it from the case of static disorder, often expected at low temperatures. We also describe long-range backbone-backbone correlation spectra obtained at various temperatures with PDSD and BASE RFDR, and demonstrate that in the latter, significant gains in polarization transfer efficiency are achieved at temperatures below the sample freezing point. This result indicates that protein dynamics may have variable effects on correlation experiments depending on the mechanism of polarization transfer employed, and suggests that molecular mobility is an important experimental consideration even in systems that appear to be highly rigid at room temperature.

Chapter 12 describes MAS NMR experiments performed at cryogenic temperatures (100 K) on amyloid fibrils formed by a small peptide. We compare room temperature and 100 K cross-polarization and homonuclear and heteronuclear correlation spectra. Interestingly, few changes to the molecular conformation of this amyloid peptide as a result of temperature are identified, as chemical shift variations. Even though resonance lines at low temperature are broad, enough resolution is available to establish sequential assignments. Furthermore, the efficiency of dipolar recoupling schemes (for one- and two-bond correlations) improves considerably at low temperature. In particular, TEDOR ^{15}N-^{13}C polarization transfer curves are highly uniform as a result of restricted molecular flexibility, which perturbs the experiment at room temperature. The MAS NMR spectra recorded at 100 K utilized signal-enhancement via dynamic nuclear polarization (DNP), with an enhancement factor of 35 for the fibril resonances. The results discussed in this chapter therefore demonstrate that MAS NMR spectroscopy at cryogenic temperatures, enhanced by DNP, provides significant advantages over room-temperature spectroscopy, both in sensitivity and due to inherent dynamics at low temperatures.

1.3 Basics of MAS NMR Spectroscopy

This section describes a few basic concepts of nuclear magnetic resonance that are necessary to understand the rationale behind commonly used experiments, with particular emphasis on interactions that are dominant in the solid state, the effect of magic-angle spinning, and dipolar recoupling.

1.3.1 Nuclear spin interactions

The Hamiltonian of nuclear spin interactions can be divided into components that are internal to the spin system and externally applied fields:

$$H = H_{int} + H_{ext}$$

In NMR spectroscopy, a uniform external magnetic field is used to generate observable magnetization via the Zeeman effect and radio frequency pulses are applied to manipulate the spins and their interactions:

$$H_{ext} = H_Z + H_{RF}$$

The internal Hamiltonian of a system of multiple nuclear spins is given by

$$H_{int} = H_{CS} + H_J + H_D + H_Q$$

which represents the interactions that each spin experiences in their local fields. They includes the chemical shift, the indirect electron-mediated spin-spin interaction (or J-coupling), the direct dipole-dipole coupling, and the quadrupolar coupling. With the exception of the J-coupling and part of the chemical shift interaction, these local fields are anisotropic, that is, their effective strength depends on their angular orientation with respect to the external magnetic field. It is therefore interesting to note that in the case of liquid-state NMR, the dipolar and quadrupolar interactions and the anisotropic component of the chemical shift average to zero due to molecular tumbling, as long as the latter is isotropic and fast compared to the timescale of the experiments, while in the solid state all interactions are effective.

1.3.1.1 The external magnetic field

A nuclear spin with angular momentum μ experiencing an external applied magnetic field $\mathbf{B_0}$ can be described with the Zeeman Hamiltonian

$$H_Z = -\mu \cdot \mathbf{B_0} = -\gamma \hbar \mathbf{I} \cdot \mathbf{B_0}$$

35

where we have written the angular momentum vector in terms of the spin angular momentum operator **I**, while γ and $\hbar$ are the gyromagnetic ratio and Planck's constant, respectively. The direction of the magnetic field is typically defined as the z-axis in laboratory reference frame. Therefore, we can write

$$H_Z = -\gamma\hbar\mathbf{B}_0\mathbf{I}_z$$

The interaction of the spin angular momentum with the external magnetic field results in the precession of the nuclear magnetization around the external field with a frequency given by the magnitude

$$\omega_0 = \gamma\mathbf{B}_0$$

called the Larmor frequency, which describes the local field experienced by an isolated spin under the sole effect of the Zeeman interaction with the magnetic field B_0.

1.3.1.2 Zeeman states and polarization

The state of a nuclear spin system is best described in terms of quantum mechanics, with the eigenfunctions corresponding to the Zeeman Hamiltonian being represented in the basis of the spin angular momentum operator I_z, that is,

$$I_z|I,m_z\rangle = m_z|I,m_z\rangle$$

In the case of I = ½, the quantum number m_z, and thus the eigenvalues expressed above, can be ½ and -½. Thus we can the spin system under the Zeeman interaction

$$H_Z|I,m_z\rangle = -\hbar\omega_0I_z|I,m_z\rangle$$

has two eigenstates denoted $|1/2,1/2\rangle$ and $|1/2,-1/2\rangle$, with energies given by $E = -\hbar\omega_0 m_z$. Thus the energy difference between these Zeeman states is

$$\Delta E = \hbar\omega_0$$

Therefore, the precession frequency of magnetization around B_0 is proportional to the energy difference between the states.

In a system of multiple non-interacting nuclei, each spin may occupy either eigenstate, with the population of each state at equilibrium being given by the Boltzmann distribution

36

$$p_m = \frac{\exp(-E_m/kT)}{\sum_m \exp(-E_m/kT)}$$

where k is the Bolzmann constant and T is the temperature of the system. The observable magnetization along the z-axis of the laboratory reference frame can be evaluated by calculating the expectation value of I_z over the ensemble of spins, which consists of $p_{1/2}|1/2,1/2\rangle$ states and $p_{-1/2}|1/2,-1/2\rangle$ states. Evaluation of $\langle I_z \rangle$ over the ensemble shows that

$$\overline{\langle \mu_z \rangle} = \gamma\hbar\overline{\langle I_z \rangle} = \frac{1}{2}\gamma\hbar\left(p_{1/2} - p_{-1/2}\right)$$

That is, the average magnetization along the z-axis is given by the population difference between the two Zeeman states and is proportional to the gyromagnetic ratio. It is important to point out that difference in population between the states itself is proportional to the external magnetic field and inversely proportional to the temperature of the system. Indeed, expanding the exponentials in the Boltzmann distributions with a Taylor series gives

$$\exp\left(-\frac{E_m}{kT}\right) \approx 1 - \frac{E_m}{kT}$$

where the approximation holds for the small value of E_m relative to kT, so that

$$\sum_{m=\pm 1/2} \exp\left(-\frac{E_m}{kT}\right) \approx \left(1 - \frac{E_{1/2}}{kT}\right) + \left(1 - \frac{E_{-1/2}}{kT}\right) \approx 2$$

and the population difference, or polarization, is

$$P = \frac{\exp(-E_{1/2}/kT) - \exp(-E_{-1/2}/kT)}{\sum_m \exp(-E_m/kT)} = \frac{1}{2}\left[\left(1 - \frac{E_{1/2}}{kT}\right) - \left(1 - \frac{E_{-1/2}}{kT}\right)\right] = \frac{\Delta E}{2kT}$$

and thus we can write the magnetic polarization as

$$P = \frac{\gamma\hbar B_0}{2kT}$$

which reflects the relations mentioned above. As a consequence, NMR experiments are typically performed using strong magnetic fields in the order of 10-20 Tesla, and low temperatures have been explored as an approach to increase the observable magnetization.

37

1.3.1.3 Internal spin interactions

The Hamiltonians of internal spin interactions are typically represented as the product of tensors and spin angular momentum operators, so that in the case of a spin system consisting of spins i and j we can write the following, along with the associated frequencies of the local fields these interactions typically produce,

Chemical Shift	$H_{CS,i} = \gamma_i \mathbf{I}_i \cdot \boldsymbol{\sigma}_i \cdot \mathbf{B}_0$	1-20 kHz
J-coupling	$H_J = \gamma_i \gamma_j \mathbf{I}_i \cdot \mathbf{J}_{ij} \cdot \mathbf{I}_j$	10-100 Hz
Dipole-Dipole	$H_D = \gamma_i \gamma_j \mathbf{I}_i \cdot \mathbf{D}_{ij} \cdot \mathbf{I}_j$	1-30 kHz
Quadrupolar	$H_{Q,i} = \dfrac{eQ_i}{2I_i(2I_i - 1)\hbar} \mathbf{I}_i \cdot \mathbf{V}_i \cdot \mathbf{I}_i$	0.1-10 MHz

In the specific case of uniformly ^{13}C and ^{15}N labeled, diamagnetic peptides and proteins at high magnetic fields ($\geq$16.4 T, or 700 MHz ^{1}H Larmor frequency) these interactions have the following magnitudes. The chemical shift range is on the order of 30 kHz for ^{13}C, 8 kHz for ^{15}N, and 5 kHz for ^{1}H nuclei. J-couplings between directly bonded ^{13}C nuclei are 20-50 Hz, while they are 120-180 Hz for directly bonded ^{1}H-^{13}C nuclei. Dipolar couplings between directly bonded nuclei are as follows, 2.1 kHz for ^{13}C-^{13}C, ~23 kHz for ^{1}H-^{13}C, and ~10 for ^{1}H-^{15}N, while the dipolar coupling between near ^{1}H spins (as in the protons in a methyl group) is up to ~35 kHz. In contrast, the dipolar coupling of a two-bond ^{13}C-^{13}C pair is 500 Hz. Finally, the quadrupolar interaction is absent from systems of spin-½ nuclei, such as those considered in this thesis. The dipolar interaction, of fundamental importance for the determination of inter-nuclear distances and torsion angles, will be considered in more detail below.

1.3.1.4 Radio-frequency pulses

The general form of the Hamiltonian the radio-frequency field B_1 interacting with a single spin I can be expressed as a Zeeman interaction:

$$H_{RF} = -\gamma \hbar \mathbf{I} \cdot \mathbf{B}_1$$

where the axes are defined by external B_0 magnetic field in the z-direction of the laboratory reference frame. For the eigenbasis of the spin angular momentum operator I_z, it can be shown that only the components of H_{RF} that are orthogonal to the magnetization axis are able to excite

transitions between the eigenstates of I_z. Therefore we can simply consider a linearly oscillating radio-frequency field along the x-axis,

$$H_{RF} = -\gamma B_1 I_x \cos\left(\omega_{rf} t\right)$$

where ω_{rf} is the reference frequency of the pulse and γB_1 is the pulse amplitude, ω_1, referred to as the nutation frequency. In the rotating reference frame, the effect of a pulse with a reference frequency equal to the nuclear Larmor frequency can be written as the operator

$$H_{RF}(\varphi) = e^{-iI_x \varphi}$$

in which $\varphi = \omega_1 \tau$ and τ is the duration of the pulse.

1.3.2 Spin evolution under a Hamiltonian

In this section we consider the evolution of spin systems under the influence of a general Hamiltonian. Expressing the spin system as a density matrix $\rho(t)$, its evolution in time can be described by solving the Liouville-von Neuman equation

$$\frac{d}{dt}\rho(t) = -\frac{i}{\hbar}\left[H, \rho(t)\right]$$

In the case of a time-independent Hamiltonian, the general solution to this equation is

$$\rho(t) = e^{-(i/\hbar)Ht}\rho(0)e^{(i/\hbar)Ht}$$

where $\rho(0)$ is the density matrix at time zero. The expectation value of a time-independent operator O is thus

$$\langle O \rangle = Tr\left\{e^{-(i/\hbar)Ht}\rho(0)e^{(i/\hbar)Ht}O\right\}$$

In the case of a time-dependent Hamiltonian, evaluation of the spin system requires approximations. Introduced by Haeberlen and Waugh, Average Hamiltonian Theory (AHT) is the principle that a spin system under the time-dependent Hamiltonian arising from a periodic interaction can be described by evolution over an average Hamiltonian based on the effective, time-independent interaction over the repetition cycle. The propagator U(t), thus, is cyclic over the period τ_c and given by

$$U(\tau_c) = \exp(-i\overline{H}\tau_c)$$

where the average Hamiltonian is obtain by the Magnus expansion

$$\overline{H} = \overline{H^{(0)}} + \overline{H^{(1)}} + \overline{H^{(2)}} + \cdots$$

which consists of a series of time-independent Hamiltonians. The first few terms are

$$\overline{H^{(0)}} = \frac{1}{\tau_c} \int_0^{\tau_c} H(t)\, dt$$

$$\overline{H^{(1)}} = \frac{-i}{2\tau_c} \int_0^{\tau_c} dt_2 \int_0^{t_2} \left[H(t_1), H(t_2) \right] dt_1$$

$$\overline{H^{(2)}} = \frac{1}{6\tau_C} \int_0^{\tau_C} dt_3 \int_0^{t_3} dt_2 \int_0^{t_2} \left\{ \left[H(t_3),[H(t_2),H(t_1)] \right] + \left[H(t_1),[H(t_2),H(t_3)] \right] \right\} dt_1$$

The average Hamiltonian formalism is particularly well suited for magic-angle spinning, since it can be used to separate the evaluation of multiple simultaneous periodic interactions, as appropriate interactions frames can be defined.

1.3.3 Dipolar couplings and magic-angle spinning

We shall now discuss in detail the dipolar Hamiltonian of two interacting spin-½ nuclei in the context of magic-angle spinning, for nuclei of the same and unlike species. While all second-rank tensors are averaged over the course of a short rotation period around the magic angle, including the chemical shift anisotropy and parts of the quadrupolar interaction, we focus here on the dipolar coupling because its dependence on inter-atomic distances provides a direct means with which to elucidate molecular structure, which is part of the objectives in the applications presented in this thesis.

1.3.3.1 The dipole-dipole interaction

The energy of the interaction between two magnetic dipoles j and k separated by a distance r has the general form

$$E_D \propto \frac{\mu_1 \cdot \mu_2}{r^3} - 3\frac{(\mu_1 \cdot \mathbf{r})(\mu_2 \cdot \mathbf{r})}{r^5}$$

in which the vector $\mathbf{r}$ is a unit vector parallel to the line that joins the centers of the dipoles. If we now consider two interacting nuclear spins I_j and I_k, with an inter-nuclear distance vector r, the diopolar Hamiltonian in Cartesian coordinates is best expressed with the dipolar coupling tensor in matrix form

$$H_D = \frac{\gamma_j \gamma_k \hbar}{r^3} \begin{vmatrix} I_{jx} & I_{jy} & I_{jz} \end{vmatrix} \begin{vmatrix} 1 - \frac{3x^2}{r^2} & -\frac{3xy}{r^2} & -\frac{3xz}{r^2} \\ -\frac{3xy}{r^2} & 1 - \frac{3y^2}{r^2} & -\frac{3yz}{r^2} \\ -\frac{3xz}{r^2} & -\frac{3yz}{r^2} & 1 - \frac{3z^2}{r^2} \end{vmatrix} \begin{vmatrix} I_{kx} \\ I_{ky} \\ I_{kz} \end{vmatrix}$$

We can see that the dipolar coupling tensor is traceless and symmetric. In the case of like-spins (with the same gyromagnetic ratio), this Hamiltonian can also be written in the form

$$H_D = -\frac{\gamma^2 \hbar}{r^5} \left(\mathbf{r}_j \cdot \mathbf{r}_k - 3 r_{jz} r_{kz} \right) \left(\mathbf{I}_j \cdot \mathbf{I}_k - 3 I_{jz} I_{kz} \right)$$

which emphasizes the mathematical similarities between the spatial an spin factors of the nuclear dipole-dipole interaction

The dipolar Hamiltonian is more conveniently written in spherical coordinates and with raising and lowering operators, to yield the dipolar alphabet

$$H_D = \omega_D \left\{ A + B + C + D + E + F \right\}$$

whose terms (with θ defined as the angle between the inter-nuclear vector and the external magnetic field) are given by

$$\omega_D = \frac{\gamma_j \gamma_k \hbar}{r^3}$$

$$A = \left(1 - 3\cos^2 \theta \right) I_{jz} I_{kz}$$

$$B = -\frac{1}{4} \left(1 - 3\cos^2 \theta \right) \left(I_j^+ I_k^- + I_j^- I_k^+ \right)$$

$$C = -\frac{3}{2} \sin\theta \cos\theta \, e^{-i\phi} \left(I_{jz} I_k^+ + I_j^+ I_{kz} \right)$$

$$D = -\frac{3}{2} \sin\theta \cos\theta \, e^{i\phi} \left(I_{jz} I_k^- + I_j^- I_{kz} \right)$$

$$E = -\frac{3}{4} \sin^2 \theta \, e^{-i2\phi} I_j^+ I_k^+$$

$$F = -\frac{3}{4} \sin^2 \theta \, e^{i2\phi} I_j^- I_k^-$$

In the high-field or secular approximation, all remaining terms of the Hamiltonian that do not commute with that of the Zeeman interaction caused by the large external magnetic field, called non-secular terms, are neglected. Since in our case the high-field approximation is achieved

41

through transformation into a frame precessing at the nuclear Larmor frequency via rotations around the z-axis, the terms C, D, E, and F above become time-dependent and periodic, while the terms A and B are unaffected by z-rotations. The average Hamiltonian for the periodic terms is zero, to a first approximation, and thus the dipolar contribution to the first-order secular Hamiltonian in this two-spin system can be written, after rearranging the terms A and B into spin operators, as

$$H_D = \omega_D \left(3\cos^2\theta - 1\right)\left[2I_{jz}I_{kz} - \left(I_{jx}I_{kx} + I_{jy}I_{ky}\right)\right]$$

or more succinctly as

$$H_D = \omega_D \left(3\cos^2\theta - 1\right)\left[3I_{jz}I_{kz} - \mathbf{I}_j \cdot \mathbf{I}_k\right]$$

Therefore we can see, from the angular dependence of this interaction, that if the spatial factor is modulated in a rapid isotropic fashion, as in the case of molecules in solution, the average given by integration over all angles is zero

$$\left\langle 3\cos^2\theta - 1\right\rangle = \int_0^{2\pi}\left(3\cos^2\theta - 1\right)\sin\theta\, d\theta = 0$$

and thus the first order dipolar Hamiltonian vanishes. In the solid state, the effective strength of the interaction depends on the angular orientation of the dipole-dipole vector, and in the case of a powder is non-zero. However, in many types of solid or semi-solid samples, the molecules may undergo anisotropic motions or local conformational flexibility that may reduce the effective dipolar interaction to varying degrees. Finally, we can see that if the angle θ is 54.73°, referred to as the magic angle, the spatial term of the dipolar Hamiltonian is zero.

1.3.3.2 The dipolar Hamiltonian during MAS

The magic-angle spinning technique involves the rapid rotation of the sample around an axis oriented 54.73° away from the external magnetic field. As we shall see, the result is the averaging of anisotropic interactions with second-rank tensors, such as the dipole-dipole coupling. The spinning process induces a time dependence in the spatial component of the dipolar Hamiltonian, which we can incorporate into the dipolar coupling coefficient and write, again for the case of two interacting spins j and k,

$$H_D(t) = \omega_D(t)\left[3I_{jz}I_{kz} - \mathbf{I}_j \cdot \mathbf{I}_k\right]$$

In the heteronuclear case, that is, when spins j and k have different gyromagnetic ratios, the secular approximation leads to oscillating terms in the Hamiltonian with respect to two different Larmor frequencies. Thus, only terms that commute with both Zeeman Hamiltonians, namly I_j and I_k, remain, and thus the heteronuclear dipolar Hamiltonian for a two-spin system can be written as

$$H_D(t) = \omega_D(t) 2 I_{jz} I_{kz}$$

In order to describe the form of the time-dependent coefficient we must first introduce the system of reference frames used to described spin interactions in rotating solids, and the angle notation used to connect them. The principal axis system (P) describes spin interactions and is defined for each type of interaction. For the dipolar coupling, the z-axis of P is chosen parallel to the internuclear vector. The crystallite-fixed frame (C) is chosen arbitrarily in our description, so it can be set to be the same as P. In a powder sample, each crystallite is oriented randomly with respect to the rotor-fixed frame (R), which is defined with its z-axis being that of the axis of rotation in our description. Finally, the laboratory frame (L) is defined with its z-axis the same as that of the external magnetic field B_0. The angles that relate the different frames are $\Omega_{PR} = (\alpha_{PR}, \beta_{PR}, \gamma_{PR})$ and $\Omega_{RL} = (\alpha_{RL}, \beta_{RL}, \gamma_{RL})$, which are Euler angles with the convention used by Rose. Therefore, in a magic-angle spinnging experiment, $\beta_{RL} = 54.74°$. Other two very important angles are β_{PR} and γ_{PR}, over which the dipolar interaction is averaged in powder samples, leading, in general to reduced values of the effective dipolar Hamiltonian.

Within this reference frame, the time-dependence of the dipolar coefficient can be expressed as a Fourier series over the components of the rank-2 dipolar tensor

$$\omega_D(t) = \sum_{m=-2}^{2} \omega_D^{(m)} \exp(im\omega_r t)$$

in which $\omega^{(m)}$, in general, depends on isotropic, anisotropic and asymmetry parameters of the tensor. Since the dipolar interaction is a symmetric tensor and has no isotropic component (it is traceless), only the anisotropic parameter, b_{jk}, is non-zero and thus

$$\omega_D^{(m)} = b_{jk} D_{0,-m}^{(2)}(\Omega_{PR}) d_{-m,0}^{(2)}(\beta_{PR})$$

where $D(\Omega)$ and $d(\beta)$ are Wigner rotation and reduced Wigner matrices, respectively, and b_{jk} is the dipolar coupling constant given entirely by the internuclear distance between spin j and k, their gyromagnetic ratios, and physical constants. In angular units,

$$b_{jk} = \left(\frac{\mu_0}{4\pi}\right)\frac{\gamma_j \gamma_k \hbar}{r_{jk}^3}$$

The Fourier components of the dipolar interaction can then be written in terms of Euler angles between the principal and rotor frames

$$\omega_D^{(0)} = \frac{1}{4}b_{jk}\left(3\cos^2 \beta_{PR}^{jk} - 1\right)\left(3\cos^2 \beta_{RL} - 1\right)$$

$$\omega_D^{(\pm 1)} = -\frac{1}{2\sqrt{2}}b_{jk}\sin\left(2\beta_{PR}^{jk}\right)\exp\left(\pm i\gamma_{PR}^{jk}\right)$$

$$\omega_D^{(\pm 2)} = \frac{1}{4}b_{jk}\sin^2 \beta_{PR}^{jk}\exp\left(\pm i2\gamma_{PR}^{jk}\right)$$

In a magic-angle spinning experiment, the rotor axis is set to $\beta_{RL} = 54.74°$ with respect to the laboratory frame, thus the ω component is zero. The remaining components depend on the crystallite angles β_{PR} and γ_{PR}, distributed over all orientations in a powder sample.

1.3.4 Dipolar recoupling

A great variety of MAS NMR pulse sequences have been designed to generate Hamiltonians with non-zero dipolar contributions, which would be otherwise averaged to zero over the period of rotation. This section describes the basic concept of dipolar recoupling illustrated in an early example of the technique.

The principle of rotary resonance recoupling (R^3) consists of reintroducing the heteronuclear dipolar interaction between an observed spin I and a second spin of a different species S that is irradiated with a continuous rf pulse with a nutation frequency equal to a small multiple of the MAS frequency. Considering an isolated spin pair I and S with a heteronuclear dipolar coupling ω_D, rotor spinning frequency ω_r, and continuous irradiatioin with x-phase and nutation frequency that is an integer multiple of the spinning frequency, $n\omega_r$, the Hamiltonian of the system is

$$H(t) = \omega_D(t)2I_z S_z + n\omega_r S_x$$

where the time-dependence of the dipolar coupling is generated by the effect of magic-angle spinning on the anisotropic component of the interaction, as described above.

The approach to evaluation of the above Hamiltonian, in the coherent averaging formalism, is to transform H(t) into an interaction frame defined by the RF pulses

$$\tilde{H}_D(t) = e^{in\omega_r t S_x} H_D(t) e^{-in\omega_r t S_x}$$

which is simplified by the straightforward relation between rotor and rf frequencies. The result of the above transformation is

$$\tilde{H}_D(t) = \sum_{m=-2}^{2} \left[\omega_D^{(m)} \left(e^{i(m+n)\omega_r t} + e^{i(m-n)\omega_r t} \right) I_z S_z - i\omega_D^{(m)} \left(e^{i(m+n)\omega_r t} - e^{i(m-n)\omega_r t} \right) I_z S_y \right]$$

from which we can calculate the zero-order average Hamiltonian, as the first term of the Magnus expansion shown above, with the a rotor period as the cycle time, over which the Hamiltonian is integrated. Such integrals vanish except at the conditions m + n = 0 and m − n = 0, which connects the integral multiple between rf and spinning frequencies to coefficients of the dipolar tensor. The average, zero-order dipolar Hamiltonian is then

$$\overline{\tilde{H}_D^{(0)}} = \left(\omega_D^{(-n)} + \omega_D^{(n)} \right) I_z S_z - i \left(\omega_D^{(-n)} - \omega_D^{(n)} \right) I_z S_y$$

Therefore, the average dipolar Hamiltonian is non-zero only when the dipolar coefficients are introduced by satisfying the conditions n = 1 or n = 2, that is, when the rf field applied to the S spin is ω_r or 2 ω_r, constituting the matching conditions for rotary resonance recoupling.

1.4 An initial perspective

This section contains parts of the author's initial thesis proposal from 2005 and is included here with the purpose of highlighting the initial outlook with which the studies described in this thesis were undertaken. Considering the current state of the art, this perspective also helps to appreciate some of the multiple developments in the field of biological solid-state NMR spectroscopy over the past five years.

A. Specific goals

Our proposed research focuses on the development of new solid-state nuclear magnetic resonance spectroscopy methods for the elucidation of homonuclear interactions, and their application to problems of biological interest. The proposed work covers three main areas.

A.1 Homonuclear recoupling methods for state-of-the-art solid-state NMR

The advent of technological improvements in solid-state NMR, such as high-field magnets and very fast magic-angle spinning probes, is expected to facilitate the application of this technique to study complex biological system, which require high sensitivity and resolution. However, these promising technologies pose a challenge to the application of many of the solid-state NMR spectroscopy experiments commonly used today, since a large fraction of them were designed under the conditions of lower magnetic fields and moderate spinning rates. This limitation is particularly significant in homonuclear recoupling experiments, which are vital in the determination of molecular structure via solid-state NMR. In the first part of this proposal, we present novel homonuclear recoupling techniques developed to satisfy the demands imposed by high-resolution solid-state NMR technologies.

A.2 Systems of multiple spins: Approaches to the dipolar truncation problem

An important unsolved problem in solid-state NMR is that of fully understanding many-spin effects in conventional experiments. As a consequence of this theoretical limitation, quantitative techniques in the field have encountered success mostly in the study of samples containing strategically placed pairs of labeled nuclei. However, this approach is extremely restrictive if large systems, such as proteins and protein complexes, are to be studied. Furthermore, the complications of multiple-spin effects are present even in a three-spin system.

For example, it is well known that the observable interaction between two weakly dipole-dipole coupled spins is dramatically quenched if a third spin, strongly coupled to one of the first two, is present. This phenomenon is known as the dipolar truncation effect. In the second part of this report, we propose experimental and computational studies designed to gain insight into how dipolar truncation works, and examine the possibility of developing new experimental schemes to circumvent, or perhaps reduce the extent of, this deleterious effect.

A.3 Protein resonance assignments via solid-state NMR

The field of structural biology, which treats fundamental questions about biochemical processes in molecular detail and has caused a breakthrough in rational drug design, is based on the premise that the first step in understanding the function of biomolecules is determining their structure. The two well established structure determination methods, X-ray crystallography and solution NMR, have proven highly effective when applied to soluble globular proteins and their complexes. However, a great number of biochemical processes involve proteins in supramolecular assemblies, such as membrane proteins and amyloid structures, which are samples that are very problematic for the two conventional methods. Solid-state NMR spectroscopy promises to emerge as an alternative to X-ray and solution NMR in studying these challenging systems. In the third part of this proposal, we illustrate methods to perform resonance assignments, the initial step in structure determination via NMR, on a 76-residue protein, the gas vesicle protein A.

B. Background and Significance

B.1 Homonuclear recoupling

A major fraction of this report describes our new developments in the technique known in the field of solid-state NMR as broadband homonuclear recoupling. A brief description and historical summary are presented, followed by an overview of the difficulties and limitations the technique currently faces.

B.1.1 Dipolar recoupling in magic angle spinning SSNMR – Early work

The ability to measure directly homonuclear and heteronuclear dipole-dipole couplings, and thus obtain highly accurate structural information, is one of the most appealing features of

solid-state NMR spectroscopy. Yet, its inherently low resolution has posed a significant challenge for the advancement of the field. In modern NMR spectroscopy of solids, high-resolution spectra are typically obtained by spinning the sample rapidly in a rotor oriented at the angle $\arccos(1/\sqrt{3})$, a process call magic-angle spinning (MAS). The effect of such spinning is the averaging out of the anisotropic interactions (of the second-rank tensor form) that cause line broadening in static samples. The result of MAS is, therefore, high-resolution SSNMR spectra, presenting narrow resonance signals at the isotropic chemical shift positions, resembling solution NMR spectra. However, the broadening interactions that are eliminated by MAS, such as the (internuclear distance-dependent) dipolar interaction, contain rich structural information and thus need to be reintroduced by interfering with the MAS averaging mechanism.

The homonuclear dipolar interaction can be reintroduced by either mechanical methods or radio frequency (rf) irradiation methods. For a pair of spins, mechanical reintroduction of their dipole-dipole coupling is achieved through the rotational resonance phenomenon, in which an integer multiple of the rotor frequency is selected to match the isotropic chemical shift separation (in Hertz) of the two spins to be recoupled. (Raleigh et al., Chem. Phys. Lett. 146) When the spinning frequency meets the rotational resonance condition for a given pair of spins, the system's dipolar Hamiltonian term survives the averaging effect of MAS. Evidence of the recoupling effect of rotational resonance can be observed in the rapid exchange of Zeeman magnetization between the pair of spins involved, which can be understood as an excitation of the flip-flop (zero-quantum) term of the interaction Hamiltonian. As Figure 1a shows, no magnetization exchange occurs while away from rotational resonance. On the other hand, when the spinning frequency satisfies the rotational resonance condition for a pair of spins, the two will exchange magnetization while any other spins will be unaffected (Figure 1b). Clearly, rotational resonance is a frequency-selective narrow-band method since only those spin pairs whose isotropic chemical shift separations match the spinning frequency will be recoupled.

Longitudinal Magnetization Exchange

a) Off Rotational Resonace (R²)   c) Off R², with Mixing Pulses

b) On Rotational Resonace (R²)   d) On R², with Mixing Pulses

τ_m (ms) τ_m (ms)
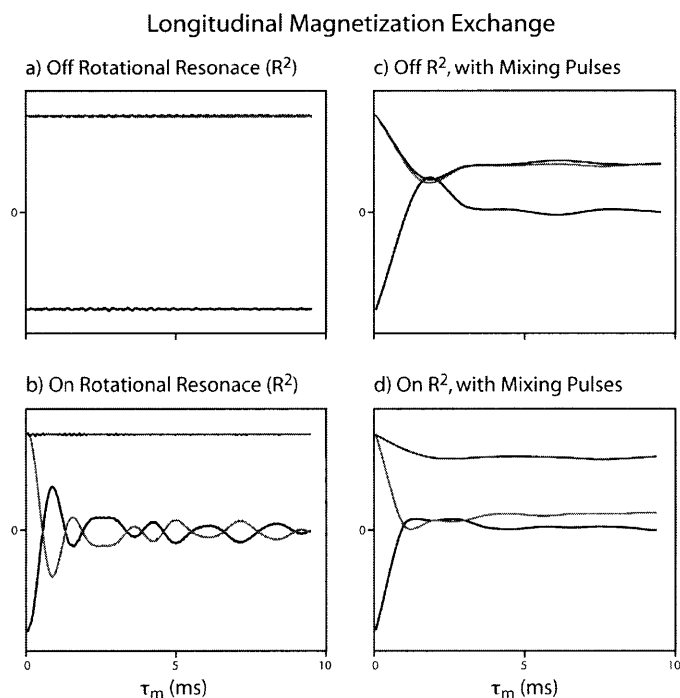
Figure 1. Simulations of polarization transfer between three strongly coupled spins. At time zero, one of the spins (black) starts inverted. As the mixing time, τ_m, is incremented, the spins, if recoupled, will exchange magnetization and come to equilibrium. (a) No exchange is observed when far from rotational resonance and in the absence of recoupling pulses. (b) When the spinning frequency matches the chemical shift separation between the inverted spin and one of the other two (black), they will be recoupled through the rotational resonance (R²) phenomenon and exchange their magnetization. As an alternative to R², rf pulses can be employed to recouple the dipole-dipole interaction. (c) Broadband mixing pulses cause magnetization exchange to occur between all coupled spins. (d) When mixing pulses are applied at R², the recoupling produced is a combination of both effects, with the spin pair at R² dominating the exchange.

The second broadly defined recoupling method is the application of rf pulses to interfere with the nuclear spin Hamiltonian. In this extremely versatile method, rf irradiation is used, during a period called the mixing time, to reintroduce specific terms in the dipolar Hamiltonian and produce an effective time-independent interaction that generates observable magnetization transfer. One of the first homonuclear recoupling techniques developed is called rf-driven dipolar recoupling (RFDR). The RFDR scheme consists of preparation of longitudinal magnetization, followed by a series of 180-degree (π) pulses applied once every rotor period throughout the mixing time. The effect of the π pulses is to reintroduce the zero-quantum term of the dipolar coupling Hamiltonian, which would otherwise remain eliminated by the averaging

49

effect of MAS. Using Average Hamiltonian Theory and the fictitious spin-1/2 operator formalism, it has been shown that the rotor-synchronized π pulses compensate for the chemical shift offset from rotational resonance, simultaneously for all spin pairs. Therefore, during the mixing period, the trajectories of the flip-flop fictitious spin operators for all the spin interactions in RFDR resemble those of individual spin pairs undergoing rotational resonance. Since RFDR recouples all dipolar interactions present, it is said to have broadband behavior. As Figures 1c and 1d illustrate, RFDR leads to longitudinal magnetization exchange between all coupled spins.

Since the first recoupling techniques were developed, many groups have worked on improving their efficiency on both selectively labeled samples and more complex, fully labeled ones. Today there exists a large number of sophisticated homonuclear recoupling pulse sequences, which, nonetheless, still face many challenges, leaving ample room for further improvements.

B.1.2 Modern homonuclear recoupling schemes and unsolved challenges

The abundant number of recoupling pulse sequences present today can be classified in a number of different ways. The most basic distinction can be made according to the coherence order recoupled by the differente schemes. For example, DRAMA (Dipolar recovery at the magic angle) and RFDR are zero-quantum (ZQ) pulse sequences, while MELODRAMA (Melding of spin-locking and DRAMA), HORROR (Homonuclear rotary resonance), DRAWS (Dipolar recoupling with a windowless sequence), and the symmetry-based schemes C7, POST-C7 (Permutationally offset stabilized C7), and SPC-5 (Supercycled POST-C5) are all double-quantum (DQ) pulse sequences.

Furthermore, pulse sequences can be classified based on their dependence on the Euler angles (α, β, γ) that describe crystallite orientations with respect to the rotor frame; where α describes molecular rotations around the internuclear vector, β describes the angle between the internuclear vector and the spinning axis, and γ describes the rotation of the internuclear vector around the spinning axis. For powder samples, the angles (α, β, γ) span the entire conformational space. The dipole-dipole coupling is inherently dependent on the angles β and γ, and is approximately independent of α. Therefore, the efficiency of dipolar recoupling techniques tends to show a strong dependence on the angles β and γ. The HORROR pulse scheme introduced the possibility of "γ-encoding", that is, elimination of the γ-angle

dependence. This reduced dependence on crystallite orientation features a lower extent of destructive geometrical averaging, thus improving the maximum theoretical recoupling efficiency from ~50% for γ-dependent sequences to ~73% for γ-independent ones. In addition to HORROR, among the pulse schemes listed above, C7, POST-C7, and SPC-5 are also γ-encoded.

Finally, a third classification can be made based on the frame of interaction used to best describe the recoupling process of a given pulse sequence. Schemes that use short rf pulses intercalated with long delays throughout the mixing time, such as RFDR, are referred to as "laboratory frame" pulse sequences, because the recoupling axis is (transversely) determined solely by the static magnetic field B_0 along the laboratory z-axis, around which the spin system evolves freely. On the other hand, "rotating frame" pulse sequences consist of mixing times in which an rf field is applied continuously, leading to a recoupling process in which the spin system evolves in the presence of a transverse rf field, thus redefining the recoupling axis. HORROR, DRAWS, C7, POST-C7, and SPC-5 are examples of rotating-frame recoupling sequences. This third classification is purely theoretical, but we stress it here in anticipation of the novel pulse schemes presented in this report. We have developed recoupling schemes in a new recoupling frame based on the phase modulation of continuous rf pulses. This new frame, dubbed the "modulation frame," is described in section 3.1 of this report.

B.1.2.1 Inherent challenges: Chemical shift effects and decoupling requirements

The evolution of recoupling techniques over the past decade has been motivated by the need to overcome difficulties that are inherent to most spin systems of interest, such as chemical shift offsets and chemical shift anisotropies and broadening effects of abundant spin species (i.e., ^{1}H) on the observed rare spins (e.g., ^{13}C). For instance, even though the first generation of rf-based recoupling schemes (DRAMA and RFDR), were shown to recover homonuclear dipolar couplings, their efficiencies are sensitive to isotropic chemical shift differences and the magnitude of the chemical shift anisotropy (CSA) of the interacting spins, resulting in reduced performance. Subsequently, pulse sequences such as MELODRAMA and DRAWS succeeded in eliminating these chemical shift dependence problems. However, being non γ-enconded, these schemes still faced a low theoretical efficiency (50% polarization loss after double quantum filtration). The concept of γ-encoding introduced by HORROR, as described in the previous section, presented an improved theoretical efficiency. The sequence C7 was the first to

apply γ-encoding to broadband recoupling, since the HORROR is limited to short bandwidths due to its low amplitude pulses. The supercycled version of C7, POST-C7, improved the pulse sequence's robustness to chemical shift offsets. Thus POST-C7 was able to produce efficient broadband homonuclear recoupling, and signified a great improvement in the technique.

In spite of its successes, POST-C7, a sevenfold symmetry sequence, requires significantly high rf power during its mixing period. More specifically, the rf power on the observe channel (e.g., ^{13}C) must match 7 times the MAS frequency. Also, during the mixing time, rf irradiation must be applied on the decouple channel (i.e., ^{1}H) in order to eliminate the deleterious heteronuclear interactions between the recoupled ^{13}C spins and the abundant protons. In order to avoid exciting ^{13}C-^{1}H cross-polarization, the rf power between the two channels must be mismatched by a ratio of about three. This means that during POST-C7's mixing period, the rf power on the decouple channel must match 21 times the MAS frequency. Thus, at 6 kHz MAS frequency, a relatively low spinning rate, POST-C7 requires 42 kHz and 126 kHz rf power simultaneously on the observe and decouple channels, respectively, in order to perform efficiently; while at 10 kHz MAS, a moderate spinning rate, those numbers become 70 kHz and 210 kHz rf power. Given that rf fields greater than 125 kHz applied for more than a few milliseconds (typical mixing times) are already potentially harmful to the instrument and to the sample, it is clear that POST-C7's high rf power demands limit its application to low MAS frequencies.

It was with this concern in mind that the sequence SPC-5 was developed. Since SPC-5 has five-fold symmetry, the recoupling pulses applied on the observe channel must match only 5 times the MAS frequency. Therefore, at a MAS of 10 kHz, for instance, SPC-5's rf power requirements for optimal efficiency are 50 kHz on the observe channel and 150 kHz on the decouple channel. Evidently, SPC-5 can be efficiently employed in experiments with moderate MAS rates, thus simultaneously recoupling the dipolar interactions while exploiting the improved resolution brought about by sufficiently fast MAS. The reduced rf power required by SPC-5 has enabled its successful application to the study of peptides and proteins via solid-state NMR. Subsequently, similar pulse sequences with symmetry numbers as low as 3.5 have been developed. Also, elaborate phase cycles (super cycles) have been introduced to compensate for the reduced bandwidth of low-symmetry sequences.

B.1.2.2 Technical challenges: High magnetic fields and very fast MAS

The employment of high magnetic fields is a straightforward way to obtain spectra with improved sensitivity and resolution. On the other hand, the increase B_0 field has the effect of expanding the range of chemical shift offsets and scaling up the magnitude of the CSA's. Also, very fast MAS probes have been developed and are effective in reducing broadening interactions, thus producing spectra with improved resolution. MAS frequencies of up to 25-30 kHz are now widely used to get high-resolution spectra of solids. For this reason, very fast MAS is complementary to utilizing high magnetic fields in order to obtain high-resolution solid-state NMR data. However, efficient homonuclear recoupling pulse sequences, such as SPC-5, cannot be applied at such extreme experimental conditions, since their rf power requirements are too high for safe operation, if at all possible to achieve. For instance, at 25 kHz MAS, the rf power requirements for the efficient SPC-5 recoupling are 125 kHz rf power on the observe channel concurrently with 375 kHz on the decouple channel, which is, first, impossible to achieve with current probe technology, and second, even if it were possible, it would certainly damage the sample, specially a biological one.

It is clear that there is a need for the development of new efficient recoupling methods that are compatible with the advances in solid-state NMR instrumentation. The combination of effective recoupling methods with high-resolution techniques is an important step in the further development of solid-state NMR as a tool for structural investigations of peptides and proteins.

B.1.3 Accuracy and limits of measurements

Finally, a few points must be made about distance measurements in solid-state NMR spectroscopy. The goal of recoupling methods is to extract an accurate estimate of the dipole-dipole coupling between interacting pairs of spins. Since the dipole-dipole coupling is a function of the internuclear distance, recoupling experiments in solid-state NMR have the capacity to extract structural information with very high accuracy, indeed, with much more accuracy than in solution NMR, for example. In a recent study about experimental accuracies in spin-pair labeled samples (Carravetta, JACS 123), certain symmetry-based homonuclear recoupling methods were found to have precisions of ± 0.02 Å, with systematic overestimates of the internuclear distance of 0.03 ± 0.01 Å. These accuracies are far superior to those that can be obtained with the two well-established structure determination methods, X-ray crystallography and solution NMR.

However, the often extreme experimental requirements of many efficient homonuclear recoupling techniques (as described in section 2.1.2.1) combined with difficulties in experimental analysis due to deleterious multi-spin effects, such as the Dipolar Truncation effect described in section 2.2, has caused these sophisticated techniques to be usually relegated in favor in much less accurate but also simpler techniques, such as proton-driven spin diffusion. Proton-driven spin diffusion is an incoherent process that transfers polarization among neighboring ^{13}C spins by means of the large coupled network formed by the abundant ^{1}H spins. The data obtained through spin diffusion experiments cannot be fit or modeled, and only very rough estimates (with errors of a few angstroms) can be made about internuclear distances. Still, many solid-state NMR researches today choose to rely on low-accuracy distance measurements with the premise that vast amounts of these measurements will lead to consistent data from which a reliable structure can be determined. These low-accuracy methods have proven to be able to lead to the determination of protein structures with the correct backbone fold, but their low resolution (evidenced, for example, in their poor side-chain distance constraints) is well below what solid-state NMR spectroscopy can potentially offer. Therefore, it is safe to say that the careful and accurate application of solid-state NMR to protein structure determination efforts remains contingent upon further developments in experimental methods, such as homonuclear recoupling schemes.

B.2 Systems of multiple spins

The second major portion of this report concerns the study of the effect of multiple-spin systems on the behavior and efficiency of solid-state NMR experiments. The vast majority of modern SSNMR techniques has been developed, and is well understood, only for systems consisting of two isolated spins, interacting only with each other. The extension of these experiments to larger systems, such as fully labeled peptides and proteins, has carried with it a great deal of inaccuracy and reduced efficiency. In this section, we describe how a system of weakly dipolar-coupled pair of spins is affected by the presence of a third spin, a phenomenon known as the dipolar truncation effect.
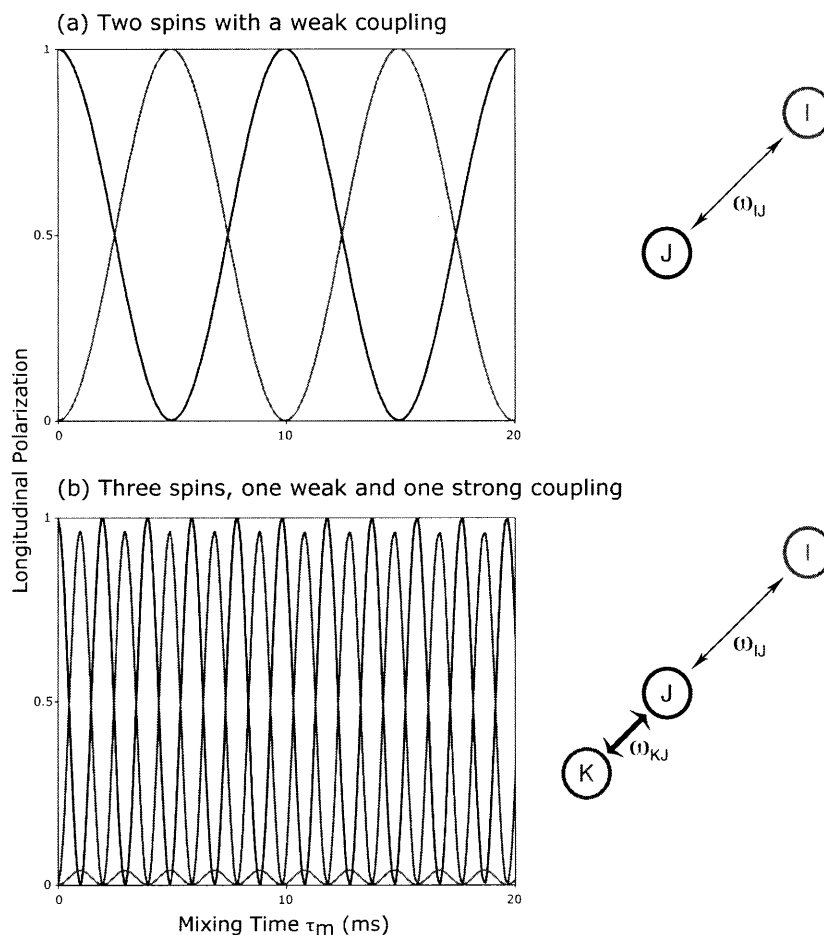
Figure 2. Single-crystallite numerical simulations of polarization exchange depicting the dipolar truncation effect. At time zero, the initial polarization is localized in a single spin, then, as the mixing time is incremented, the polarization will undergo an oscillatory exchange behavior between the coupled spins. The dynamics of this magnetization exchange is dictated by the nature of the dipolar couplings driving the evolution of the system. (a) In the case of two coupled spins (J and I), polarization is fully exchanged between the two sites, even if the coupling is weak. The strength of the coupling determines the frequency of the exchange cycle, in which polarization goes completely from one spin to the other and back again. The case of three spins, with two couplings is significantly more complicated. (b) When a third spin (K), close in space to one of the first two (J), is added to the system, the two couplings have very different magnitudes (in this case $c_{KJ} = 5c_{IJ}$). As a consequence, the polarization exchange dynamics is dominated by the stronger coupling, and the oscillatory exchange frequency of the entire three-spin system is determined solely by the stronger interaction. Furthermore, the maximum polarization accumulated by the distant is spin (I) is reduced by a factor given by the square of the ratio of the two couplings (in this case, 1/25).

B.2.1 The dipolar truncation effect

Let us consider a pair of spins (J and I) interacting under the effect of a weak dipole-dipole coupling c_{IJ}, a situation characteristic of two distant nuclei, in a single crystallite orientation. If the initial system consists of longitudinal magnetization on spin J only ($\rho(0) = J_z$) and the spin pair is allowed to evolve under a zero-quantum process, the polarization will be transferred back and forth between the two spins according to the equations:

$$< J_z > (t) = \cos^2(c_{IJ}t) = (1 + \cos(2c_{IJ}t))/2 \qquad (2.1)$$

$$< I_z > (t) = \sin^2(c_{IJ}t) = (1 - \cos(2c_{IJ}t))/2$$

As the numerical simulation in Figure 2a shows, there is complete polarization exchange between the two spins with a frequency of $2c_{IJ}$ (neglecting any relaxation effects).

Let us now add a third spin to the system: spin K interacting with spin J under the strong dipole-dipole coupling c_{KJ}, where $c_{KJ} = 5c_{IJ}$. For the moment, we will neglect any coupling between spins I and K. Starting again with $\rho(0) = J_z$ the polarization transfer equations are now:

$$< J_z > (t) = \cos^2(\omega t)$$

$$< I_z > (t) = \chi^2 \sin^2(\omega t) \qquad (2.2)$$

$$< K_z > (t) = (1 - \chi^2)\sin^2(\omega t)$$

where $\omega = \sqrt{c_{IJ}^2 + c_{KJ}^2}$, and $\chi = c_{IJ} / c_{KJ}$. The magnetization transfer behavior is significantly different from that in the two-spin case. The polarization transfer curves, shown in Figure 2b, present two important features that follow the form of equations 2.2 above. First, the maximum polarization transfer to spin I is about 4%, which can be understood by the predicted $<I_z>$ amplitude χ^2, which takes a value of $(1/5)^2 = 0.04$ for $c_{KJ} = 5c_{IJ}$. The second important feature is the fact that the oscillation frequency of the polarization exchange is determined primarily by the stronger coupling, c_{KJ}. Indeed, the magnetization exchange oscillation frequency for this case is about five times that of the two-spin case of Figure 2a. This behavior can be explained as the effect of an effective dipolar coupling $\omega = (c_{IJ}^2 + c_{KJ}^2)^{1/2} = (c_{KJ}^2 / 25 + c_{KJ}^2)^{1/2} \approx c_{KJ}$.

It is evident that, in the three-spin, two-couplings case, the stronger dipole-dipole coupling vastly dominates the polarization transfer dynamics of the entire three-spin system. It is also clear that the presence of the third, strongly coupled spin quenches the maximum possible

polarization transfer to the distant spin. These two observations constitute the phenomenon we refer to as the dipolar truncation effect.

B.2.2 Multi-spin effects and the study of biomolecules by SSNMR

The dipolar truncation effect has been a significant limitation in the study of fully labeled biological samples. Internuclear distances that are useful for generating structural constraints are in the range of 4-6 Å, which represent dipole-dipole couplings on the order of 100 Hz. The directly bonded spins in fully labeled samples carry dipole-dipole couplings of ~2.2 kHz, which quench polarization transfer to distant spins to about $(0.1 \text{ kHz} / 2.2 \text{ kHz})^2 = 0.2\%$ of their maximum possible value.

Recently, an alternating labeling technique has been developed and employed in the study of proteins via solid-state NMR. In this technique, proteins are expressed in media enriched with either [2-^{13}C] glycerol or [1,3-^{13}C] glycerol, leading to protein samples that have an alternating labeling pattern between adjacent carbons, with the two media producing complementary patterns. Even though this technique results in spectra with narrower lineshapes due to a reduced number of directly bonded carbons (and thus reduced broadening by scalar couplings), the deleterious effect of dipolar truncation may still be present. For instance, the dipolar interaction between a spin J and a spin I separated by 4.5 Å has a coupling constant of about 100 Hz. Then a spin K separated by two bonds from J has a coupling of about 500 Hz. As we noted in the preview section, this stronger coupling will quench the distant spin coupling by a factor of $(100/500)^2 = 1/25$, and it will dominate the polarization exchange dynamics, making it still difficult to measure the long distance accurately.

Alternating labeling schemes reduce the number of strong couplings but they do not eliminate dipolar truncation to the extent that long distances can be measured accurately. However, these labeling techniques might prove extremely useful when applied with novel pulse schemes (yet to be) designed to reduce dipolar truncation effects.

B.2.3 Circumventing the dipolar truncation problem: Rotational resonance approaches

The problems presented by dipolar truncation can be avoided to a large extent through the use of selective recoupling techniques. In particular, rotational resonance methods are highly selective in isotropic chemical shift differences, and they have yielded very accurate distance

measurements in isolated spin-pair samples. Furthermore, it has been shown that by selecting a specific pair of spins to recouple, the effects of strong couplings are reduced and long distances can be measured with a fair level of accuracy.

A variation of the basic rotational resonance experiment, denoted rotational resonance tickling (R^2T), is an appealing scheme for systems of many spins. In R^2T, the MAS frequency is set close to the rotational resonance condition for one or more pair of spins, and then a continuous weak rf pulse is ramped in amplitude. The rf pulse creates different effective fields for different resonances, allowing for passage through the R^2 condition for multiple pairs of spins at different points in the ramp. The MAS frequency and the initial and final amplitudes of the rf ramp determine which dipolar coupled spin pairs will go through R^2 and be recoupled. Thus R^2T can be used to recouple not just one pair of spins but also a series of them.

Obviously, R^2T is limited to specific chemical shift ranges, and it is mostly applied to recouple interactions between carbonyl and aliphatic carbons. Also, even though R^2T can be used to select for weak couplings (based on their chemical shift separations), the effects of strong couplings near R^2 are not fully suppressed, and they complicate the recoupling dynamics. However, R^2T is one of the most promising options currently available to deal with dipolar truncation effects.

B.3 Protein resonance assignments

Although an iterative task, the process of protein structure determination via NMR spectroscopy can be divided into three main steps: resonance assignment, distance and torsion angle constraint generation, and structure calculation. The third and final part of this proposal describes the application of solid-state NMR methodology to obtain backbone and side-chain resonance assignments of gas vesicle protein A, a 76-residue protein for which there is not a known structure yet. This protein sample is an example of a system that is not amenable to structural investigations using solution NMR or X-ray crystallography, since it is highly insoluble and it does not form diffraction-quality crystals. It therefore constitutes a challenging system with the potential of demonstrating the power of solid-state NMR to offer structural insight where conventional methods cannot be applied.

B.3.1 Resonance assignments in solid-state NMR

Significant efforts have been made in recent years to develop methods for carrying out resonance assignments in fully labeled peptides and proteins via solid-state NMR. The first completed site-specific resonance assignments in proteins have proven the feasibility of the technique, but they have also shown that the process is far from routine and that further improvements would be of great benefit.

Besides being the first step in structure determination via NMR spectroscopy, site-specific resonance assignments offer insights into chemical and conformational information. For instance, chemical shifts and correlation spectra can provide valuable information about secondary structure, hydrogen bonding, and protonation states.

B.3.2 Commonly used strategies

The typical experiments used to assign backbone and sidechain resonances consist of homonuclear (^{13}C-^{13}C) and heteronuclear (^{15}N-^{13}C) correlations. ^{13}C-^{13}C correlations can be used to assign intraresidue resonances, while ^{15}N-^{13}C correlations (in its specific cross polarization forms NCA and NCO) can be used to perform sequential assignments. Additionally, NCACX and NCOCX experiments can be employed to obtain and corroborate assignments. 2D NCACX (NCOCX) experiments have ^{15}N$_i$ in one dimension and CA$_i$ (CO$_{i-1}$) and its correlated ^{13}C resonances in the other dimension. In 3D NCACX (NCOCX) experiments, resolution is improved by performing ^{13}C-^{13}C correlations in two dimensions, and evolving on ^{15}N in the third dimension, producing 2D ^{13}C-^{13}C planes at the associated ^{15}N chemical shift.

Part I

Development of Magic-Angle Spinning NMR Methodology for the Analysis of Protein Structure

Chapter 2. Heteronuclear Interference and Homonuclear Mixing without Decoupling in High-Resolution MAS NMR Spectroscopy

Summary

  We describe solid-state NMR homonuclear recoupling experiments at high magic-angle spinning (MAS) frequencies using the radio frequency-driven recoupling (RFDR) scheme. The effect of heteronuclear decoupling interference during RFDR recoupling at high spinning frequencies is investigated experimentally and via numerical simulations, resulting in the identification of optimal decoupling conditions. The effects of MAS frequency, RF field amplitude, bandwidth, and chemical shift offsets are examined. Most significantly, it is shown that broadband homonuclear correlation spectra can be efficiently obtained using RFDR without decoupling during the mixing period in fully protonated samples. The utility of RFDR *sans* decoupling is demonstrated with broadband correlation spectra of a peptide and a model protein at high MAS frequencies and high magnetic field.

2.1 Introduction

Solid-state NMR has emerged as a powerful and versatile technique for the study of biological molecules in large part due to advances in high-resolution techniques such as high-frequency magic-angle spinning (MAS). In recent work,[1,2] solid-state NMR spectra of proteins recorded at MAS frequencies of 20 kHz and above have shown considerably improved resolution when compared to spectra observed at moderate spinning rates (~10 kHz). Concurrently, high-frequency MAS rotors hold small amounts of sample; therefore, the beneficial application of high MAS rates has been contingent on parallel advances in spectroscopic methods for improving signal intensities, such as heteronuclear decoupling and dipolar recoupling. Various heteronuclear decoupling schemes[3-14] have been shown to perform very efficiently under rapid spinning. On the other hand, the investigation of dipolar recouplng techniques compatible with high spinning frequencies, has only recently begun to show progress.[15-19]

One of the main limitations in the application of well-established homonuclear dipolar recoupling pulse sequences, such as symmetry-based schemes,[20-27] is the typical requirement of high-power irradiation, which scales as a multiple of the spinning frequency ($\omega_r/2\pi$). Moreover, in most pulse sequences, the ^1H decoupling irradiation during the mixing period must exceed the recoupling field by a factor of three or more in order to avoid interference effects.[25,28,29] In recent studies, certain symmetry-based recoupling schemes have been found to preserve some efficiency in the absence of ^1H decoupling irradiation, thus allowing their application at moderate MAS frequencies ($\omega_r/2\pi \leq 20$ kHz).[17,18] At very high spinning frequencies (above 40 kHz), the DREAM[30,31] scheme has been shown to perform well in the absence of decoupling,[16] yet its narrow bandwidth remains a practical limitation at high magnetic fields. Finally, a new approach to recoupling at high MAS frequencies, introduced with the pulse sequence CMAR,[19] is to simultaneously recouple homonuclear interactions and eliminate heteronuclear couplings with a single-channel irradiation, thereby reducing power requirements. This double-quantum, cosine-modulated pulse sequence strategy is currently under further development in our laboratory.[32]

The π-pulse based homonuclear dipolar recoupling scheme radio frequency-driven recoupling[29,33] (RFDR), which emerged in the early 1990's,[33-36] has been used in a variety of cases to record correlation spectra of small molecules and proteins[29,37-40] and has given rise to multiple applications and theoretical developments.[41-44] More recently, RFDR has been shown to

64

function well at high spinning frequencies[45] and in the absence of ^{1}H decoupling with non-protonated ^{13}C nuclei[15] and with a heavily deuterated protein.[46] However, the effect of ^{1}H decoupling interference during RFDR mixing at high MAS frequencies has not been described in detail. In this chapter, we investigate the effect of proton decoupling during RFDR at high spinning frequencies ($\omega_r/2\pi \geq 20$ kHz) and observe that RFDR can function efficiently both with carefully calibrated high-power decoupling and in the absence of decoupling, even for protonated ^{13}C nuclei with large chemical shift offsets. We show that, at high spinning rates, the ^{1}H decoupling field creates narrow depolarization conditions that depend on the strength of the ^{13}C RFDR pulses and the MAS frequency, and thus must be calibrated carefully for optimal operation. In addition, we demonstrate that RFDR at high spinning frequencies can achieve high broadband efficiency in the complete absence of decoupling irradiation and that its efficiency improves with increasing ^{13}C RF field strengths. The observations presented here describe the optimal implementation of RFDR at high MAS frequencies and demonstrate, with ^{13}C-^{13}C correlation spectra of a peptide and a model protein, that it is a powerful scheme for high-resolution structural investigations of biomolecules.

2.2 Theory

A. The recoupled Hamiltonian of a heteronuclear spin system

The Hamiltonian of a system comprising of ^{1}H and ^{13}C spins in a magic-angle spinning experiment is represented by

$$H(t) = H_{csa}(t) + H_{CC}(t) + H_{HH}(t) + H_{CH}(t) + H_{RF}(t) \qquad (1)$$

where the chemical shift and radio-frequency Hamiltonians are represented by single spin operators, while the dipolar interactions (both homonuclear $H_{CC}(t)$, $H_{HH}(t)$ and heteronuclear $H_{CH}(t)$) are represented by two spin operators:

$$H_{csa}(t) = \sum_{i=1}^{N} \omega_{C,i}(t) S_{iz} + \sum_{i=1}^{M} \omega_{H,i}(t) I_{iz} \qquad (2a)$$

$$H_{CC}(t) = \sum_{\substack{i,j=1, \\ i \neq j}}^{N} \omega_{CC,ij}(t) \left[2S_{iz}S_{jz} - \tfrac{1}{2}\left(S_i^+ S_j^- + S_i^- S_j^+ \right) \right] \qquad (2b)$$

$$H_{HH}(t) = \sum_{\substack{i,j=1, \\ i \neq j}}^{M} \omega_{HH,ij}(t) \left[2I_{iz}I_{jz} - \tfrac{1}{2}\left(I_i^+ I_j^- + I_i^- I_j^+ \right) \right] \qquad (2c)$$

$$H_{CH}(t) = \sum_{i=1}^{N} \sum_{j=1}^{M} \omega_{CH,ij}(t) 2S_{iz}I_{jz} \qquad (2d)$$

For the sake of convenience, the above spin Hamiltonian is often described in the rotating frame, wherein the time-dependent components of the RF field are neglected under the secular approximation. To describe the effects of MAS and RF irradiation on the performance of recoupling experiments, we employ a model system comprising of four spins, two ^{13}C and two ^{1}H, in our treatment. Furthermore, the spin Hamiltonian is expressed in terms of irreducible spherical tensor operators[47-50] to elucidate the intricacies in such experiments. Expressing the spin Hamiltonian of Eq. 1 in terms of irreducible spherical tensor operators leads to the following rotating frame Hamiltonian for our four-spin model system,

$$H'(t) = \sum_{i,j=1}^{2} \sum_{\substack{m=-2,\\m\neq0}}^{2} \omega_{C_i}^{(m)} T^{(1)0}(C_i) e^{im\omega_r t} + \omega_{H_j}^{(m)} T^{(1)0}(H_j) e^{im\omega_r t} + \sum_{\substack{m=-2,\\m\neq0}}^{2} \omega_{CC}^{(m)}(t) T^{(2)0}(C_1 C_2) e^{im\omega_r t}$$

$$+ \sum_{\substack{m=-2,\\m\neq0}}^{2} \omega_{HH}^{(m)}(t) T^{(2)0}(H_1 H_2) e^{im\omega_r t} + 2\sum_{i,j=1}^{2} \sum_{\substack{m=-2,\\m\neq0}}^{2} \omega_{C_i H_j}^{(m)}(t) T^{(1)0}(C_i) T^{(1)0}(H_j) e^{im\omega_r t} \qquad (3)$$

$$+ \omega_{1C}\left(S_{1x} + S_{2x}\right) + \omega_{1H}\left(I_{1x} + I_{2x}\right)$$

where the single spin and two-spin operators are represented by $T^{(k)q}(X_i)$ and $T^{(k)q}(X_i X_j)$, and ω_{1C} and ω_{1H} represent the RF fields employed on the ^{13}C and ^{1}H channel, respectively. To describe the simultaneous effects of RF irradiation and sample spinning on the spin system, the Hamiltonian in Eq. 3 is further transformed into the RF interaction frame defined by the transformation operator $U = \exp(i\omega_{1C}S_{ix}t)\exp(i\omega_{1H}I_{ix}t)$, yielding,

$$H''(t) = \sum_{i,j=1}^{2} \sum_{q=-1}^{1} \sum_{\substack{m=-2,\\m\neq0}}^{2} \omega_{C_i,m}^{(1)q} T^{(1)q}(C_i) e^{im\omega_r t} e^{iq\omega_{rf,C}t} + \omega_{H_j,m}^{(1)q} T^{(1)q}(H_j) e^{im\omega_r t} e^{iq\omega_{rf,H}t}$$

$$+ \sum_{\substack{m=-2,\\m\neq0}}^{2} \sum_{q=-2}^{2} \omega_{CC,m}^{(2)q}(t) T^{(2)q}(C_1 C_2) e^{im\omega_r t} e^{iq\omega_{1C}t} + \sum_{\substack{m=-2,\\m\neq0}}^{2} \sum_{q=-2}^{2} \omega_{HH,m}^{(2)q}(t) T^{(2)q}(H_1 H_2) e^{im\omega_r t} e^{iq\omega_{1H}t} \qquad (4)$$

$$+ 2\sum_{i,j=1}^{2} \sum_{q=-1}^{1} \sum_{\substack{m=-2,\\m\neq0}}^{2} \omega_{C_i H_j,m}(t) T^{(1)q}(C_i) T^{(1)q}(H_j) e^{im\omega_r t} e^{iq\omega_{1C}t} e^{iq\omega_{1H}t}$$

In this interaction frame, the spin Hamiltonian exhibits a complicated dependence due to MAS and RF fields besides involving operators with different rank and coherence orders. Depending on the nature of the recoupling experiment, the experimental RF field strengths and the spinning frequency are suitably adjusted to result in a Hamiltonian that is time-independent, often referred to as the recoupled Hamiltonian in solid-state NMR. In dipolar recoupling experiments the dipolar interactions among the nuclear spins are reintroduced either selectively or in a broadband fashion by adjusting the sample spinning frequency and the RF field strengths carefully such that the chemical shift anisotropy (CSA) and heteronuclear dipolar interactions are suppressed in the experiment. The optimum conditions for a particular recoupling experiment can be deduced sequentially from the model four spin system described above within the framework of average Hamiltonian theory. To simplify the description of the spin dynamics the RF fields on both the ^{13}C and ^{1}H channels are synchronized with the sample spinning frequency.

Employing the Magnus expansion formula,[51] the first and second order corrections to the effective Hamiltonian are calculated in the usual manner,

$$H^{(1)} = \frac{1}{\tau_r} \int_0^{\tau_r} H''(t)dt$$

$$H^{(2)} = -\frac{i}{2\hbar\tau_r} \int_0^{\tau_r} dt_2 \int_0^{t_2} dt_1 \left[H''(t_1), H''(t_2) \right]$$

(5)

Depending on the recoupling conditions, the effective Hamiltonian describing the spin dynamics may vary from a typical zero-quantum (ZQ) or double-quantum (DQ) Hamiltonian. In either case the effective subspace that describes the overall spin physics in the Liouville space involves only the zero-quantum operators along with the recoupled first order (ZQ or DQ) Hamiltonian. Hence, to describe the performance of a particular recoupling sequence we need to calculate all the higher order terms, which contribute to the ZQ operators in the effective Liouville subspace.[49,50] Although, we do not make a quantitative description of the underlying spin dynamics in the Liouville space, the general philosophy involved in the design of such experiments can be deduced from the various terms present in the effective Hamiltonians described in the next section. In this article we confine our analysis only to the second order terms in the Magnus expansion. Although, the description provided below predicts similar trends in all homonuclear recoupling pulse sequences, specific schemes need to be investigated individually to elucidate their exact behavior and will be described in forthcoming publications.

B. Contributions from first order terms

In an ideal dipolar recoupling experiment, only the dipolar interactions are present to first order. All other effects representative of the CSA and heteronuclear dipolar interactions are effectively averaged to first order by careful manipulation of the spinning frequency and RF fields on both channels. The desired and undesired conditions in a particular experiment can be deduced from the interaction Hamiltonian represented in Eq. 3. For example, when the effective field on the carbon channel is matched to the spinning frequency, i.e. $m\omega_r = \omega_{1C}$, the CSA interactions of the ^{13}C spins are reintroduced to first order. Hence, to avoid undesirable rotary resonance conditions,[52-55] the RF field strengths employed on the ^{13}C channel are often chosen to be greater than the sample spinning frequency i.e. $\omega_{1C} > 2\omega_r$. In a similar vein, the heteronuclear dipolar interactions (say ^{13}C-^1H) also have matching conditions that need to be avoided while

68

recoupling the homonuclear dipolar interactions among the ^{13}C spins. For example when the sum of the effective fields (RF fields) on both channels is matched to an integer multiple of the spinning frequency ($m\omega_r = \omega_{1C} + \omega_{1H}$) the heteronuclear dipolar interactions involving DQ operators, $T^{(2)\pm2}(CH)$, are reintroduced to first order in the interaction frame. The single-quantum operators involving heteronuclear dipolar interactions, $T^{(k)\pm1}(CH)$, are reintroduced when the RF field strengths on either of the channels is matched to an integer multiple of the spinning frequency, i.e. $m\omega_r = \omega_{1C}$ or $m\omega_r = \omega_{1H}$. Finally, the heteronuclear dipolar interactions involving ZQ operators ($T^{(k)0}(CH)$, $k = 0,1,2$) are reintroduced when the difference in the effective fields is matched to an integer multiple of the spinning frequency, $m\omega_r = |\omega_{1C} - \omega_{1H}|$. Hence, dipolar recoupling experiments are carefully designed and implemented so as to avoid such conditions.

The conditions described above, based on first order average Hamitonian theory, can be readily implemented in experiments at low to moderate spinning frequencies by employing intense RF fields so as to avoid strong matching conditions, that is, $\omega_1 > m\omega_r$ with $m = 3$ or 4. However, at high spinning frequencies it is difficult to avoid such conditions since the ω_r dependence requires prohibitively intense RF fields. Moreover, the overall efficiency of a sequence depends on the compensation of the higher order terms (second order in particular) involved in the Magnus expansion. In particular the ZQ single-spin, two-spin, and three-spin operators resulting from second order corrections are primarily responsible for the depolarization observed in both zero-quantum and double-quantum recoupling experiments in MAS solid-state NMR.

C. Contributions from second order terms

Single-spin longitudinal operators

The longitudinal operators corresponding to the carbon spins, $T^{(1)0}(C_i)$, result from cross-terms among the single-spin and two-spin interactions[49,50] and are summarized by the following set of equations.

(a) CSA × CSA

$$\left[T^{(1)q}(C_i), T^{(1)-q}(C_i)\right] \simeq \omega_{C_i,m}^{(1)q} \omega_{C_i,-m}^{(1)-q} \frac{1}{\left(m\omega_r \pm \omega_{1C}\right)} T^{(1)0}(C_i)$$ (6a)

(b) Heteronuclear × Heteronuclear dipolar interactions

$$\left[T^{(2)2}(C_iH_j), T^{(2)-2}(C_iH_j)\right] \simeq \omega_{C_iH_j,m}^{(2)2} \omega_{C_iH_j,-m}^{(2)-2} \frac{1}{m\omega_r \pm \left(\omega_{1C} + \omega_{1H}\right)} T^{(1)0}(C_i)$$ (6b)

$$\left[T^{(k)q}(C_iH_j), T^{(k')-q}(C_iH_j)\right] \simeq \omega_{C_iH_j,m}^{(k)q} \omega_{C_iH_j,-m}^{(k')-q} \frac{1}{\left(m\omega_r \pm \omega_{1H}\right)} T^{(1)0}(C_i)$$

$$\omega_{C_iH_j,m}^{(k)q} \omega_{C_iH_j,-m}^{(k')-q} \frac{1}{\left(m\omega_r \pm \omega_{1C}\right)} T^{(1)0}(C_i)$$ (6c)

$$\omega_{C_iH_j,m}^{(k)q} \omega_{C_iH_j,-m}^{(k')-q} \frac{1}{\left(m\omega_r \pm \left|\omega_{1C} - \omega_{1H}\right|\right)} T^{(1)0}(C_i)$$

(c) Homonuclear × Homonuclear dipolar interactions

$$\left[T^{(2)q}(C_iC_j), T^{(2)-q}(C_iC_j)\right] \simeq \omega_{C_iC_j,,m}^{(2)q} \omega_{C_iC_j,,-m}^{(2)-q} \frac{1}{\left(m\omega_r + q\omega_{1C}\right)} T^{(1)0}(C_i)$$ (6d)

The contributions arising from these terms can be minimized significantly by increasing the magnitude of the terms present in the denominator, which must be selected taking into account first-order heteronuclear matching conditions.

Two-spin zero-quantum operators

These operators result from cross-terms between the two-spin and single-spin operators in the second order Magnus expansion formula. For example, the cross-terms between the ^{13}C homonuclear dipolar interaction and CSA interactions result in two-spin homonuclear operators. Similarly the cross-terms between the heteronuclear dipolar interactions and CSA interactions result in two-spin heteronuclear operators. The possible combinations involving a single-spin and two-spin operators are summarized below.

(a) Homonuclear interactions $\times$ CSA interactions

$$\left[T^{(2)q}(C_iC_j), T^{(1)-q}(C_i)\right] \simeq \omega_{C_iC_j,m}^{(2)q}\omega_{C_i,-m}^{(1)-q}\frac{1}{\left(m\omega_r + q\omega_{1C}\right)}T^{(k)0}(C_iC_j) \tag{7a}$$

(b) Heteronuclear interactions $\times$ CSA interactions

$$\left[T^{(2)q}(C_iH_j), T^{(1)-q}(C_i)\right] \simeq \omega_{C_iH_j,m}^{(2)q}\omega_{C_i,-m}^{(1)-q}\frac{1}{\left(m\omega_r \pm \omega_{1C}\right)}T^{(k)0}(C_iH_j)$$

$$\omega_{C_iH_j,m}^{(2)q}\omega_{C_i,-m}^{(1)-q}\frac{1}{\left(m\omega_r \pm \omega_{1H}\right)}T^{(k)0}(C_iH_j) \tag{7b}$$

Three-spin zero-quantum operators

The three-spin interactions arise only from cross-terms between two-spin interactions that are only present in second and higher order terms in the Magnus expansion. The following equations summarize the various second order contributions resulting in three-spin operators.

(a) Heteronuclear $\times$ Heteronuclear dipolar interactions

$$\left[T^{(k)q}(C_iH_j), T^{(k')-q}(C_kH_j)\right] \simeq \omega_{C_iH_j,m}^{(k)q}\omega_{C_kH_j,-m}^{(k')-q}\frac{1}{\left(m\omega_r \pm \left|\omega_{1C} - \omega_{1H}\right|\right)}T^{(k)0}(C_iC_kH_j) \tag{8a}$$

(b) Heteronuclear $\times$ Homonuclear dipolar interactions (protons)

$$\left[T^{(k)q}(C_iH_j), T^{(k')-q}(H_kH_j)\right] \simeq \omega_{C_iH_j,m}^{(k)q}\omega_{H_kH_j,-m}^{(k')-q}\frac{1}{\left(m\omega_r \pm \omega_{1H}\right)}T^{(k)0}(C_iH_kH_j) \tag{8b}$$

(c) Heteronuclear $\times$ Homonuclear dipolar interactions (carbons)

$$\left[T^{(k)q}(C_iH_j), T^{(k')-q}(C_iC_k)\right] \simeq \omega_{C_iH_j,m}^{(k)q}\omega_{C_iC_k,-m}^{(k')-q}\frac{1}{\left(m\omega_r \pm \omega_{1C}\right)}T^{(k)0}(C_iC_kH_j) \tag{8c}$$

Since the magnitude of the heteronuclear dipolar interactions are much stronger than the CSA interactions, compensation of the second order terms involving heteronuclear dipolar interactions is essential for improving the overall efficiency of a particular recoupling sequence. In the slow spinning regime (say $\omega_r/2\pi \leq 10$ kHz), the RF requirements on both channels are less demanding ($\omega_{1C}/2\pi \leq 50$ kHz, $\omega_{1H}/2\pi \leq 100$ kHz) and hence can be implemented with commercial NMR probes. In this regime many of the existing recoupling pulse sequences perform efficiently due to the compensation of the second order terms involving both CSA and

71

heteronuclear dipolar interactions. On the other hand, in the intermediate ($\omega_r/2\pi = 10 - 20\,\text{kHz}$) and fast ($\omega_r/2\pi > 20\,\text{kHz}$) spinning regimes, the performance of many rotor-synchronized recoupling sequences begins to deteriorate in spite of the improved averaging provided by higher spinning frequencies. This behavior can be explained from the second order contributions arising from single-spin, two-spin and three-spin zero-quantum operators. For example, in order to minimize the effect of the heteronuclear dipolar interactions resulting in single spin (Eq. 6c) and three-spin zero-quantum operators (Eq. 8a) the difference in the effective fields on the two RF channels must be greater than the spinning frequency, i.e. $|\omega_{1C} - \omega_{1H}| > 2\omega_r$, while the ^{13}C RF field itself must be greater than the spinning frequency (say, $\omega_{1C} > 2\omega_r$) to effectively compensate for CSA interactions, although in the case of RFDR, these interactions are averaged to zero by phase alternation of the recouping pulses. This condition is difficult to maintain since it requires the simultaneous application of high RF powers on both channels, which commercial NMR probes do not tolerate well. In addition, intense RF irradiation may lead to excessive sample heating, which is of particular concern in biological applications. As a result, the applied RF fields are limited to a regime where they fail to fully compensate for second-order heteronuclear dipolar interactions in homonuclear dipolar recoupling experiments at high MAS frequencies, leading to a decrease in efficiency.

Alternatively, employing high MAS frequencies and high RF fields on the ^{13}C channel without any decoupling irradiation on the ^{1}H channel could in principle minimize the second-order terms involving heteronuclear dipolar interactions. This can be readily identified in the second order terms presented in Eqs. 7 and 8. Such an approach alleviates the demanding requirement of high RF powers on both channels, leading to an experimental scheme suitable for high spinning frequencies and application at high magnetic fields.

2.3 Experimental Methods

A. Pulse sequence

The RFDR recoupling pulse sequence[29,33,34] consists of rotor-synchronized π pulses placed at the middle of each rotor period during a mixing period following the preparation of longitudinal magnetization. The phases of consecutive π pulses are set to follow the XY-4, XY-8, or XY-16 compensated pulse schemes,[29,56,57] with the basic block comprising two π pulses. Figure 1 shows the usual implementation of RFDR in the mixing period of a homonuclear correlation two-dimensional (2D) experiment. In the limit of short ^{13}C π pulses and at low to moderate spinning rates (and thus long window periods without ^{13}C irradiation), optimal heteronuclear decoupling is achieved using the TPPM scheme[3] during the windows, as in the case of rotational-echo double resonance experiments.[58] During the π pulses, strong cw decoupling irradiation is necessary in order to avoid first-order heteronuclear depolarization conditions, in particular Hartmann-Hahn matching between the ^{1}H and ^{13}C RF fields.[28] Indeed, this is an essential experimental consideration since it has been demonstrated that, for low to moderate MAS frequencies, the decoupling ^{1}H RF field during the recoupling π pulses needs to exceed the ^{13}C RF field by a factor of three or more to fully avoid depolarization.[28,29] At higher MAS frequencies (above 10 kHz), more complicated depolarization patterns arise, as illustrated by our results described below.
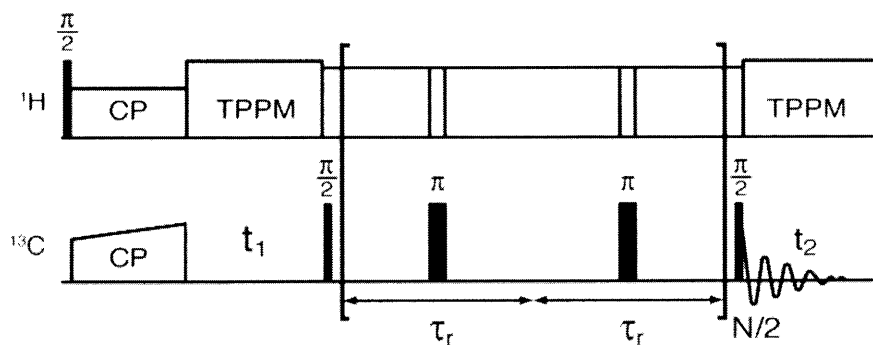
Figure 1. RFDR pulse sequence with t_1 evolution period. During the mixing time, one ^{13}C π pulse is applied per rotor period, while ^{1}H decoupling irradiation is applied simultaneously. At low to moderate spinning frequencies, the decoupling scheme during the RFDR mixing period typically consists of strong CW ^{1}H fields during the recoupling π pulses and moderate CW or TPPM decoupling during the windows between the π pulses. The basic building block of RFDR consists of two rotor periods.

B. Samples and instruments

In order to investigate the performance of RFDR at high MAS frequencies, we first examined the heteronuclear depolarization phenomena for a single ^{13}C spin during a train of ^{13}C π pulses at different MAS frequencies. These studies were performed on a sample of [2-^{13}C, ^{15}N]glycine diluted to 10% in natural abundance glycine, using a custom-designed MAS probe with a 3.2 mm Revolution NMR stator on a custom-designed spectrometer operating at 400 MHz ^{1}H frequency. We then explored the dependence of heteronuclear interference effects on varying decoupling fields, recoupling RF fields, and frequency offsets at high MAS frequencies and high magnetic field. These data sets were collected on the tripeptide N-formyl-[U-^{13}C, ^{15}N] Met-Leu-Phe-OH (N-f-MLF-OH)[59] and the microcrystalline model protein beta-1 immunoglobulin binding domain of protein G (GB1, prepared according to recent literature[60]), using a Bruker 2.5 mm MAS probe on a custom-designed spectrometer operating at 750 MHz ^{1}H frequency. Two-dimensional correlation spectra of these two samples demonstrate the efficient polarization transfer attainable with RFDR mixing at high spinning frequencies. For experiments at high spinning frequency, a 2.5 mm rotor was fully packed and the probe utilized presented high RF homogeneity (~90 % intensity ratio between 450 and 90 degree pulses), which is important for the observation of narrow matching conditions.

2.4 Results and Discussion

A. The effect of the decoupling field during RFDR π pulses at increasing MAS frequencies

As delineated in the theory section, inadequate decoupling fields can lead to depolarization during the RFDR pulse train due to heteronuclear interference between concurrent the ^{1}H and ^{13}C fields. As the theory predicts, heteronuclear interference has a strong dependence on the spinning frequency. Figure 2 demonstrates the process of ^{13}C magnetization loss due to ^{1}H-^{13}C interference during RFDR at increasing MAS frequencies. In these experiments, the ^{13}C magnetization of [2-^{13}C, ^{15}N] glycine remaining after 1.6 ms of RFDR pulsing was monitored as a function of the ^{1}H decoupling field during the ^{13}C π pulses, while maintaining the ^{1}H decoupling field during the windows at a fixed value. In particular, the RF field of the ^{13}C π pulses was 33.3 kHz, and the ^{1}H decoupling field during the windows between π pulses was 85 kHz (a decoupling field sufficiently strong to suppress proton-driven spin diffusion regardless of the spinning frequency), while the ^{1}H decoupling field during the π pulses was varied from 0 to 125 kHz. In order to minimize the effects of relaxation and ^{13}C pulse errors, a short RFDR period of 1.6 ms was employed.

The interference effect of the decoupling field during RFDR pulses at 5 kHz MAS was examined in detail by Bennett et al,[29] using similar magnetization monitoring experiments. A significant difference in their implementation is that a very strong decoupling field (150 kHz) was applied during the windows between pulses, with the intention of isolating the source of depolarization to the interference effect between the simultaneous ^{13}C and ^{1}H fields during the π pulses, in addition to relaxation and ^{13}C pulse imperfections. On the other hand, in the experiments presented in Figure 2, the decoupling field during the windows was set to a ^{1}H nutation frequency of 85 kHz in order to examine experimental conditions typically employed when working with biological samples. As a result of this relatively moderate decoupling field, the observed magnetization losses can be expected to arise from depolarization during both the π pulses and the windows between them. However, despite this additional complexity, the magnetization profile observed at the moderate spinning rate of 10 kHz (Figure 2a) is similar to that observed by Bennett et al[29] at 5 kHz in that it follows approximately a smooth sigmoidal shape and that a mismatch ratio of 3:1 between ^{1}H and ^{13}C RF fields, reached in our case at 100 kHz ^{1}H field, efficiently attenuates heteronuclear interference and preserves most (~95 %) of the magnetization.

In contrast, at higher MAS frequencies the magnetization profile presents significant changes, as illustrated in Figures 2b and 2c. At $\omega_r/2\pi = 15$ kHz, the effect of increasing the decoupling field is no longer a smooth increase of the magnetization (due to a decreasing extent of depolarization), but instead we observe plateaus around 50 and 80 kHz. More importantly, the 3:1 mismatch ratio is less effective in preserving the magnetization than at spinning frequencies of 10 kHz and below. Further increasing the MAS frequency to 20 kHz yields heteronuclear interference conditions that are dispersed and sufficiently narrow that they create local minima in the magnetization profile, and a 3:1 mismatch ratio between ^{1}H and ^{13}C RF fields only preserves 80 % of the initial magnetization.

The observation of rather narrow depolarization conditions at high spinning rates that persist at strong decoupling levels can be expected from first-order heteronuclear interference terms.[28,29] At moderate MAS rates, the strong ^{1}H-^{13}C and ^{1}H-^{1}H couplings (~20 kHz) yield broad heteronuclear interference conditions that can be attenuated with a strong decoupling field, relative to the ^{13}C RF field. As the MAS frequency is raised, however, it becomes more efficient at averaging the dipolar interactions, leading to narrower first-order matching conditions, which persist under strong decoupling levels due to their ω_r dependence. In order to further examine the effect of MAS on heteronuclear interference observed during the RFDR recoupling period, we performed similar experiments at 30 kHz spinning frequency, described below.

(a) $\omega_r/2\pi = 10$ kHz   (b) $\omega_r/2\pi = 15$ kHz   (c) $\omega_r/2\pi = 20$ kHz

Magnetization

1.0
0.8
0.6
0.4
0.2
0

0 20 40 60 80 100 120   0 20 40 60 80 100 120   0 20 40 60 80 100 120

Decoupling field during ^{13}C π pulses, $\omega_{1H}^{pul}/2\pi$ (kHz)
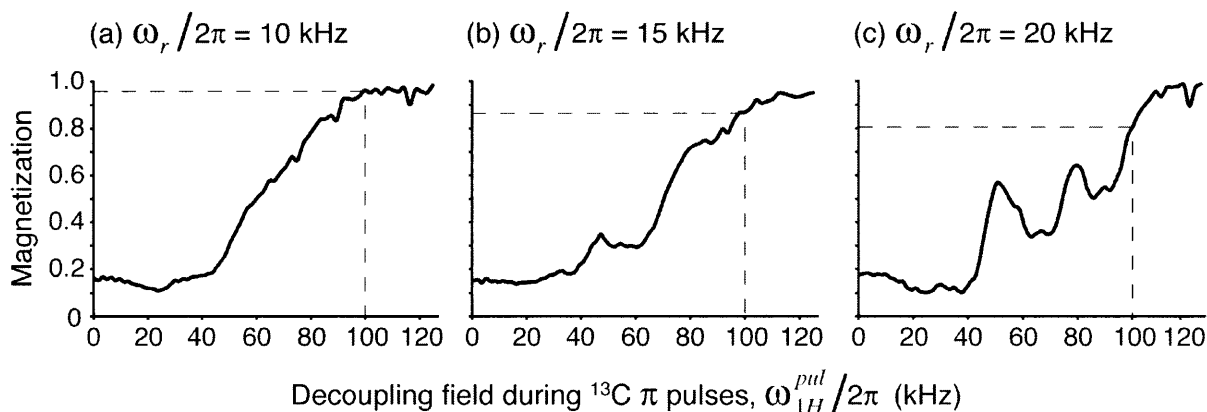
Figure 2. Experimental ^{13}C magnetization monitoring profiles showing the effect of the decoupling field strength during the π pulses of RFDR at various spinning frequencies for 2-^{13}C-Glycine. The RFDR period was 1.6 ms and the ^{13}C recoupling pulses applied ω_{1C} = 33.3 kHz near resonance. The decoupling field during the recoupling windows ($\omega_{1H}^{win}/2\pi$) was kept constant at 85 kHz, while the decoupling field during the recoupling pulses ($\omega_{1H}^{pul}/2\pi$) was increased from 0 to 125 kHz (in ~1.5 kHz intervals) to generate each profile. We can observe that as the MAS frequency is increased, the interference between ω_{1H}^{pul} and ω_{1C} persists to high decoupling levels and creates local minima for the magnetization. The dashed line emphasizes the decreasing decoupling effectiveness of a 3:1 mismatch between ω_{1H}^{pul} and ω_{1C} as the spinning frequency increases.

B. Decoupling interference at high spinning frequencies: Experimental results

We studied the effect of ^{1}H decoupling during RFDR at a high MAS frequency and high magnetic field with the uniformly [^{13}C, ^{15}N] labeled tripeptide N-f-MLF-OH. At a static magnetic field of 17.6 T, the ^{13}C chemical shifts of amino acids span 30 kHz, and dictate the optimal experimental conditions with which to obtain efficient RFDR recoupling. In particular, the choice of bandwidth of the recoupling pulses, and therefore the ^{13}C RF field amplitudes applied during the recoupling period, depend strongly on proper decoupling settings.

While heteronuclear interference has an effect on the recoupling efficiency of a pulse sequence, its effect during the RFDR pulse train is readily observable on the magnetization of a single ^{13}C spin, as illustrated above. The conditions that efficiently avoid ^{1}H-^{13}C depolarization losses are essential for minimizing interference terms during RFDR recoupling of multiple ^{13}C spins. Therefore it is important to examine the depolarization profile of uniformly ^{13}C labeled systems in order to elucidate optimal decoupling settings to be employed during the mixing

period. Indeed, conducting the experiments illustrated in Figure 2 on uniformly ^{13}C labeled N-f-MLF-OH yield similar profiles as those obtained with singly ^{13}C labeled glycine (data not shown). At higher MAS frequencies, the magnetization profile becomes increasingly complicated due to a multitude of narrow interference conditions, as illustrated in Figure 3 for $\omega_r/2\pi$ = 30.303 kHz. Figures 3a and 3b show the same type of magnetization profile as presented in Figure 2, in which the ^{13}C magnetization remaining after 1.6 ms of RFDR pulses is monitored as the decoupling field during the π pulses is increased while the decoupling field during the windows is maintained constant at 85 kHz. Figure 3a illustrates that when applying 33.3 kHz ^{13}C pulses, as in Figure 2, a very fine structure is observed in the magnetization profile at $\omega_r/2\pi$ = 30.303 kHz, with many maxima and minima separated by only a few kHz. It is evident from Figure 3a that application of a strong decoupling field does not guarantee minimal depolarization at a 30.303 kHz MAS frequency, since significant fluctuations in the magnetization can be observed even in the range of 90 to 110 kHz ^{1}H decoupling. When the ^{13}C RF pulses are set to a nutation frequency of 100 kHz, as illustrated in Figure 3b, the magnetization profile appears to be smoother, with fewer maxima and minima, and presents significant magnetization at low decoupling fields. Figure 3c shows a second kind of magnetization profile, in which the decoupling field during the [13]C π pulses is maintained equal to the decoupling field during the windows between the π pulses, and both decoupling fields are varied simultaneously, from 0 to 110 kHz. Interestingly, Figure 3c demonstrates that a significant amount of magnetization is preserved in the absence of ^{1}H decoupling irradiation.

These results at $\omega_r/2\pi$ = 30.303 kHz suggest that, at least in the case of strong ^{13}C RF fields, the magnetization conserved during RFDR in the absence of ^{1}H decoupling can be of similar or greater magnitude than that with strong decoupling irradiation. The horizontal dashed line in Figure 3 highlights this observation. The results also indicate the existence of sharp depolarization conditions that we proceeded to investigate via numerical simulations.

(a) $\omega_{1C}/2\pi = 33.3$ kHz,

$\omega_{1H}^{win}/2\pi = 85$ kHz

(b) $\omega_{1C}/2\pi = 100$ kHz,

$\omega_{1H}^{win}/2\pi = 85$ kHz

(c) $\omega_{1C}/2\pi = 100$ kHz,

$\omega_{1H}^{win} = \omega_{1H}^{pul}$

$\omega_{1H}^{pul}/2\pi$ (kHz)
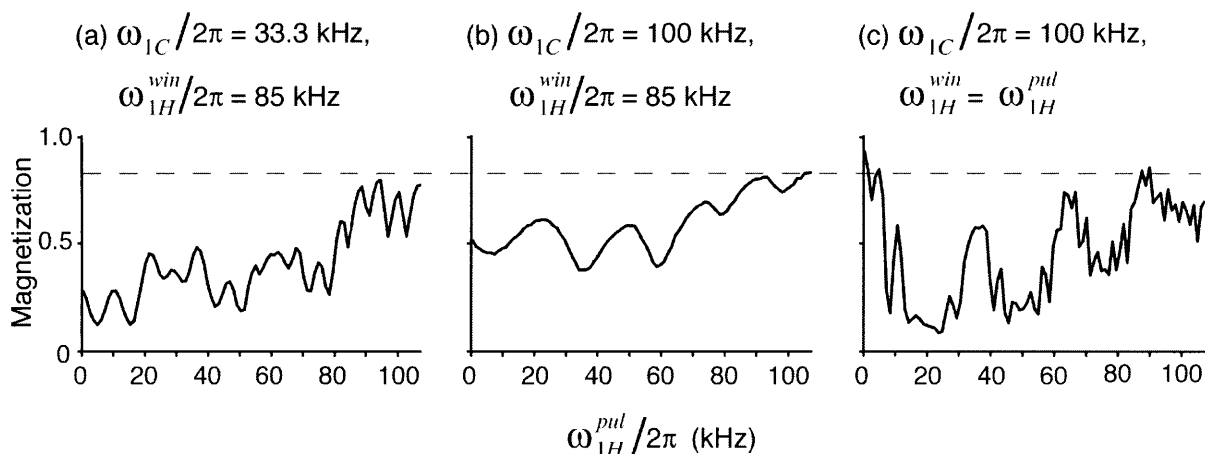
Figure 3. Experimental ^{13}C magnetization monitoring profiles at $\omega_r/2\pi = 30.303$ kHz for [U-^{13}C, ^{15}N] N-f-MLF-OH illustrating the complexity of the depolarization conditions encountered during RFDR mixing at high spinning frequencies. The profiles in (a) and (b) were carried out as described for those in Fig. 2, with $\omega_{1C}/2\pi = 33.3$ kHz and $\omega_{1C}/2\pi = 100$ kHz, respectively. The profile in (c) was obtained by simultaneously increasing $\omega_{1H}^{pul}/2\pi$ and $\omega_{1H}^{win}/2\pi$, in ~1.5 kHz intervals. These data plot the ^{13}C magnetization of the Met C^γ resonance after 1.6 ms of RFDR pulses. All the resonances in the N-f-MLF-OH spectrum follow the same magnetization pattern. The depolarization effects are minimal for carbonyl nuclei but nevertheless follow a similar pattern.

C. Decoupling interference at high spinning frequencies: Numerical simulations

The procedure for the optimization of decoupling irradiation during RFDR mixing involves the parallel calibration of the decoupling fields during the ^{13}C pulses (ω_{1H}^{pul}) and during the windows between them (ω_{1H}^{win}), considering the limited power applicable when working with biological samples. While at slow to moderate spinning frequencies this procedure consists of selecting a moderately strong ω_{1H}^{win} and the highest possible ω_{1H}^{pul} in order to avoid heteronuclear interference, the results in the previous section demonstrate that this strategy is not appropriate at high spinning frequencies. In order to thoroughly characterize the effect of these two decoupling fields we performed two-dimensional maps of numerical simulations in which the magnetization is monitored as a function of ω_{1H}^{pul} (x-axis) and ω_{1H}^{win} (y-axis).

Figure 4 summarizes our numerical simulations of heteronuclear interference during RFDR, which were performed using the SPINEVOLUTION package[61] with a four-spin system consisting of the CH_2 group of glycine plus the amide proton, and the RFDR mixing time was 1.6 ms of XY-16 pulses. Figure 4a shows that the basic procedure for decoupling optimization is

valid at 10 kHz spinning frequency, since efficient decoupling is achieved with a moderate ω_{1H}^{win} and a strong ω_{1H}^{pul}. In addition, increasing both decoupling fields simultaneously (along the diagonal in figure 4a) works well for high values of ω_{1H}^{pul} and ω_{1H}^{win}, although it is a suboptimal procedure due to the requirement of high decoupling fields.

On the other hand, at higher spinning frequencies such as 20 kHz (Figure 4b) and 30.303 kHz (Figure 4c), the magnetization maps show a substantial dependence on the precise values of ω_{1H}^{pul} and ω_{1H}^{win} due to multiple interference conditions with severe effects even at high decoupling fields. These simulations demonstrate that the appropriate procedure to avoid depolarization at high spinning frequencies is to carefully optimize both ω_{1H}^{pul} and ω_{1H}^{win}. Taking into account limited available RF powers, this optimization may consist of applying the highest possible value of ω_{1H}^{pul} and then carefully optimizing ω_{1H}^{win}. For example, following this procedure in Figures 4b and 4c, we find that for $\omega_{1H}^{pul}/2\pi$ = 100 kHz, the optimal value of $\omega_{1H}^{win}/2\pi$ is between 75 to 80 kHz. A decoupling calibration consisting of simply increasing ω_{1H}^{pul} and ω_{1H}^{win} simultaneously will tend to yield suboptimal decoupling settings.

In order to further characterize the depolarization conditions observed at high spinning frequencies, additional magnetization maps were simulated for a MAS frequency of 30.303 kHz and ^{13}C RF pulses of 50, 71, and 100 kHz, illustrated in Figures 4d, 4e, and 4f, respectively. These simulations follow similar patterns as those for 33 kHz ^{13}C π pulses (figure 4c), with multiple depolarization conditions and narrow regions of efficient decoupling. The regions of maxima and minima can be understood as matching conditions for the effective ^{1}H and ^{13}C RF fields, with their slopes given by the ratio of the duration of the applied ^{13}C π pulses to the duration of the windows between them.
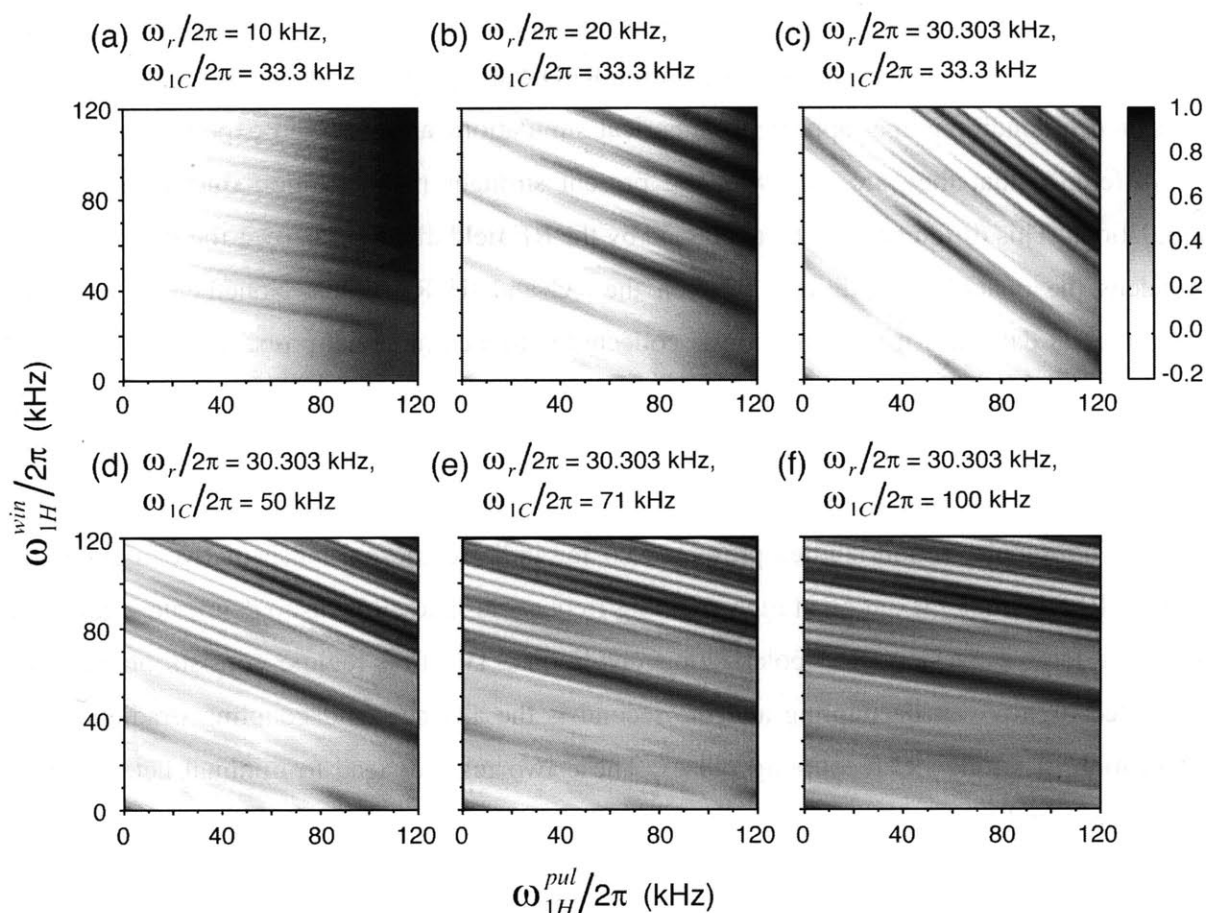
Figure 4. Contour plots of simulated ^{13}C magnetization remaining after a 1.6 ms RFDR period as a function of the decoupling fields ω_{1H}^{pul} and ω_{1H}^{win} for a four-spin system consisting of one ^{13}C spin and three ^{1}H spins, for increasing values of spinning frequency ω_r (a-c) and ^{13}C π pulse strength ω_{1C} (c-f). Corresponding cross-sections of these simulations are in good agreement with the experimental results illustrated in Figures 2 and 3.

Another important result from these simulations is the observation that significant magnetization is conserved in the absence of decoupling irradiation at high spinning frequencies (bottom left corners of Figures 4c-f). The absence of decoupling fields precludes heteronuclear interference effects, while the ^{13}C π pulses effectively attenuate heteronuclear dipolar interactions with an efficiency that improves with higher ^{13}C RF fields. The application of RFDR without decoupling irradiation during the mixing time is therefore a powerful method, specially in cases of limited RF fields or in order to avoid excessive sample heating, a common concern when working with biological samples.

The simulations thus confirm the experimental results of Figures 2 and 3, revealing the complexity of heteronuclear interference during RFDR at high MAS rates. All the heteronuclear interference conditions that appear in numerical simulations are observed experimentally, with the difference that they are narrower and present stronger magnetization fluctuations in the simulations. This difference can be explained by the RF field distribution over the sample, which broadens the matching conditions between the ^{13}C and ^{1}H RF fields. Nonetheless, since the experimental data shown in Figure 3 was collected with a commercial probe and a fully packed rotor, these sharp matching conditions can be expected to readily appear at high spinning frequency in most experimental setups, allowing the identification of optimal decoupling settings during the RFDR mixing period.

The depolarization studies presented so far indicate the existence of two regimes for optimal RFDR implementation at high spinning frequencies and allow us to delineate a guideline for the effective evasion of depolarization conditions. The first regime consists of carefully calibrated high-power decoupling and the second is the absence of decoupling irradiation and application of strong ^{13}C recoupling pulses. These two regimes lead to minimal heteronuclear interference and are a prerequisite for efficient RFDR recoupling. The choice between one and the other depends on experimental conditions, although the second regime (*sans* decoupling) can be expected to be the optimal one at MAS frequencies above 30 kHz.

D. RF bandwidth and finite-pulse effects

In addition to avoiding depolarization conditions, optimal experimental settings must include a judicious choice of ^{13}C RF bandwidth, taking into account the distribution of chemical shifts one wishes to observe. This choice is most relevant at high magnetic fields, where ^{13}C chemical shifts can span over 30 kHz. Figure 5 illustrates three different optimized settings for RFDR mixing at 30.303 kHz MAS and 17.6 T, where the ^{13}C resonances of N-f-MLF-OH span 32 kHz. The carrier frequency can be placed either near 95 ppm for broadband recoupling or near 35 ppm for recoupling of the aliphatic region alone, as indicated by the arrows in Figure 5a. Using high-power decoupling, 50 kHz ^{13}C pulses can cover the entire ^{13}C spectrum (Figure 5b), but it is more efficient to employ 33 kHz ^{13}C pulses in a narrow-bandwidth fashion (Figure 5c), focusing only in a small region of the spectrum, in this case the aliphatic region. However it is yet more efficient to turn off the decoupling fields during the RFDR period and employ 100 kHz

[13]C pulses in a broadband fashion (Figure 5d). Strong [13]C π pulses have the twofold advantage of preserving polarization in the absence of decoupling (and thus allowing the evasion of heteronuclear interference) and covering a large ^{13}C bandwidth. Furthermore, as the correlation spectra presented below suggest, at high MAS frequencies such as 30 kHz, even the shortest π pulses commonly available (up to 100 kHz) show excellent recoupling efficiencies, likely as a result of persistent finite-pulse effects.[15]
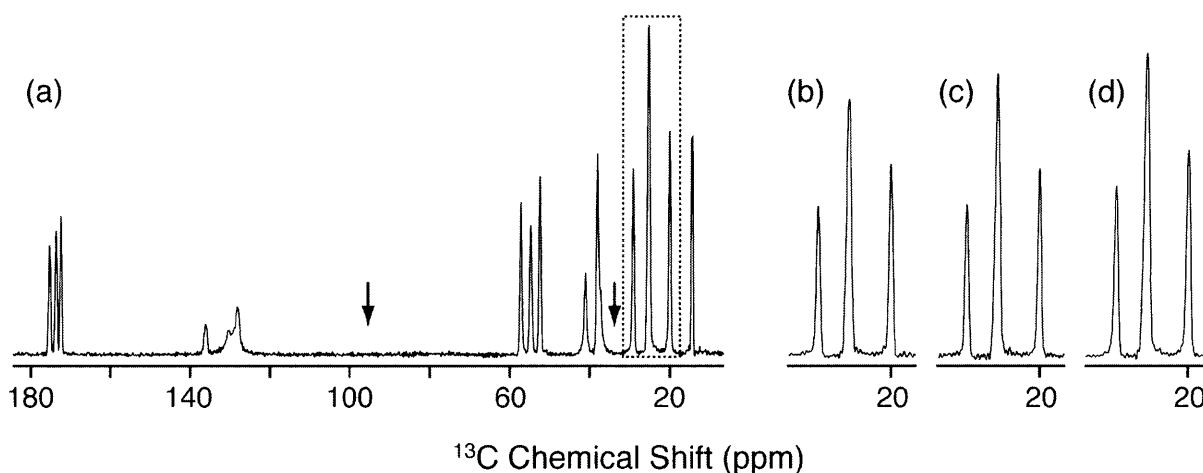
Figure 5. Comparison of different RFDR pulse bandwidths at $\omega_r/2\pi$ = 30.303 kHz and 750 MHz ^{1}H frequency. (a) Cross-polarization spectrum of N-f-MLF-OH. (b-d) Comparison of the aliphatic region around 25 ppm after a 1.6ms RFDR period with three different implementations: (b) high-power decoupling and 50 kHz ^{13}C pulses with the carrier at 90 ppm, (c) high-power decoupling and 33.3 kHz ^{13}C pulses with the carrier at 30 ppm, and (d) no decoupling during RFDR and 100 kHz ^{13}C pulses with the carrier at 90 ppm.

The comparison depicted in Figure 5 illustrates possible experimental conditions related to ^{13}C pulse bandwidth and heteronuclear depolarization. Avoiding depolarization conditions and utilizing appropriate ^{13}C RF bandwidths are prerequisites for efficient RFDR recoupling. However, the actual recoupling efficiency is determined by the effective homonuclear dipole-dipole recoupled Hamiltonian, which includes finite pulse effects. Finite pulse effects were recognized as being important for efficient recoupling in early π-pulse experiments.[35] Furthermore, the effects were responsible for obtaining excellent [13]C-^{13}C RFDR correlation spectra of various biological samples.[37-40] Subsequently, finite pulse effects were described in detail by Ishii[15] in RFDR experiments at high MAS frequencies ($\omega_r/2\pi$ ≥ 20 kHz). Specifically,

in the high spinning frequency regime, the RFDR recoupling mechanism is optimal for ^{13}C recoupling pulses that comprise a significant fraction of the rotor period, an effect that Tycko and co-workers have demonstrated and employed in several applications.[62-64] For instance, at $\omega_r/2\pi$ = 30 kHz, the theoretical efficiency of finite-pulse RFDR is optimal for ^{13}C RF fields of 30 kHz and decreases for stronger pulses as the finite-pulse effect is attenuated.[15,64]

However, when working with fully protonated samples, the deleterious effects of heteronuclear depolarization are of primary concern. As the experiments and simulations presented in the previous sections indicate, heteronuclear interference conditions vary for different ^{13}C RF pulse strengths. In particular, at high spinning frequency and in the absence of decoupling, stronger ^{13}C RF pulses are more efficient at preserving the ^{13}C magnetization than weaker ones. Therefore the choice of ^{13}C recoupling RF field must account not only for the desired chemical shift bandwidth, but also for the efficiency of the finite-pulse effect and the avoidance of heteronuclear depolarization. In the experiments at $\omega_r/2\pi$ = 30.303 kHz we describe below, the shortest ^{13}C RFDR pulses available with current probe technology (100 kHz) showed significant finite-pulse effects, as evidenced by their high efficiency when compared to longer pulses, and were thus utilized to obtain broadband correlation spectra with and without decoupling during the RFDR period.

E. Correlation spectra without decoupling during the mixing period

In order to evaluate the efficiency of the RFDR mixing scheme during high-frequency spinning, we performed two-dimensional (2D) correlation experiments on the tripeptide N-f-MLF-OH and the protein [U-^{13}C, ^{15}N] GB1. These spectra were recorded at $\omega_r/2\pi$ = 30.303 kHz on a spectrometer operating at 750 MHz ^{1}H Larmor frequency.

Figure 6 presents aliphatic regions of two 2D RFDR correlation spectra of N-f-MLF-OH recorded with carefully optimized high-power decoupling (a) and without decoupling (b) during the RFDR recoupling period of 8 ms. These spectra were recorded in a broadband fashion, with the carrier placed at 95 ppm and applying 100 kHz RFDR recoupling pulses, and present similar features, including the number of crosspeaks, crosspeak intensities (recoupling efficiency), and linewidths of crosspeaks and diagonal peaks. Similar characteristics are also encountered in the carbonyl-aliphatic region of the spectra (data not shown). The comparable qualities of these two

long-range correlation spectra and their high broadband efficiency demonstrate that RFDR can be implemented effectively at high MAS frequencies in both the high-power decoupling regime and the *sans* decoupling regime.

(a) High-power decoupling   (b) *Sans* decoupling

^{13}C Chemical Shift (ppm)

^{13}C Chemical Shift (ppm)
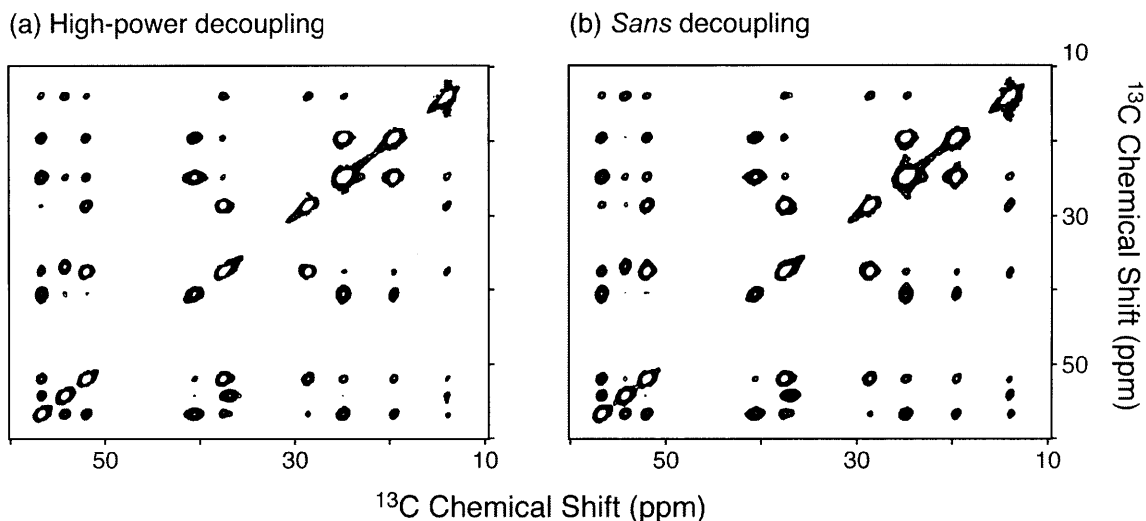
Figure 6. Long-range RFDR correlation spectra of [U-^{13}C, ^{15}N] N-f-MLF-OH recorded at $\omega_r/2\pi$ = 30.303 kHz and 750 MHz ^{1}H frequency with high-power decoupling (a) and without decoupling irradiation (b) during the mixing period, which consisted of 8.0 ms of $\omega_{1C}/2\pi$ = 100 kHz RFDR pulses. For the spectrum (a), the decoupling fields applied were set to $\omega_{1H}^{pul}/2\pi$ = 110 kHz and $\omega_{1H}^{win}/2\pi$ = 90 kHz. Both spectra were recorded with spectral widths of 40 kHz in both dimensions, 512 t_1 points with 4 scans each, and 1024 points in t_2. The data were processed with shifted sine bell apodization and zero filling to 1024 and 2048 points in t_1 and t_2, respectively.

In both cases, a careful choice of RF fields is imperative. In the high-power decoupling case, the ^{1}H decoupling fields must be optimized for both the short periods during the ^{13}C RFDR pulses and the longer periods during the windows between ^{13}C pulses. As described in the previous section, the optimal decoupling values will depend on the choice of ^{13}C RF field. In the absence of decoupling, the combined effect of avoiding depolarization and achieving efficient polarization transfer shows an overall improvement with increasing ^{13}C recoupling pulse strengths. The varying recoupling efficiencies of different ^{13}C RF fields are less pronounced than the superior decoupling properties of strong ^{13}C π pulses. To some extent this is expected, since at high MAS frequencies (say 30.303 kHz) even the strongest ^{13}C RF fields available (say 100 kHz) have a significant duration (5 μs) compared to the rotor period (30 μs) and present considerable finite-pulse effects that assist the recoupling mechanism.
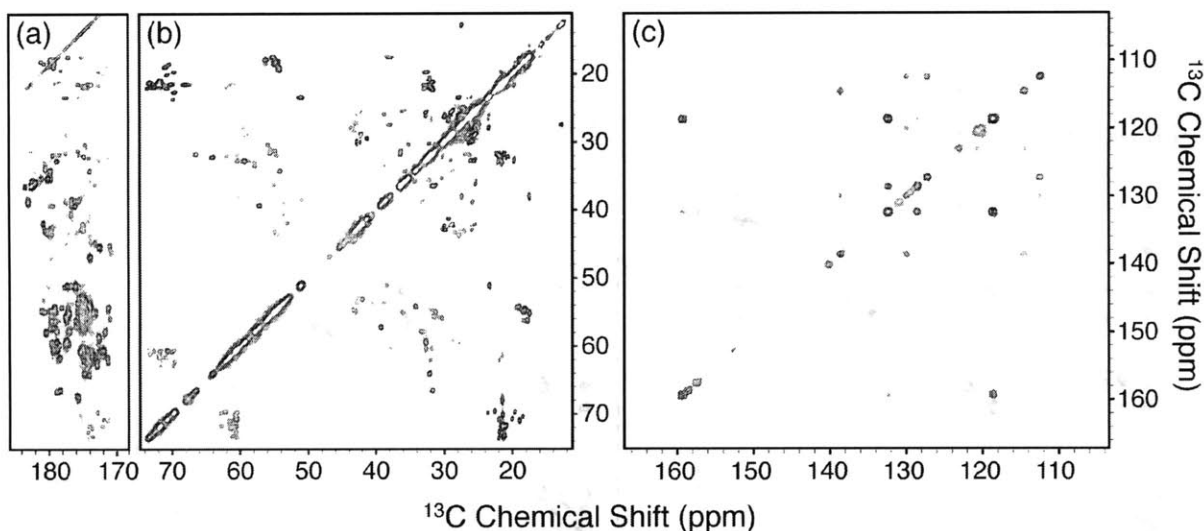
Figure 7. RFDR correlation spectrum of protein [U-^{13}C, ^{15}N] GB1 recorded at $\omega_r/2\pi$ = 30.303 kHz and 750 MHz ^{1}H frequency. (a) Carbonyl-aliphatic, (b) aliphatic, and (c) aromatic regions of a single 2D spectrum. The 2 ms mixing period was performed with 100 kHz ^{13}C RFDR pulses and without ^{1}H decoupling irradiation. During t_1 and t_2 periods of chemical shift evolution, 83 kHz TPPM decoupling was applied. The spectrum was recorded with spectral widths of 40 kHz in both dimensions, 380 t_1 points with 16 scans each, and 1024 points in t_2. The spectrum was processed with shifted sine bell apodization and zero filling to 1024 and 2048 points in t_1 and t_2, respectively.

The application of RFDR mixing at high MAS frequencies without simultaneous ^{1}H decoupling is of particular interest for studies of biological molecules since it can afford high-resolution correlation spectra while significantly attenuating sample heating. The efficiency of RFDR without decoupling is demonstrated in the high-resolution ^{13}C correlation spectrum of the 56-residue microcrystalline protein GB1 shown in Figure 7, recorded at 30.303 kHz MAS frequency employing 100 kHz RFDR pulses. Strong one-bond crosspeaks are found in all the expected regions of the spectrum, indicating highly broadband efficiency for RFDR *sans* decoupling. The slightly lower spectral quality of Figure 7 compared with previously published spectra[60] is likely the result of differences in sample preparation, and not related to experimental conditions. Indeed, high spinning frequencies afford optimal resolution, and in combination with RFDR, can yield high-quality protein correlation spectra.

2.5 Conclusions

The experimental data and numerical simulations presented in this article illustrate the complex patterns of heteronuclear interference at high spinning frequencies predicted from the theory of homonuclear recoupling in heteronuclear spin system and allow us to delineate optimal decoupling conditions to avoid these deleterious effects during the RFDR recoupling period. As in recent studies for double-quantum recoupling sequences,[17,18] we can define two regimes that avoid depolarization conditions, one utilizing high-power ^{1}H decoupling and one in the absence of ^{1}H decoupling fields. However, in the case of RFDR at high spinning frequencies, due to its intricate heteronuclear interference patterns, it is not sufficient to employ the highest ^{1}H decoupling powers commonly available, but instead the decoupling settings must be carefully calibrated. In the absence of decoupling, we have shown that strong ^{13}C RFDR pulses effectively preserve magnetization while circumventing heteronuclear depolarization and yield efficient homonuclear correlation spectra. The efficient implementation of RFDR recoupling at high spinning frequencies without decoupling during the mixing period, described in this article, should be of particular relevance to biological applications.

References

1. Ernst, M., Detken, A., Bockmann, A. & Meier, B. H. (2003). NMR Spectra of a Microcrystalline Protein at 30 kHz MAS. *J. Am. Chem. Soc.* **125**, 15807-15810.

2. Siemer, A. B., Ritter, C., Ernst, M., Riek, R. & Meier, B. H. (2005). High-Resolution Solid-State NMR Spectroscopy of the Prion Protein HET-s in its Amyloid Conformation. *Angew. Chem. Int. Ed.* **44**, 2441-2444.

3. Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V. & Griffin, R. G. (1995). Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* **103**, 6951-6958.

4. Tekely, P., Palmas, P. & Canet, D. (1994). Effect of Proton Spin Exchange on the Residual 13C MAS NMR Linewidths. Phase-Modulated Irradiation for Efficient Heteronuclear Decoupling in Rapidly Rotating Solids. *J. Magn. Reson., Ser. A* **107**, 129-133.

5. Gan, Z. & Ernst, R. R. (1997). Frequency- and phase-modulated heteronuclear decoupling in rotating solids. *Solid State Nucl. Magn. Reson.* **8**, 153-159.

6. Eden, M. & Levitt, M. H. (1999). Pulse sequence symmetries in the nuclear magnetic resonance of spinning solids: Application to heteronuclear decoupling. *J. Chem. Phys.* **111**, 1511-1519.

7. Fung, B. M., Khitrin, A. K. & Ermolaev, K. (2000). An improved broadband decoupling sequence for liquid crystals and solids. *J. Magn. Reson.* **142**, 97-101.

8. Khitrin, A. & Fung, B. M. (2000). Design of heteronuclear decoupling sequences for solids. *J. Chem. Phys.* **112**, 2392-2398.

9. Takegoshi, K., Mizokami, J. & Terao, T. (2001). 1H decoupling with third averaging in solid NMR. *Chem. Phys. Lett.* **341**, 540-544.

10. Detken, A., Hardy, E. H., Ernst, M. & Meier, B. H. (2002). Simple and efficient decoupling in magic-angle spinning solid-state NMR: the XiX scheme. *Chem. Phys. Lett.* **356**, 298-304.

11. De Paëpe, G., Lesage, A. & Emsley, L. (2003). The performance of phase modulated heteronuclear dipolar decoupling schemes in fast magic-angle-spinning nuclear magnetic resonance experiments. *J. Chem. Phys.* **119**, 4833-4841.

12. De Paëpe, G., Elena, B. & Emsley, L. (2004). Characterization of heteronuclear decoupling through proton spin dynamics in solid-state nuclear magnetic resonance spectroscopy. *J. Chem. Phys.* **121**, 3165-80.

13. Ramachandran, R., Bajaj, V. S. & Griffin, R. G. (2005). Theory of heteronuclear decoupling in solid-state nuclear magnetic resonance using multipole-multimode Floquet theory. *J. Chem. Phys.* **122**, 164503.

14. Thakur, R. S., Kurur, N. D. & Madhu, P. (2006). Swept-frequency two-pulse phase modulation for heteronuclear dipolar decoupling in solid-state NMR. *Chem. Phys. Lett.* **426**, 459-463.

15. Ishii, Y. (2001). 13C-13C dipolar recoupling under very fast magic angle spinning in solid-state nuclear magnetic resonance: Applications to distance measurements, spectral assignments, and high-throughput secondary-structure determination. *J. Chem. Phys.* **114**, 8473-8483.

16. Ernst, M., Meier, M. A., Tuherm, T., Samoson, A. & Meier, B. H. (2004). Low-Power High-Resolution Solid-State NMR of Peptides and Proteins. *J. Am. Chem. Soc.* **126**, 4764-4765.

17. Hughes, C. E., Luca, S. & Baldus, M. (2004). Radio-frequency driven polarization transfer without heteronuclear decoupling in rotating solids. *Chem. Phys. Lett.* **385**, 435-440.

18. Marin-Montesinos, I., Brouwer, D. H., Antonioli, G., Lai, W. C., Brinkmann, A. & Levitt, M. H. (2005). Heteronuclear decoupling interference during symmetry-based homonuclear recoupling in solid-state NMR. *J. Magn. Reson.* **177**, 307-317.

19. DePaëpe, G., Bayro, M. J., Lewandowski, J. & Griffin, R. G. (2006). Broadband Homonuclear Correlation Spectroscopy at High Magnetic Fields and MAS Frequencies. *J. Am. Chem. Soc.* **128**, 1776-1777.

20. Sun, B., Costa, P. R., Kocisko, D., Lansbury, J. & Griffin, R. G. (1995). Internuclear distance measurements in solid state nuclear magnetic resonance: Dipolar recoupling via rotor synchronized spin locking. *J. Chem. Phys.* **102**, 702-707.

21. Gregory, D. M., Mitchell, D. J., Stringer, J. A., Kiihne, S., Shiels, J. C., Callahan, J., Mehta, M. A. & Drobny, G. P. (1995). Windowless dipolar recoupling: the detection of weak dipolar couplings between spin 1/2 nuclei with large chemical shift anisotropies. *Chem. Phys. Lett.* **246**, 654-663.

22. Lee, Y. K., Kurur, N. D., Helmle, M., Johannessen, O. G., Nielsen, N. C. & Levitt, M. H. (1995). Efficient dipolar recoupling in the NMR of rotating solids. A sevenfold symmetric radiofrequency pulse sequence. *Chem. Phys. Lett.* **242**, 304-309.

23. Hohwy, M., Jakobsen, H. J., Eden, M., Levitt, M. H. & Nielsen, N. C. (1998). Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: A compensated C7 pulse sequence. *J. Chem. Phys.* **108**, 2686-2694.

24. Rienstra, C., Hatcher, M., Mueller, L., Sun, B., Fesik, S. & Griffin, R. (1998). Efficient Multispin Homonuclear Double-Quantum Recoupling for Magic-Angle Spinning NMR: 13C-13C Correlation Spectroscopy of U-13C-Erythromycin A. *J. Am. Chem. Soc.* **120**, 10602-10612.

25. Hohwy, M., Rienstra, C. M., Jaroniec, C. P. & Griffin, R. G. (1999). Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy. *J. Chem. Phys.* **110**, 7983-7992.

26. Brinkmann, A., Eden, M. & Levitt, M. H. (2000). Synchronous helical pulse sequences in magic-angle spinning nuclear magnetic resonance: Double quantum recoupling of multiple-spin systems. *J. Chem. Phys.* **112**, 8539-8554.

27. Hohwy, M., Rienstra, C. M. & Griffin, R. G. (2002). Band-selective homonuclear dipolar recoupling in rotating solids. *J. Chem. Phys.* **117**, 4973-4987.

28. Ishii, Y., Ashida, J. & Terao, T. (1995). 13C-1H dipolar recoupling dynamics in 13C multiple-pulse solid-state NMR. *Chem. Phys. Lett.* **246**, 439-445.

29. Bennett, A. E., Rienstra, C. M., Griffiths, J. M., Zhen, W., Lansbury, J. P. T. & Griffin, R. G. (1998). Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys.* **108**, 9463-9479.

30. Verel, R., Baldus, M., Nijman, M., van Os, J. W. M. & Meier, B. H. (1997). Adiabatic homonuclear polarization transfer in magic-angle-spinning solid-state NMR. *Chem. Phys. Lett.* **280**, 31-39.

31. Verel, R., Ernst, M. & Meier, B. H. (2001). Adiabatic Dipolar Recoupling in Solid-State NMR: The DREAM Scheme. *J. Magn. Reson.* **150**, 81-99.

32. De Paëpe, G., Lewandowski, J. & Griffin, R. (2007). Cosine-modulated rotary resonance. *In preparation.*

33. Bennett, A. E., Griffin, R. G., Ok, J. H. & Vega, S. (1992). Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys.* **96**, 8624-8627.

34. Gullion, T. & Vega, S. (1992). A simple magic angle spinning NMR experiment for the dephasing of rotational echoes of dipolar coupled homonuclear spin pairs. *Chem. Phys. Lett.* **194**, 423-428.

35. Ok, J. H., Spencer, R. G. S., Bennett, A. E. & Griffin, R. G. (1992). Homonuclear correlation spectroscopy in rotating solids. *Chem. Phys. Lett.* **197**, 389-395.

36. Sodickson, D. K., Levitt, M. H., Vega, S. & Griffin, R. G. (1993). Broad band dipolar recoupling in the nuclear magnetic resonance of rotating solids. *J. Chem. Phys.* **98**, 6742-6748.

37. Griffiths, J. M., Lakshmi, K. V., Bennett, A. E., Raap, J., van der Wielen, C. M., Lugtenburg, J., Herzfeld, J. & Griffin, R. G. (1994). Dipolar Correlation NMR Spectroscopy of a Membrane Protein. *J. Am. Chem. Soc.* **116**, 10178-10181.

38. Boender, G. J., Raap, J., Prytulla, S., Oschkinat, H. & de Groot, H. J. M. (1995). MAS NMR structure refinement of uniformly 13C enriched chlorophyll a/water aggregates with 2D dipolar correlation spectroscopy. *Chem. Phys. Lett.* **237**, 502-508.

39. Griffiths, J. M., Bennett, A. E., Engelhard, M., Siebert, F., Raap, J., Lugtenburg, J., Herzfeld, J. & Griffin, R. G. (2000). Structural investigation of the active site in bacteriorhodopsin: geometric constraints on the roles of Asp-85 and Asp-212 in the proton-pumping mechanism from solid state NMR. *Biochem.* **39**, 362-371.

40. Pauli, J., van Rossum, B., Forster, H., de Groot, H. J. & Oschkinat, H. (2000). Sample optimization and identification of signal patterns of amino acid side chains in 2D RFDR spectra of the alpha-spectrin SH3 domain. *J. Magn. Reson.* **143**, 411-416.

41. Boender, G. J. & Vega, S. (1998). Phase sensitive detection of 2D homonuclear correlation spectra in MAS NMR. *J. Magn. Reson.* **133**, 281-285.

42. Boender, G. J., Vega, S. & de Groot, H. J. M. (2000). Quantized field description of rotor frequency-driven dipolar recoupling. *J. Phys. Chem.* **112**, 1096-1106.

43. Goobes, G., Boender, G. J. & Vega, S. (2000). Spinning-frequency-dependent narrowband RF-driven dipolar recoupling. *J. Magn. Reson.* **146**, 204-219.

44. Goobes, G. & Vega, S. (2002). Improved narrowband dipolar recoupling for homonuclear distance measurements in rotating solids. *J. Magn. Reson.* **154**, 236-251.

45. Balbach, J. J., Ishii, Y., Antzutkin, O. N., Leapman, R. D., Rizzo, N. W., Dyda, F., Reed, J. & Tycko, R. (2000). Amyloid fibril formation by A beta 16-22, a seven-residue fragment of the Alzheimer's beta-amyloid peptide, and structural characterization by solid state NMR. *Biochem.* **39**, 13748-13759.

46. Morcombe, C. R., Gaponenko, V., Byrd, R. A. & Zilm, K. W. (2004). Diluting abundant spins by isotope edited radio frequency field assisted diffusion. *J. Am. Chem. Soc.* **126**, 7196-7.

47. Sanctuary, B. C. (1976). Multipole operators for an arbitrary number of spins. *J. Chem. Phys.* **64**, 4352-4361.

48. Ramachandran, R. & Griffin, R. G. (2005). Multipole-multimode Floquet theory in nuclear magnetic resonance. *J. Chem. Phys.* **122**, 164502.

49. Ramachandran, R., Lewandowski, J. R., van der Wel, P. C. A. & Griffin, R. G. (2006). Multipole-multimode Floquet theory of rotational resonance width experiments: 13C-13C distance measurements in uniformly labeled solids. *J. Chem. Phys.* **124**, 214107.

50. Ramachandran, R. & Griffin, R. G. (2006). Description of depolarization effects in double-quantum solid state nuclear magnetic resonance experiments using multipole-multimode Floquet theory. *J. Chem. Phys.* **125**, 44510.

51. Magnus, W. (1954). On the exponential solution of differential equations for a linear operator. *Comm. Pure Appl. Math.* **7**, 649-673.

52. Levitt, M. H., Oas, T. G. & Griffin, R. G. (1988). Rotary Resonance Recoupling in Heteronuclear Spin Pair Systems. *Isr. J. Chem.* **28**, 271-282.

53. Oas, T. G., Griffin, R. G. & Levitt, M. H. (1988). Rotary resonance recoupling of dipolar interactions in solid-state nuclear magnetic resonance spectroscopy. *J. Chem. Phys.* **89**, 692-695.

54. Nakai, T. & McDowell, C. A. (1994). Spinning-frequency-dependent linewidths in 1H-decoupled 13C magic-angle spinning NMR spectra. *Chem. Phys. Lett.* **227**, 639-644.

55. Ernst, M., Bush, S., Kolbert, A. C. & Pines, A. (1996). Second-order recoupling of chemical-shielding and dipolar-coupling tensors under spin decoupling in solid-state NMR. *J. Chem. Phys.* **105**, 3387-3397.

56. Maudsley, A. A. (1986). Modified Carr-Purcell-Meiboom-Gill sequence for NMR fourier imaging applications. *J. Magn. Reson., Ser. A* **69**, 488-491.

57. Gullion, T., Baker, D. B. & Conradi, M. S. (1990). New, compensated Carr-Purcell sequences. *J. Magn. Reson., Ser. A* **89**, 479-484.

58. Jaroniec, C. P., Tounge, B. A., Rienstra, C. M., Herzfeld, J. & Griffin, R. G. (2000). Recoupling of heteronuclear dipolar interactions with rotational-echo double-resonance at high magic-angle spinning frequencies. *J. Magn. Reson.* **146**, 132-9.

59. Rienstra, C. M., Tucker-Kellogg, L., Jaroniec, C. P., Hohwy, M., Reif, B., McMahon, M. T., Tidor, B., Lozano-Pérez, T. & Griffin, R. G. (2002). De novo determination of peptide structure with solid-state magic-angle spinning NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10260-5.

60. Franks, W. T., Zhou, D. H., Wylie, B. J., Money, B. G., Graesser, D. T., Frericks, H. L., Sahota, G. & Rienstra, C. M. (2005). Magic-angle spinning solid-state NMR spectroscopy of the beta1 immunoglobulin binding domain of protein G (GB1): 15N and 13C chemical shift assignments and conformational analysis. *J. Am. Chem. Soc.* **127**, 12291-305.

61. Veshtort, M. & Griffin, R. G. (2006). SPINEVOLUTION: a powerful tool for the simulation of solid and liquid state NMR experiments. *J. Magn. Reson.* **178**, 248-82.

62. Ishii, Y., Balbach, J. J. & Tycko, R. (2001). Measurement of dipole-coupled lineshapes in a many-spin system by constant-time two-dimensional solid state NMR with high-speed magic-angle spinning. *Chem. Phys.* **266**, 231-236.

63. Oyler, N. A. & Tycko, R. (2002). Multiple Quantum 13C NMR Spectroscopy in Solids under High-Speed Magic-Angle Spinning. *J. Phys. Chem. B* **106**, 8382-8389.

64. Tycko, R. (2007). Symmetry-based constant-time homonuclear dipolar recoupling in solid state NMR. *J. Chem. Phys.* **126**, 064506.

Chapter 3. Methods to Improve the Experimental Efficacy of MAS NMR Pulse Sequences

Section 3.1 Simultaneous Recoupling and Decoupling with an Adiabatic Sweep: The CMAR Scheme

Summary

We present a new homonuclear recoupling sequence, cosine modulated adiabatic recoupling (CMAR), that allows observation of 2D $^{13}C-^{13}C$ correlation spectra at high magnetic fields and MAS frequencies (10–30 kHz). The main advantages of the sequence are that it provides efficient, broadband dipolar recoupling and concurrently decouples the ^1H spins from the ^{13}C's. Thus, no additional ^1H decoupling is required during the mixing period, thereby significantly reducing the radio frequency power requirements for the experiment. Thus, CMAR significantly extends the range of applicability of the usual homonuclear recoupling techniques and should be of major interest for structure determinations of biomolecules at high magnetic fields.

3.1.1 Introduction

Structural studies of amyloid,[1] membrane[2] and nanocrystalline peptides and proteins[3, 4] are currently being performed with high resolution solid state NMR, and the initial step in the investigation of these systems is assignment of the spectra.[5, 6] Spectral assignments of ^{13}C and ^{15}N backbone and side chain resonances, and subsequently ^{13}C-^{13}C and ^{13}C-^{15}N distance and torsion angle measurements, are accomplished by recoupling dipolar interactions during magic angle spinning (MAS) experiments,[7-14] and a large number of broadband homonuclear recoupling techniques have been introduced which successfully accomplish this goal.[8, 15-18] However, essentially all of these experimental techniques were developed and function at low magnetic fields ($B_0 \le 12$ T, ≤ 500 MHz) and spinning frequencies ($\omega_r/2\pi \le 10$ kHz). Thus, for reasons detailed below, they are not applicable to experiments performed at high magnetic fields ($B_0 \sim$ 16-21 T, 700-900 MHz) and elevated spinning frequencies ($\omega_r/2\pi \sim$ 10-30 kHz) where structural studies are performed. Here we report the development of a novel dipolar recoupling scheme, cosine modulated (CM) adiabatic recoupling (CMAR), that performs broadband homonuclear dipolar recoupling at $\omega_r/2\pi$ ranging from 10 to 30 kHz and high B_0, and provides a means to perform spectral assignments. *The scheme has the interesting property that it simultaneously recouples the ^{13}C's and decouples the ^1H spins; thus, there is no ^1H decoupling present during the mixing period.*

3.1.2 Results and Discussion

It is axiomatic that MAS experiments at high B_0 require high $\omega_r/2\pi$ to attenuate rotational sidebands from the increased shift anisotropies, and therefore to fully manifest the available resolution and sensitivity of the spectra. However, most recoupling techniques do not function well at high $\omega_r/2\pi$ because the ratio of the ^{13}C nutation frequency (γB_1) to $\omega_r/2\pi$ (denoted as *N*) is typically 5 (SPC-5[17]), 7 (POST-C7[15]), or even 8.5 (DRAWS[8]) times $\omega_r/2\pi$, and concurrently, the ^1H nutation frequency should be ≥ 3 times the ^{13}C γB_1.[17] The high r.f. powers required to satisfy these constraints result in sample heating and challenge the technical integrity of commercial MAS probes.

These limitations stimulated the development of sequences with a reduced *N*.[17, 18] In addition, dipolar recoupling using an *adiabatic sweep* of the Hamiltonian spin system and an attenuated r.f. field on the recoupling channel was recently introduced.[16, 19] While the latter

approach concurrently reduces the r.f. power requirements at high MAS frequencies, it introduces a sensitivity to isotropic and anisotropic chemical shift offsets, which are amplified at high B_0.[16, 19] Finally, a recent study showed that some double quantum (DQ) recoupling sequences remain efficient in the *absence of ^1H decoupling* at moderate MAS rates.[20]

To address these limitations and to perform recoupling at high B_0 and $\omega_r/2\pi$ simultaneously, we have developed a novel dipolar recoupling scheme to observe broadband homonuclear chemical shift correlations. This new scheme, CMAR, combines rapid cosine modulation of the ^{13}C r.f. phase together with an adiabatic sweep of the cosine modulation amplitude, without ^1H decoupling. It is illustrated schematically in Figure 1 and functions as follows.

Recently, it was shown that CM[21] of the phase of the ^1H *decoupling* r.f. leads to the reintroduction of the ^1H–^1H couplings through a DQ homonuclear rotary resonance (HORROR)[7] mechanism.[22] The recoupling mechanism functions in a *modulation frame* (MF) defined by the modulation frequency and the mean axis of the irradiation.[21, 22] This suggests that the application of CM irradiation to the ^{13}C spin system can be employed to perform homonuclear recoupling of the ^{13}C-^{13}C interactions, and concurrently perform heteronuclear decoupling of the ^1H-^{13}C interactions. Furthermore, in this frame, effects of ^{13}C chemical shift anisotropies and r.f. inhomogeneity are attenuated leading to efficient DQ excitation even at high B_0.

The HORROR[7] condition in the MF[22] is defined by the following two equations:

$$ a_H = \frac{v_r}{v_1} \quad ; \quad \frac{v_c^H}{v_1} = 1 - \frac{1}{4}\left(\frac{v_r}{v_1}\right)^2 $$

where a_H is the amplitude of the modulation (rad), v_c^H the frequency (Hz) of the modulation, v_r the MAS frequency, and v_1 the ^{13}C nutation frequency. Under these conditions, the r.f. irradiation can be expressed in the MF as a static r.f. irradiation of amplitude $v_{eff} \approx a_H v_1/2$ that matches half the MAS frequency. This cosine-modulated rotary recoupling (CMRR) experiment is being explored as a part of a homonuclear recoupling technique for low-γ spins.

The CMAR scheme reported in this Section is an extension of this new recoupling scheme, in which an adiabatic sweep of the modulation amplitude a through the CMRR matching condition a_H is performed. The modulation frequency v_c is constant throughout the sweep of the amplitude a. As shown in Figure 1b, this irradiation scheme consists of a fast

cosine phase modulation envelope inside an adiabatic function (typically a tangent). The slow variation of the cosine phase modulation amplitude leads to the introduction of a fictitious Zeeman field that facilitates an adiabatic passage through the recoupling condition. The spin dynamics induced by this r.f. phase sweep in the MF is analogous to the r.f. amplitude sweep through the HORROR condition carried out in the rotating frame with the DREAM experiment.[16, 19, 23] The adiabatic process involved does not rely on the exact value of the dipolar coupling, and thus constitutes a robust and efficient approach for correlation experiments in multiply labeled samples. Furthermore, recoupling in the MF permits the use of CMAR without concurrent ^{1}H decoupling enabling its application at high MAS rates.
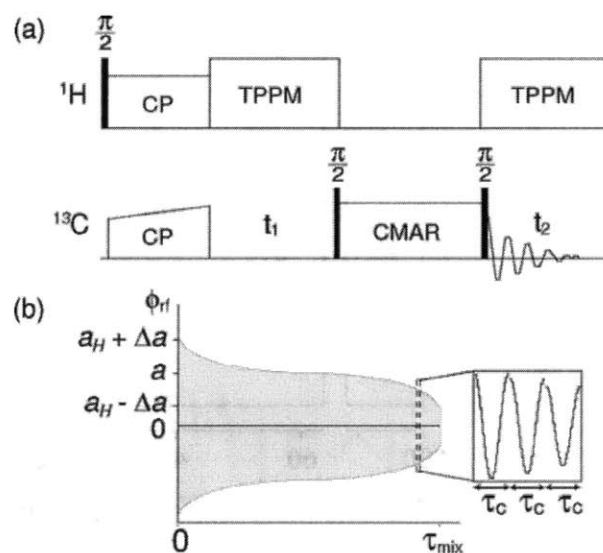
Figure 1. (a) 2D ^{13}C-^{13}C correlation pulse sequence used with CMAR recoupling sequence Note that no ^{1}H decoupling r.f. field is applied during the mixing time. (b).^{13}C r.f. phase modulation applied during the recoupling period. Note that τ_c stands for the inverse of the modulation frequency ν_c.

The experimental realization of CMAR as an efficient homonuclear correlation technique at high $\omega_r/2\pi$ is illustrated in Figure 2, which shows ^{13}C-^{13}C correlation spectra of a sample of the tripeptide [U-^{13}C, ^{15}N] N-f-MLF-OH,^{13} at $\omega_r/2\pi$ = 28.6 kHz and a magnetic field of 17.6 T (corresponding to 750 MHz ^{1}H Larmor frequency). γB_1 (^{13}C) = 100 kHz and no ^{1}H decoupling was employed. This data set was collected using a 2.5 mm, triple-channel Bruker probe.
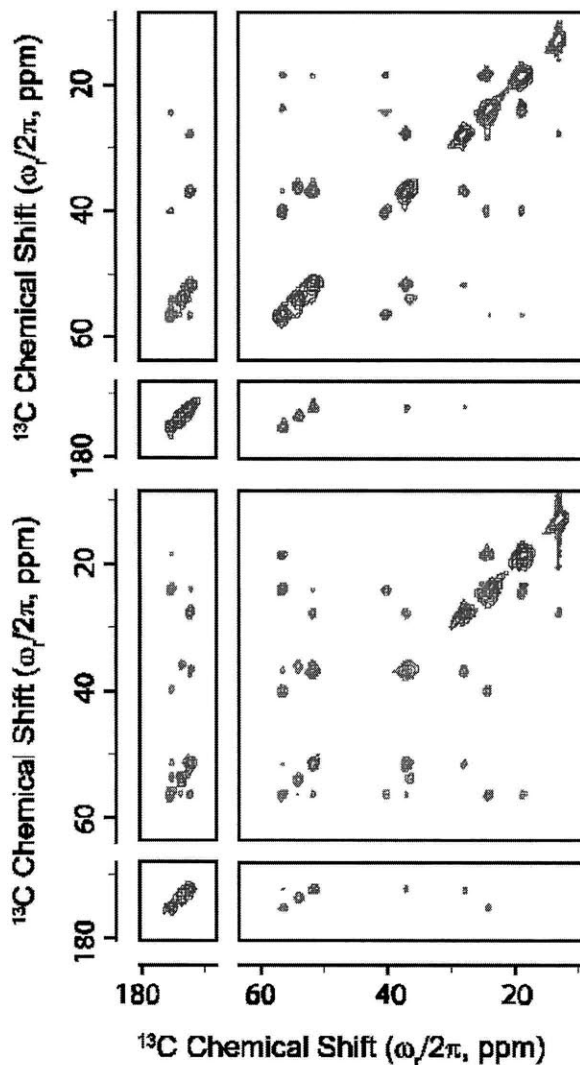
Figure 2. 2D correlation spectrum of the tripeptide [U-^{13}C, ^{15}N] N-f-MLF-OH using CM$_{3.5}$AR recoupling at 28.6 kHz MAS and ~100 kHz of ^{13}C r.f. field strength using (top) 2 ms and (bottom) 5 ms mixing time, with no ^{1}H decoupling. The evolution and acquisition periods used 83 kHz of TPPM decoupling, and 512 points were collected in the direct and 128 points in the indirect dimension, with 8 scans per transient.

These high B$_0$ and $\omega_r/2\pi$ spectra present correlations for all directly bonded spins, whose isotropic chemical shifts span ~35 kHz, illustrating the efficient broadband performance of the CMAR recoupling scheme. These correlation spectra also demonstrate one of the remarkable features of this recoupling technique, namely, the fact that high intensity r.f. recoupling pulses can be safely applied during the mixing time because there is no concurrent proton decoupling irradiation. Furthermore, application of a sweep through the resonant matching condition

improves the robustness of the CMAR pulse sequence with respect to phase imperfections, which are of practical concern in the application of the CMRR scheme.

In summary we have introduced a novel adiabatic double-quantum recoupling scheme. The advantages of this new technique are several. First, the experiment can be performed with a ^{13}C r.f. field of sufficient strength to uniformly recouple the entire ^{13}C spectrum at high B_0 resulting in true broadband behavior for the CMAR experiment. This characteristic represents an improvement over the DREAM experiment, in which the mean amplitude of the sweep is only half the MAS frequency. The recently developed DREAM-C7[24] sequence addresses this issue, but is limited to low MAS frequencies (< 15 kHz) due to high r.f. power requirements. With CM_pAR, the optimal ratio (p) of the modulation frequency (approximately the r.f. field strength) to the MAS frequency lies in the range of 3 to 8. The exact value is a compromise between direct interference between the MAS and r.f. averaging (small p value) and a small phase modulation amplitude a_H (high p value). The spectra in Figure 2 demonstrate that $CM_{3.5}AR$ functions efficiently at high MAS frequencies (28.6 kHz) with p = 3.5. Furthermore, without ^1H irradiation, higher r.f. fields can be applied to the ^{13}C's allowing more efficient recoupling at high MAS rates and magnetic fields, which in turn results in lower overall levels of sample heating. These characteristics extend the range of applicability of homonuclear recoupling techniques to high-resolution conditions, and should be of major interest for structure determination of biomolecules.

References

(1)   Jaroniec, C. P.; MacPhee, C. E.; Bajaj, V. S.; McMahon, M. T.; Dobson, C. M.; Griffin, R. G., Proc. Natl. Acad. Sci. U. S. A. 2004, 101, 3, 711-716.

(2)   Griffiths, J. M.; Lakshmi, K. V.; Bennett, A. E.; Raap, J.; Vanderwielen, C. M.; Lugtenburg, J.; Herzfeld, J.; Griffin, R. G., Journal of the American Chemical Society 1994, 116, 22, 10178-10181.

(3)   Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H., Nature 2002, 420, 6911, 98-102.

(4)   Zech, S. G.; Wand, A. J.; McDermott, A. E., Journal of the American Chemical Society 2005, 127, 24, 8618-8626.

(5)   Rienstra, C. M.; Hohwy, M.; Hong, M.; Griffin, R. G., Journal of the American Chemical Society 2000, 122, 44, 10979-10990.

(6)   Detken, A.; Hardy, E. H.; Ernst, M.; Kainosho, M.; Kawakami, T.; Aimoto, S.; Meier, B. H., J. Biomol. NMR 2001, 20, 3, 203-221.

(7)   Nielsen, N. C.; Bildsoe, H.; Jakobsen, H. J.; Levitt, M. H., J. Chem. Phys. 1994, 101, 3, 1805-1812.

(8)   Gregory, D. M.; Mitchell, D. J.; Stringer, J. A.; Kiihne, S.; Shiels, J. C.; Callahan, J.; Mehta, M. A.; Drobny, G. P., Chem. Phys. Lett. 1995, 246, 6, 654-663.

(9)   Costa, P. R.; Sun, B. Q.; Griffin, R. G., Journal of the American Chemical Society 1997, 119, 44, 10821-10830.

(10)   Jaroniec, C. P.; Tounge, B. A.; Herzfeld, J.; Griffin, R. G., Biophys. J. 2001, 80, 1, 368A-368A.

(11)   Carravetta, M.; Eden, M.; Johannessen, O. G.; Luthman, H.; Verdegem, P. J. E.; Lugtenburg, J.; Sebald, A.; Levitt, M. H., Journal of the American Chemical Society 2001, 123, 43, 10628-10638.

(12)   Jaroniec, C. P.; Filip, C.; Griffin, R. G., Journal of the American Chemical Society 2002, 124, 36, 10728-10742.

(13)   Rienstra, C. M.; Tucker-Kellogg, L.; Jaroniec, C. P.; Hohwy, M.; Reif, B.; McMahon, M. T.; Tidor, B.; Lozano-Perez, T.; Griffin, R. G., Proc. Natl. Acad. Sci. U. S. A. 2002, 99, 16, 10260-10265.

(14)   Ramachandran, R.; Ladizhansky, V.; Bajaj, V. S.; Griffin, R. G., Journal of the American Chemical Society 2003, 125, 50, 15623-15629.

(15)   Hohwy, M.; Jakobsen, H. J.; Eden, M.; Levitt, M. H.; Nielsen, N. C., J. Chem. Phys. 1998, 108, 7, 2686-2694.

(16)   Verel, R.; Baldus, M.; Ernst, M.; Meier, B. H., Chem. Phys. Lett. 1998, 287, 3-4, 421-428.

(17)   Hohwy, M.; Rienstra, C. M.; Jaroniec, C. P.; Griffin, R. G., J. Chem. Phys. 1999, 110, 16, 7983-7992.

(18)   Brinkmann, A.; Eden, M.; Levitt, M. H., J. Chem. Phys. 2000, 112, 19, 8539-8554.

(19)   Verel, R.; Ernst, M.; Meier, B. H., J. Magn. Reson. 2001, 150, 1, 81-99.

(20)   Hughes, C. E.; Luca, S.; Baldus, M., Chem. Phys. Lett. 2004, 385, 5-6, 435-440.

(21)   De Paepe, G.; Hodgkinson, P.; Emsley, L., Chem. Phys. Lett. 2003, 376, 3-4, 259-267.

(22)   De Paepe, G.; Elena, B.; Emsley, L., J. Chem. Phys. 2004, 121, 7, 3165-3180.

(23)   Verel, R.; Baldus, M.; Nijman, M.; van Os, J. W. M.; Meier, B. H., Chem. Phys. Lett. 1997, 280, 1-2, 31-39.

(24)   Verel, R.; Meier, B. H., Chemphyschem 2004, 5, 6, 851-862.

Section 3.2 Approaches to Compensate for Pulse Imperfections in Dipolar Recoupling Experiments

Adapted from "Dipolar recoupling in solid-state NMR by phase alternating pulse sequences" by James Lin, Marvin J. Bayro, Robert G. Griffin, and Navin Khaneja. Journal of Magnetic Resonance 197, 145-152 (2009) and from "Long-range correlations between aliphatic ^{13}C nuclei in protein MAS NMR spectroscopy" by Marvin J. Bayro, Thorsten Maly, Neil R. Birkett, Christopher M. Dobson, and Robert G. Griffin. Angewandte Chemie International Edition 48, 5708-5710 (2009).

Summary

We describe some new developments in the methodology of making heteronuclear and homonuclear recoupling experiments in solid state NMR insensitive to rf-inhomogeneity by phase alternating the irradiation on the spin system every rotor period. By incorporating delays of half rotor periods in the pulse sequences, these phase alternating experiments can be made χ encoded. The proposed methodology is conceptually different from the standard methods of making recoupling experiments robust by the use of ramps and adiabatic pulses in the recoupling periods. We show how the concept of phase alternation can be incorporated in the design of homonuclear recoupling experiments that are both insensitive to chemical shift dispersion and rf-inhomogeneity. Finally, a novel, phase-shifted XY-32 phase scheme is introduced for rf-inhomogeneity compensation of π-pulse echo recoupling sequences.

3.2.1 Introduction

An important application of solid-state nuclear magnetic resonance (NMR) spectroscopy is in structural analysis of `insoluble' protein structures such as membrane proteins, fibrils, and extracellular matrix proteins, which are exceedingly difficult to analyze using conventional atomic-resolution structure determination methods, including liquid-state NMR and X-ray crystallography (1-5). The goal of studying increasingly complex molecular systems is a strong motivation for the development of improved solid-state NMR methods. This section describes a set of principles for the design of heteronuclear and homonuclear dipolar recoupling experiments that are both broadband and insensitive of rf-field inhomogeneity. The merits of the proposed techniques with respect to state of the art methods are demonstrated.

For solids, the internal Hamiltonian not only contains isotropic interactions, such as isotropic chemical shifts and scalar couplings, but also anisotropic (i.e., orientation dependent) chemical shifts and dipole-dipole coupling interactions in the case of coupled spin-½ nuclei. This implies that each molecule/crystallite in a "powder" sample may exhibit different nuclear spin interactions leading to severe line broadening and thereby reduced spectral resolution and sensitivity. This problem may be alleviated using magic-angle spinning (MAS), which averages these interactions and hereby results in high-resolution conditions for solid samples. However, this also results in loss of useful parts of the anisotropic interactions like dipolar couplings, which carry information about distances between nuclei and can help in obtaining structural information. This has triggered the development of dipolar recoupling techniques (6-12), which selectively reintroduce these couplings to enable measurement of internuclear distances, torsion angles, and transfer of magnetization from spin to spin in the molecule. Such recoupling experiments are the building blocks of essentially all biological solid-state NMR experiments using `powder' samples. However, the recoupling experiments are sensitive to the amplitude of the radio-frequency fields and the orientation dependence of the dipolar coupling interaction.

The latter challenge motivates the present paper, where we present methods for making dipolar recoupling experiments less sensitive to rf-inhomogeneity and Hartmann-Hahn mismatch by phase alternating the irradiation on the spins every rotor period. In a previous study (13), it was shown that this simple technique makes recoupling experiments less sensitive to the dispersion in the strength of the radio-frequency field. These experiments however do not achieve uniform transfer efficiency for all γ, where γ represents the rotation of the crystallite

around the rotor axis. By introducing, half rotor period delays in the middle of the proposed phase-alternating recoupling blocks, the recoupling is made insensitive to the angle γ. One standard technique for making recoupling experiments robust to Hartmann Hahn mismatch is by using ramps or shaped adiabatic pulses (14, 15) on the rf-power during the recoupling period. A similar example is the adiabatic sweep through the phase modulation recoupling condition emplyed in the CMAR scheme (16) discussed in the preceding section. The methodology in this section is conceptually different. The main contribution of this section is the development of broadband homonuclear recoupling experiments that are robust to rf-inhomogeneity by using the concept of phase alternation. The first example improves the performance of heteronuclear double cross-polarization experiments. The second case in an improvement over broadband homonuclear mixing via cosine-modulated recoupling (16, 17). We show that by using phase alternation, we can make CMRR experiments less sensitive to rf-inhomogeneity. The third and last methodological development described is an improved phase sequence for consecutive echoes with π pulses, such as those used in the RFDR experiments described in chapter 2, that renders the sequence more robust with respect to rf field offsets.

3.2.2 Heteronuclear experiments

Phase alternation is employed in the MOIST (mismatch optimized I-S transfer) experiment, making its matching condition considerably more broad than that of (non-ramped) double cross-polarization (DCP) in ^{15}N-^{13}C polarization transfer. However, while the DCP experiment is independent of the crystallite angle γ (thus called γ-enconded), alternation of phases removes this independence, resulting in the averaging of the effective recoupled dipolar interaction over γ (in the principal axis to rotor-fixed frame) and decreased theoretical efficiency. The strategy undertaken here is to use phase alternation to make a DCP-type experiment more robust against rf inhomogeneity and at the same time reintroduce the γ-encoding into the sequence. This is achived through a series of delays with a duration of half a rotor period, $\tau_r/2$, during which the rf fields are set off the Hartmann Hahn condition. This scheme is called PATCHED, phase alternating experiment compensated by half-rotor period delays. The pulse schemes of MOIST and PATCHED are depicted in Figure 1.
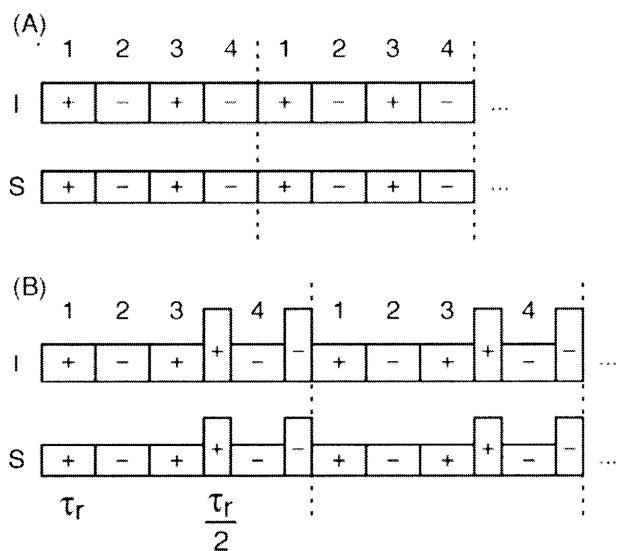
Figure 1. In (A) is shown the basic idea of the MOIST experiment, where the phase of the rf-field is switched every τ_r units of time. In (B) is shown the PATCHED experiment, where introducing delays of $\tau_r/2$, preceding and past every fourth period keeps the γ-encoding property of the recoupled Hamiltonian. The rf-power on both channels during half rotor period delays is chosen to avoid Hartmann-Hahn matching and rotary resonance and prevent dipolar recoupling.

Experimental ^{13}C spectra for DCP, Phase-alternation (MOIST) and PATCHED techniques applied for ^{15}N-^{13}Cα transfer in [U-^{13}C, ^{15}N] glycine, following ^{1}H-^{15}N cross-polarization, are shown in Figure 2. The experiments were performed in a spectrometer operating at 360 MHz 1H Larmor frequency, and with a spinning frequency of 8 kHz. During the mixing period, 100 kHz continuous wave decoupling on ^{1}H was applied. The nominal powers on ^{13}Cα and ^{15}N were 22.2 kHz and 30.2 kHz, respectively, calibrated by optimizing the transfer efficiency of the DCP experiment. The bottom panel in Fig. 4 shows the transfer profile for the DCP, MOIST and PATCHED experiments, when the carbon power is varied from its nominal value. The top panel shows sample spectra from the transfer profile in the bottom panel for discrete values of the ^{13}C rf field around the nominal value. The nominal powers on ^{13}Cα and ^{15}N during the half rotor period delays were calibrated to be 22.2 kHz and 30.2 kHz, respectively. The powers must be chosen large enough to prevent chemical shift evolution and avoid Hartmann-Hahn and rotary resonance conditions and thereby prevent dipolar recoupling. The exact values were optimized experimentally.

As expected from theory and simulations, the MOIST experiment is much more robust to rf-inhomogeneity compared to the standard DCP experiment. However, it is not a γ-encoded experiment and therefore has lower transfer efficiency. The PATCHED experiment is both insensitive to rf-inhomogeneity and γ-encoded, and as seen in Figure 2, is both robust and efficient.
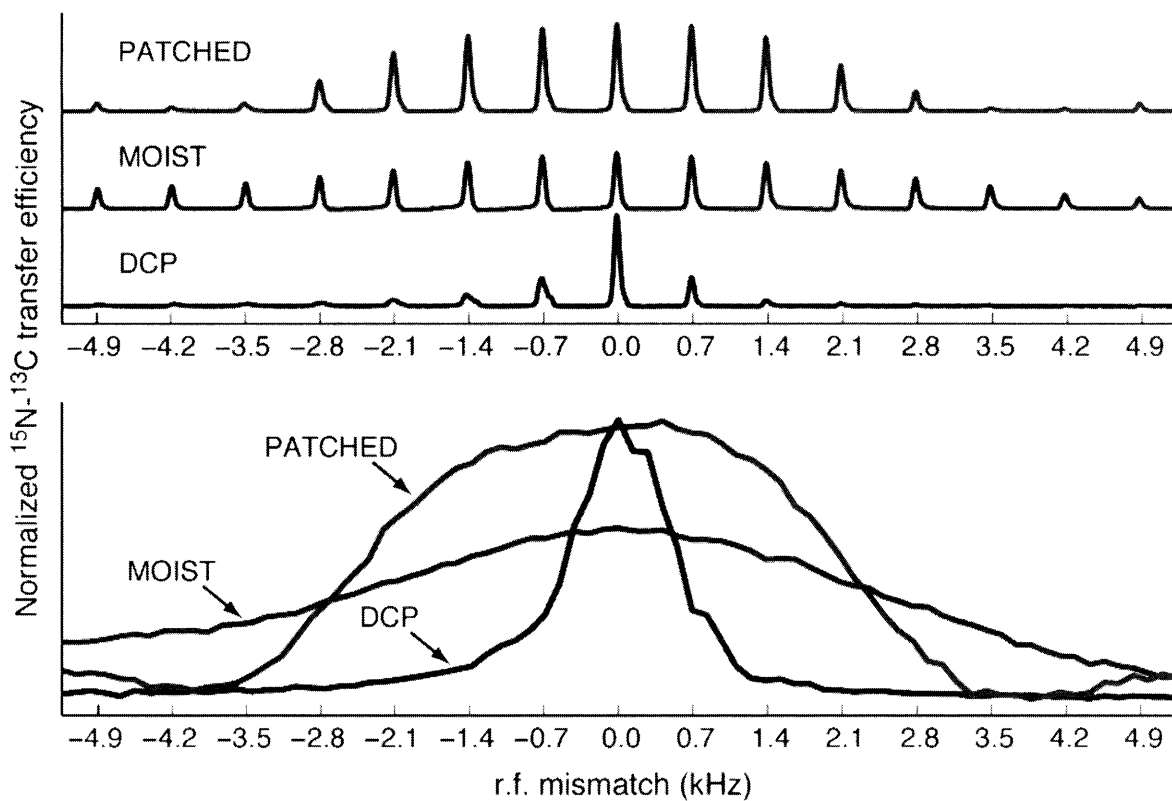
Figure 2. Profiles of recoupling efficiency versus radio frequency field offset for DCP, MOIST, and PATCHED. Experimental spectra for ^{15}N-^{13}C transfer at 8kHz MAS frequency. The bottom panel shows the transfer profile when the carbon power is varied from its nominal value. The top panel shows examples of spectra from the transfer profile in the bottom panel for discrete values of the ^{13}C rf field round its optimal value.

3.2.2 Homonuclear experiments

Similar phase-alternating approaches to compensate for rf imperfections as the ones described above can be applied to homonuclear recoupling sequences such as HORROR (homonuclear rotary resonance) recoupling (9), which similar to DCP, consists of single phase irradiation. However, the principle can also be applied to sequences with inherent frequencies other than the rotor period. A good candidate is CMRR (16, 17), in which the phase of the irradiation is modulated in a cosine fashion with period τ_c, usually a fraction of the rotation period. The approach presented here is to shift the phase of the rf field every other modulation period such that two consecutive periods have phases $\phi(t)$ and $\phi(t + \tau_c) = \pi - \phi(t)$, respectively. We refer to this pulse sequence as PAMORE, phase alternating modulated recoupling. The compensation for rf mismatch afforded by PAMORE is compared to a CMRR experiment from which it is derived in Figure 3 below.
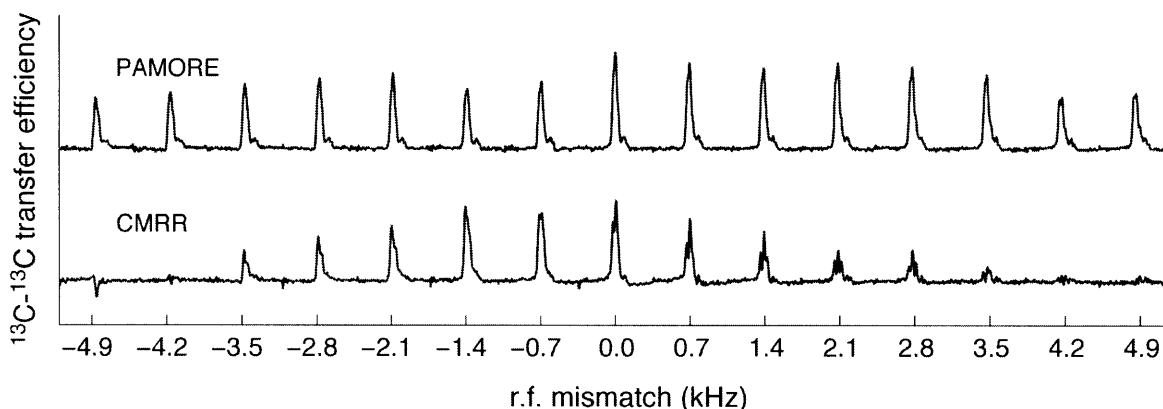
Figure 3. The top panel shows experimental spectra for CMRR and PAMORE pulse sequences, $^{13}C\alpha$-^{13}CO, transfers for a powder [U-^{13}C, ^{15}N] glycine using, recorded with 8 kHz MAS frequency and 100 kHz c.w. ^{1}H decoupling. The nominal power on the ^{13}C channel was calibrated to 48 kHz, with the carrier placed midway between $C\alpha$ and CO resonances that are 12 kHz apart. Phase modulation had a period of 20.83 μs, that is, one sixth of the rotor period. Both experiments used 15 rotor periods for the full transfer and started with initial y-magnetization on $^{13}C\alpha$ following and initial ^{15}N-$^{13}C\alpha$ DCP preparation step.

3.2.3 Phase-shifted XY-32 scheme for π-pulse recoupling

This section presents a new phase sequence for π-pulse echo trains found to compensate for radio frequency field deviations in RFDR experiments. The scheme is based on the XY-8 and XY-16 schemes by Gullion et al (18), and reduces rf imperfections particularly well for low-power ^{13}C pulses, down to the windowless limit. While several sequences have been proposed by permutations starting from XY-8, up to 64, 128 and even 256 steps, they are usually not effective beyond XY-16. Our 32-step sequence is simply XY-16 followed by the same sequence shifted by 90°, that is $Y\overline{X} - 16$. Thus the full sequence is

$$XYXY \ YXYX \ \overline{XYXY} \ \overline{YXYX} \ Y\overline{XY}X \ \overline{XY}XY \ \overline{YX}\overline{Y}X \ X\overline{YX}\overline{Y} \ .$$

Below is a demonstration of the robustness of this sequence compared to the widely used XY-16 scheme. The performance of XY-8 and XY-4 was worse than that of XY-16, as may be expected.
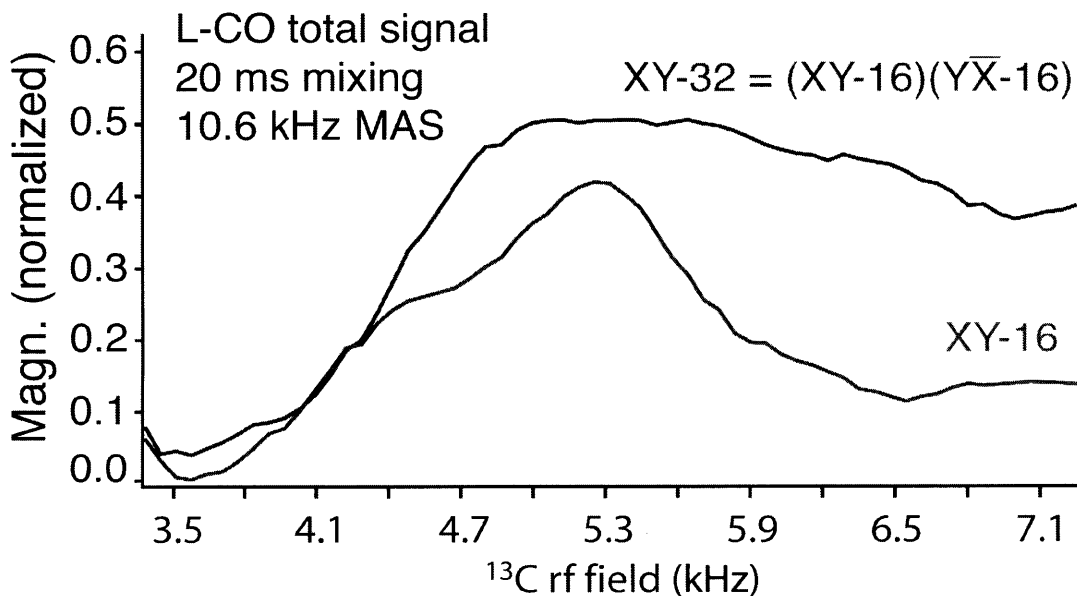
Figure 4. Robustness of phase-shifted XY-32 against deviations in ^{13}C rf field for windowless low-power RFDR (described in chapter 5). The observed signal is the remaining magnetization in the Leu-CO resonance of a [U-^{13}C,^{15}N] N-acetyl-VL sample after 20 ms of ^{13}C irradiation with a nominal field of one half the MAS frequency, 5.3 kHz. Phase-shifted XY-32 (not to be confuse with permutated XY-32) performs better than XY-16, especially for low-power ^{13}C π pulses.

3.3 Conclusions

In this chapter we have described two general and widely applied approaches to compensate for experimental imperfections during multiple-pulse sequences, and we have used them to generate novel dipolar recoupling schemes. The first general approach is to use an adiabatic sweep through an effective matching condition, as is typically used during cross-polarization and in the DREAM (15) experiment, in which the r.f. field is adiabatically varied around a center value. In the CMAR scheme presented here, the matching condition consists of a specific value of phase modulation frequency (the CMRR recoupling condition) through which the applied phase modulation frequency is adiabatically swept. We have used the CMAR scheme successfully in various protein systems such as the ones described in Part II of this thesis. The second general approach described in this chapter is phase alternation to reduce the deleterious effect of radio-frequency offsets. By applying this principle to Hartmann-Hahn matching and to CMRR recoupling, we demonstrated the superior performance of the resulting heteronuclear and homonuclear (respectively) dipolar recoupling pulse sequences with respect to pulse imperfections. Similarly, we have extended the XY-16 phase scheme to 32 steps by applying a simple XY pattern over XY-16, resulting in superior compensation for deviations in the nominal r.f. field used for rotor-synchronized π-pulse reocupling trains. This improved phase sequence is used successfully in some of the experiments described in chapters 5 and 8.

Without proper compensation, pulse imperfections can lead to serious signal losses, with a dramatic cumulative effect in long irradiation periods. The r.f. field strength typically depends on the position along the rotor axis, which often results in the need to confine the sample to the middle third of the rotor where deviations are minimal. The approaches presented here are aimed to relax some of the stringent experimental requirements often associated with MAS NMR pulse sequences.

References

[1] S.J. Opella, NMR and membrane proteins, Nat. Struct. Biol. 4 (1997), pp. 845–848.

[2] R.G. Griffin, Dipolar recoupling in MAS spectra of biological solids, Nat. Struct. Biol. 5 (1998), pp. 508–512.

[3] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein and H. Oschkinat, Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy, Nature 420 (2002), pp. 98–102.

[4] A.T. Petkova, Y. Ishii, J.J. Balbach, O.N. Antzutkin, R.D. Leapman, F. Deglaglio and R. Tycko, A structural model for alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR, Proc. Natl. Acad. Sci. USA 99 (2002), pp. 16742–16747.

[5] C.P. Jaroniec, C.E. MacPhee, V.S. Baja, M.T. McMahon, C.M. Dobson and R.G. Griffin, High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy, Proc. Natl. Acad. Sci. USA 101 (2004), pp. 711–716.

[6] J. Schaefer, R.A. McKay and E.O. Stejskal, Double-cross-polarization NMR of solids, J. Magn. Reson. 34 (1979), pp. 443–447.

[7] T. Gullion and J. Schaefer, Rotational-echo double-resonance NMR, J. Magn. Reson. 81 (1989), pp. 196–200.

[8] D.P. Raleigh, M.H. Levitt and R.G. Griffin, Rotational resonance in solid state NMR, Chem. Phys. Lett. 146 (1988), pp. 71–76.

[9] N.C. Nielsen, H. Bildsøe, H.J. Jakobsen and M.H. Levitt, Double-quantum homonuclear rotary resonance: Efficient dipolar recovery in magic-angle spinning nuclear magnetic resonance, J. Chem. Phys. 101 (1994), pp. 1805–1812.

[10] Y.K. Lee, N.D. Kurur, M. Helmle, O.G. Johannessen, N.C. Nielsen and M.H. Levitt, Efficient dipolar recoupling in the NMR of rotating solids. A sevenfold symmetric radiofrequency pulse sequence, Chem. Phys. Lett. 242 (1995), pp. 304–309.

[11] M. Hohwy, H.J. Jakobsen, M. Eden, M.H. Levitt and N.C. Nielsen, Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: A compensated C7 pulse sequence, J. Chem. Phys. 108 (1998), pp. 2686–2694.

[12] M.H. Levitt, Symmetry-based pulse sequences in magic-angle spinning solid-state NMR, Encyclopedia of NMR, Wiley, Chichester (2002), pp. 165–196.

[13] M.H. Levitt, D. Suter and R.R. Ernst, Spin dynamics and thermodynamics in solid-state NMR cross polarization, J. Chem. Phys. 84 (8) (1986).

[14] M. Baldus, D.G. Geurts, S. Hediger and B.H. Meier, Efficient 15N-13C polarization transfer by adiabatic-passage Hartmann–Hahn cross polarization, J. Magn. Reson. A 118 (1996), pp. 140–144.

[15] R. Verel, M. Baldus, M. Nijman, J.W.M. Vanus and B.H. Meier, Adiabatic homonuclear polarization transfer in magic-angle-spinning solid-state NMR, Chem. Phys. Lett. 280 (1997), pp. 31–39.

[16] G. De Paëpe, M.J. Bayro, J. Lewandowski and R.G. Griffin, Broadband homonuclear correlation spectroscopy at high magnetic fields and MAS Frequencies, J. Am. Chem. Soc. 128 (2006), pp. 1776–1777.

[17] G. De Paëpe, J. Lewandowski and R.G. Griffin, Spin dynamics in the modulation frame: Application to homonuclear recoupling in magic angle spinning solid-state NMR, J. Am. Chem. Soc. 128 (2008), p. 124503.

[18] T. Gullion, D.B. Baker and M.S. Conradi. New, compensate Carr-Purcell sequences. J. Magn. Reson 89 (1900), pp. 479-484.

Chapter 4. The Dipolar Truncation Effect

Summary

Quantitative solid-state NMR distance measurements in strongly coupled spin systems are often complicated due to the simultaneous presence of multiple non-commuting spin interactions. In the case of zeroth-order homonuclear dipolar recoupling experiments, the recoupled dipolar interaction between distant spins is attenuated by the presence of stronger couplings to nearby spins, an effect known as dipolar truncation. In this article, we quantitatively investigate the effect of dipolar truncation on the polarization-transfer efficiency of various homonuclear recoupling experiments with analytical theory, numerical simulations and experiments. In particular, using selectively ^{13}C-labeled tripeptides, we compare the extent of dipolar truncation in model three-spin systems encountered in protein samples produced with uniform and alternating labeling. Our observations indicate that while the extent of dipolar truncation decreases in the absence of directly bonded nuclei, two-bond dipolar couplings can generate significant dipolar truncation of small, long-range couplings. Therefore, while alternating labeling alleviates the effects of dipolar truncation, and thus facilitates the application of recoupling experiments to large spin systems, it does not represent a complete solution to this outstanding problem.

4.1 Introduction

In the last decade high resolution solid-state NMR has emerged as the method of choice for structural investigations of complex, insoluble systems such as membrane- and amyloid proteins. Advances in solid-state NMR methodology that are essential for these experiments include the improved resolution available with magic-angle spinning (MAS), heteronuclear decoupling schemes, and optimized sample preparation techniques. Significant advances have also been achieved with the development of efficient dipolar recoupling schemes, which enable the measurement of internuclear distances by reintroducing the dipole-dipole coupling otherwise averaged out by MAS.[1] In particular, a large number of heteronuclear and homonuclear recoupling pulse sequences are now available and applicable to a variety of experimental conditions.[2-21] Most of the recoupling sequences generate in a zeroth-order approximation an effective dipolar-coupling Hamiltonian that contains zero-quantum or double-quantum two-spin operators. In this publication we focus primarily on zeroth-order sequences. Second-order or higher-order recoupling sequences[22-25] have different properties because they contain cross-terms between two interactions and the effective coupling constants depend on both interactions.

In spite of the advances in pulse sequence design, the successful application of recoupling methods to the accurate measurement of ^{13}C-^{13}C distance in macromolecules has typically been possible only when employing either specifically labeled spin-pair samples[26-37] or frequency-selective recoupling methods[38-47], experiments in which a single internuclear distance is measured at a time. Although highly accurate (±5% is not difficult to achieve), both of these approaches are of limited practical utility in protein structure determinations that benefit from the availability of hundreds of constraints even for small proteins. For instance, the most efficient rotational resonance based methods[43,46,48] yield only a modest number of relevant constraints[48], and are thus most useful at the refinement stage of a structure determination process and for addressing specific mechanistic questions. On the other hand, broadband recoupling schemes applied to uniformly labeled samples offer the possibility of measuring many internuclear distances simultaneously. However, the deleterious effects of multiple spin interactions interfere with the efficiency and analysis of zeroth-order broadband homonuclear dipolar recoupling schemes.[11,15,49-52] In particular, the recoupling of a small dipolar coupling between a distant spin pair is severely attenuated in the presence of a third spin strongly coupled to one of the first two, an effect known as dipolar truncation[52]. As a result of dipolar truncation, the multiple-spin

recoupled dipolar interactions are dominated by the strong dipolar coupling, prohibiting the measurement of the long-range distances that are important in constraining structures.

The dipolar truncation effect has impeded the application of many highly efficient zeroth-order broadband homonuclear recoupling sequences to the simultaneous measurement of multiple long internuclear ^{13}C-^{13}C distances in uniformly labeled proteins. Therefore, a potential approach to attenuating dipolar-truncation effects is to employ sparse isotopic labeling methods, such as the alternating labeling scheme introduced by LeMaster and Kushlan[53], which produces few directly bonded pairs of ^{13}C nuclei. Indeed, this labeling scheme has been employed in solid-state protein NMR with the intention of eliminating dipolar truncation by generating spin systems without strong dipole-dipole couplings for most amino acid residues.[54-56] However, as we will see below, the remaining two-bond couplings (0.5 kHz as opposed to 2.2 kHz) already result in significant attenuation of the structurally interesting long-range constraints.

In this chapter we describe a study of the dipolar truncation effect in homonuclear dipolar recoupling experiments aimed at long-range polarization transfer. First, we present a theoretical description of dipolar truncation, and then we investigate its effect on various zeroth-order recoupling pulse sequences via numerical simulations. Finally, using a set of three selectively labeled samples, we demonstrate dipolar truncation experimentally and illustrate its dependence on relative dipolar coupling magnitudes. The model systems studied allow us to compare the extent of dipolar truncation observed in homonuclear dipolar recoupling of protein spin systems produced by uniform and alternating ^{13}C labeling.

4.2 Theoretical Description

In this paper we focus on the effects of the reintroduction of dipolar interactions using rotor-synchronized pulse sequences[17] on the time evolution of the spin system during MAS. Depending on the form of the effective Hamiltonian, dipolar recoupling sequences are classified into double-quantum (DQ) and zero-quantum (ZQ) experiments. In the case of DQ recoupling sequences, the recoupled zeroth-order effective Hamiltonian for a three-spin system has been shown to be[4,6,7]

$$\bar{H}_{DQ} = \sum_{\substack{i,j=1 \\ i<j}}^{3} C_{ij,DQ} \left[I_i^+ I_j^+ + I_i^- I_j^- \right] \tag{1}$$

where $C_{ij,DQ}$ represents the scaled dipolar coefficient between spins i and j and is a function of several variables including the spin-system parameters as well as the scaling factors due to the RF pulses. In a similar vein the recoupled zeroth-order effective Hamiltonian in the ZQ case is represented by[8,14,38]

$$\bar{H}_{ZQ} = \sum_{\substack{i,j=1 \\ i<j}}^{3} C_{ij,ZQ} \left[I_i^+ I_j^- + I_i^- I_j^+ \right] \tag{2}$$

Since, the recoupled dipolar Hamiltonians (Eqs. 1 and 2) among the various spin pairs are homogeneous in nature (i.e., the interactions do not commute with one another), the resulting spin dynamics is more complicated and difficult to analyze.

To elucidate the effects of dipolar truncation and the mechanism of polarization transfer in a strongly coupled spin system, we employ a state-space approach (or indirect method) to describe the spin dynamics. Using the standard basis functions for a three-spin system, the ZQ and DQ spin Hamiltonian derived in the appropriate interaction frame to zeroth order is represented by Eqs. 3 and 4, respectively.

	$\|\beta\beta\beta\rangle$	$\|\alpha\beta\alpha\rangle$	$\|\beta\alpha\alpha\rangle$	$\|\alpha\alpha\beta\rangle$	$\|\alpha\alpha\alpha\rangle$	$\|\beta\alpha\beta\rangle$	$\|\alpha\beta\beta\rangle$	$\|\beta\beta\alpha\rangle$
$\langle\beta\beta\beta\|$	0	0	0	0				
$\langle\alpha\beta\alpha\|$	0	0	C_{12}	C_{23}				
$\langle\beta\alpha\alpha\|$	0	C_{12}	0	C_{13}				
$\langle\alpha\alpha\beta\|$	0	C_{23}	C_{13}	0				
$\langle\alpha\alpha\alpha\|$					0	0	0	0
$\langle\beta\alpha\beta\|$					0	0	C_{12}	C_{23}
$\langle\alpha\beta\beta\|$					0	C_{12}	0	C_{13}
$\langle\beta\beta\alpha\|$					0	C_{23}	C_{13}	0

(3)

	$\|\beta\alpha\beta\rangle$	$\|\alpha\alpha\alpha\rangle$	$\|\beta\beta\alpha\rangle$	$\|\alpha\beta\beta\rangle$	$\|\alpha\beta\alpha\rangle$	$\|\beta\beta\beta\rangle$	$\|\alpha\alpha\beta\rangle$	$\|\beta\alpha\alpha\rangle$
$\langle\beta\alpha\beta\|$	0	C_{13}	0	0				
$\langle\alpha\alpha\alpha\|$	C_{13}	0	C_{12}	C_{23}				
$\langle\beta\beta\alpha\|$	0	C_{12}	0	0				
$\langle\alpha\beta\beta\|$	0	C_{23}	0	0				
$\langle\alpha\beta\alpha\|$					0	C_{13}	0	0
$\langle\beta\beta\beta\|$					C_{13}	0	C_{12}	C_{23}
$\langle\alpha\alpha\beta\|$					0	C_{12}	0	0
$\langle\beta\alpha\alpha\|$					0	C_{23}	0	0

(4)

In this representation of the effective Hamiltonian, the recoupling sequences are assumed to be ideal, i.e., the chemical shifts are ignored and the Hamiltonian is block-diagonal. If the coupling between spins 1 and 3 is set to zero ($C_{13} = 0$) then we have a representation for the Hamiltonian that is identical for both ZQ and DQ recoupling sequences. Such an approach allows derivation of analytical formulas that are common to both types of sequences. Following the standard procedure, the time-dependence of any observable is calculated by:

$$\langle O \rangle (t) = Tr\left[O\rho(t)\right] = Tr\left[OU(t,0)\rho(0)U^{-1}(t,0)\right]$$

(5)

Since the effective Hamiltonian is block-diagonal, the effective propagator $U(t,0)$ (i.e., $U(t,0) = \exp(-i\bar{\bar{H}}t)$) for both types of sequences can be represented as depicted in Eq. 6.

117

$$U(t,0) =$$

	$\lvert 1\rangle$	$\lvert 2\rangle$	$\lvert 3\rangle$	$\lvert 4\rangle$
$\langle 1\rvert$	1	0	0	0
$\langle 2\rvert$	0	$\cos(\omega t)$	$-i\dfrac{C_{12}}{\omega}\sin(\omega t)$	$-i\dfrac{C_{23}}{\omega}\sin(\omega t)$
$\langle 3\rvert$	0	$-i\dfrac{C_{12}}{\omega}\sin(\omega t)$	$\dfrac{C_{12}^2}{\omega^2}\cos(\omega t)+\dfrac{C_{23}^2}{\omega^2}$	$\dfrac{C_{12}C_{23}}{\omega^2}(\cos(\omega t)-1)$
$\langle 4\rvert$	0	$-i\dfrac{C_{23}}{\omega}\sin(\omega t)$	$\dfrac{C_{12}C_{23}}{\omega^2}(\cos(\omega t)-1)$	$\dfrac{C_{23}^2}{\omega^2}\cos(\omega t)+\dfrac{C_{12}^2}{\omega^2}$

$+$

	$\lvert 5\rangle$	$\lvert 6\rangle$	$\lvert 7\rangle$	$\lvert 8\rangle$
$\langle 5\rvert$	1	0	0	0
$\langle 6\rvert$	0	$\cos(\omega t)$	$-i\dfrac{C_{12}}{\omega}\sin(\omega t)$	$-i\dfrac{C_{23}}{\omega}\sin(\omega t)$
$\langle 7\rvert$	0	$-i\dfrac{C_{12}}{\omega}\sin(\omega t)$	$\dfrac{C_{12}^2}{\omega^2}\cos(\omega t)+\dfrac{C_{23}^2}{\omega^2}$	$\dfrac{C_{12}C_{23}}{\omega^2}(\cos(\omega t)-1)$
$\langle 8\rvert$	0	$-i\dfrac{C_{23}}{\omega}\sin(\omega t)$	$\dfrac{C_{12}C_{23}}{\omega^2}(\cos(\omega t)-1)$	$\dfrac{C_{23}^2}{\omega^2}\cos(\omega t)+\dfrac{C_{12}^2}{\omega^2}$

$$(6)$$

The ket vectors $\lvert 1\rangle$, $\lvert 2\rangle$...$\lvert 8\rangle$ are identical to the basis functions employed in the corresponding zero and double-quantum Hamiltonians illustrated in Eqs. 3 and 4, respectively. The coupling between spins 1 and 3 is set to zero.

With this formulation, the mechanism of polarization transfer and the effects of dipolar truncation are transparent. Starting with polarization on spin 2 only (i.e. $\rho(0)=I_{2z}$), the expectation values for ZQ polarization transfer to spins 1 and 3 from spin 2 are given by:

$$\langle I_{1z}\rangle(t) = \frac{C_{12}^2}{\left(C_{12}^2 + C_{23}^2\right)}\sin^2(\omega t)$$

$$\langle I_{2z}\rangle(t) = \cos^2(\omega t) \tag{7}$$

$$\langle I_{3z}\rangle(t) = \frac{C_{23}^2}{\left(C_{12}^2 + C_{23}^2\right)}\sin^2(\omega t)$$

where $\omega = \sqrt{C_{12}^2 + C_{23}^2}$. The equations illustrated above are identical for DQ transfer, except for a minus sign in the denominator of the expressions for $\langle I_{1z}\rangle$ and $\langle I_{3z}\rangle$. It is important to note that there is only a single frequency present in the time evolution under such a zeroth-order effective Hamiltonian. This is even true if all three dipolar couplings are present in the three spin system.

118

Only the relative intensities will change with a change in the relative magnitude of the dipolar coefficients. Neglecting relative orientations, when the two dipolar couplings are of equal magnitude ($C_{23} = C_{12}$), polarization transfer is identical between the two pairs of coupled spins. However, when one of the dipolar couplings is stronger, for instance between spins 2 and 3, it dominates the dynamics and causes polarization transfer to occur mostly from spin 2 to spin 3, reducing the extent of transfer from spin 2 and spin 1. For example if $C_{23} = 2C_{12}$ then 80% of the polarization from spin 2 is transferred to spin 3 and only 20% is transferred to spin 1. And if the couplings differ by an order of magnitude then the extent of transfer drops by two orders of magnitude, even in the quasi-equilibrium state[57].

The above description of dipolar truncation illustrates the magnitude of the expected effects in an idealized three-spin system that was initially outlined by Costa in 1996[52]. Detailed equations for the full three-spin system under double-quantum recoupling can be found in Hohwy *et al.*[11] However, a more realistic analysis of dipolar truncation requires consideration of higher-order contributions to the effective Hamiltonians, which contain cross terms between different interactions in the spin-system Hamiltonian. For this reason, we performed numerical simulations and dipolar recoupling experiments on model systems, the results of which are described in Section IV. The zeroth-order average Hamiltonian generated by RFDR recoupling is of the general form shown in Eq. (2) with an effective coupling constant that depends not only on the magnitude of the dipolar coupling but also on the ratio of the isotropic chemical shift differences and the spinning frequency.[17]

4.3 Experimental Methods

A. NMR Experiments

NMR experiments were performed on a custom designed spectrometer operating at a field of 11.7 T (500.125 MHz for ^{1}H, courtesy of D. J. Ruben) using a Chemagnetics triple resonance MAS probe equipped with a 4.0 mm spinning module. The peptide samples (~20 mg) were packed in the center third of the rotor to reduce the inhomogeneity of the rf field to ~5%. All the experiments consisted of ^{1}H-^{13}C cross polarization (CP) followed by a mixing period and detection. In 2D RFDR experiments, a chemical shift evolution period was inserted between the CP and the mixing period. The RFDR mixing period consisted of rotor-synchronized 50 kHz ^{13}C π pulses, with concurrent ^{1}H cw decoupling fields of 115 kHz and 90 kHz during ^{13}C pulses and the windows between them, respectively, optimized to avoid polarization losses due to heteronuclear interference[8,21]. During chemical shift evolution and detection periods, 83 kHz TPPM[58] decoupling was used applied. For each 2D RFDR spectrum, 280 t_1 points, with 4 scans each, were collected, with 512 points in the direct dimension. Data processing included shifted sine bell apodization and zero filling to 1024 points in each dimension. Spectra were processed and analyzed using NMRPipe[59].

B. Selectively Labeled Samples

The experimental results reported in this article were obtained utilizing three selectively labeled tripeptides illustrated in Fig. 1: Ala-Gly-Gly[60] labeled as Ala-[2-^{13}C]Gly-[1-^{13}C]Gly (AGG-2), Ala-[1,2-^{13}C$_2$]Gly-[1-^{13}C]Gly (AGG-3), and Gly-Gly-Val[61] labeled as [1-^{13}C]Val-[2-^{13}C]Gly-[1-^{13}C]Gly (GGV-3). All three nitrogen sites were ^{15}N labeled in all three compounds.
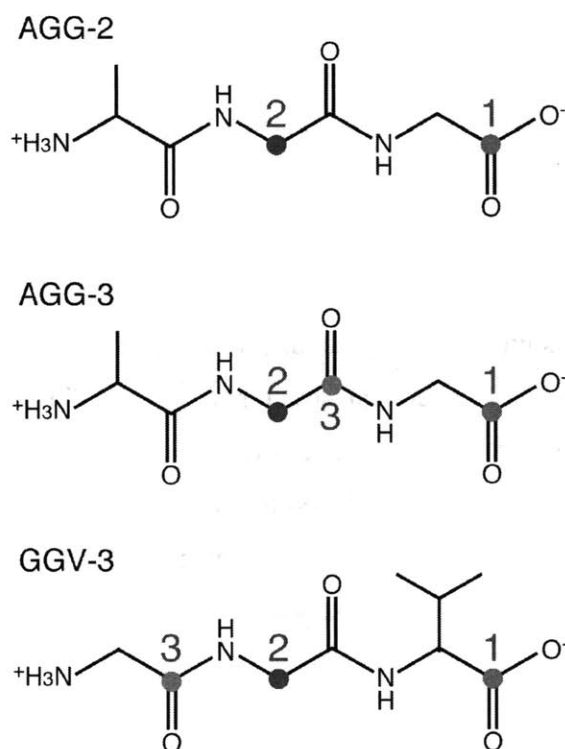
Figure 1. Labeling schemes employed in the model tri-peptides used to illustrate the effect of dipolar truncation. AGG-2 (Ala-[2-^{13}C]Gly-[1-^{13}C]Gly) is a two-spin system with a dipolar coupling of 66 Hz, corresponding to a distance of 4.86 Å between nuclei 1 and 2. AGG-3 (Ala-[1,2-^{13}C$_2$]Gly-[1-^{13}C]Gly) is a three-spin system formed by adding a third labeled nucleus (spin 3) to AGG-2, forming a strong dipolar coupling of 2.15 kHz with spin 2. The coupling between spins 1 and 3 is 150 Hz. GGV-3 ([1-^{13}C]Val-[2-^{13}C]Gly-[1-^{13}C]Gly) is a three-spin system similar to AGG-3, but with the third labeled nucleus (spin 3) two bonds away from spin 2. GGV-3 has a weak coupling of 80 Hz corresponding to a distance of 4.56 Å between spins 1 and 2, and a medium coupling of 550 Hz corresponding to a distance of 2.43 Å between spins 2 and 3.

The AGG-2 sample is a weakly coupled isolated spin pair with a dipole-dipole coupling of 66 Hz, and serves as a reference system for polarization transfer efficiency. The samples AGG-3 and GGV-3 are three-spin systems formed by incorporating a third spin into the two-spin system of AGG-2. In AGG-3, the third spin is directly bonded to spin 2, thus exhibiting a strong ^{13}C-^{13}C coupling (2.150 kHz) while the coupling between spins 1 and 3 is only 150 Hz. The third spin in GGV-3 is two bonds away from spin 2 and thus represents a moderate (550 Hz) coupling. AGG-3 and GGV-3 are simple models for the geometries encountered in uniform and odd-even labeling of amino acids, respectively. All the samples were diluted to 10% in natural abundance lattice to minimize the effects of intermolecular ^{13}C-^{13}C dipolar couplings.

C. Numerical Simulations

To evaluate dipolar truncation in various homonuclear recoupling sequences, numerical simulations were programmed in C++ using the GAMMA spin-simulation environment[62]. All simulations were powder averages of 1154 orientations, where the individual orientations were determined by the method of Cheng *et al.*[63,64] The time dependence of the Hamiltonian was approximated by subdividing each rotor period into at least 100 time steps with a time-constant Hamiltonian. Only the ^{13}C spins were simulated. The CSA parameters were estimated using the SIMMOL package[65], based on the crystal structures[60,61]. In addition, the SPINEVOLUTION program[66] was used to simulate the experimental RFDR data. Because of the efficiency of this program, this second set of simulations included not only ^{13}C spins, but several ^{1}H spins and the exact rf-irradiation scheme employed in the experiments in order to reproduce the experimentally observed polarization transfer profiles as closely as possible.

4.4 Results and Discussion

A. Dipolar truncation in homonuclear recoupling schemes: Numerical simulations

The results of numerical simulations demonstrating the effect of dipolar truncation for various recoupling sequences are presented in Fig. 2. These simulations start with magnetization on spin 2 and monitor the transfer to the distant spin (spin 1) and to the near spin (spin 3). The polarization transfer efficiency in an isolated spin pair (AGG-2 in Fig. 2) varies for different recoupling schemes depending on their particular characteristics such as scaling factor and compensation for chemical-shift terms, as examined previously. Recoupling of the weak dipole-dipole coupling corresponding to the 4.86 Å internuclear distance in the AGG-2 spin system requires long mixing times, which makes the contributions from small error terms that appear in higher-order contributions to the effective Hamiltonian more visible. Nevertheless, a variety of recoupling schemes show significant polarization transfer in such an isolated two-spin system.

Upon introduction of a third spin (spin 3) presenting a one-bond coupling to one of the spins in the original spin pair (spin 2), a dramatic reduction of polarization transfer to the distant spin (spin 1) is observed, as the simulations for AGG-3 in Fig. 2 illustrate. This happens despite the fact that in AGG-3 a relay polarization-transfer pathway from spin 2 via spin 3 to spin 1 is possible and most likely faster than the direct polarization transfer from spin 2 to spin 1. Indeed, dipolar truncation of a weak coupling by a one-bond ^{13}C-^{13}C coupling has a severe effect in all the recoupling schemes we studied, with polarization transfer to the distant ^{13}C spin reduced to a few percent in the most favorable cases ($SR4_4^1$ and RFDR). In the case of a third spin with a two-bond coupling (GGV-3 in Fig. 2), the transfer efficiency to the distant spin increases for all schemes compared to the one-bond dipolar truncation case, with considerable improvements for the zero-quantum recoupling sequences RFDR and $SR4_4^1$. However, comparing the simulations for GGV-3 to those for the isolated distant spin-pair (AGG-2), it is evident that a two-bond coupling is sufficient to cause significant dipolar truncation of the weak dipolar coupling in this model three-spin system. These results are in agreement with the expectations from the analytical solutions, which predict significant truncation effects for two dipolar couplings that differ by a factor of seven. However, the process that finally leads to the transfer to the distant spin is not described by the zero-order three spin model because it has a slower characteristic time constant than the direct transfer, in contrast to the expectations from this model Hamiltonian which predicts all transfers to proceed with the same frequency.

123
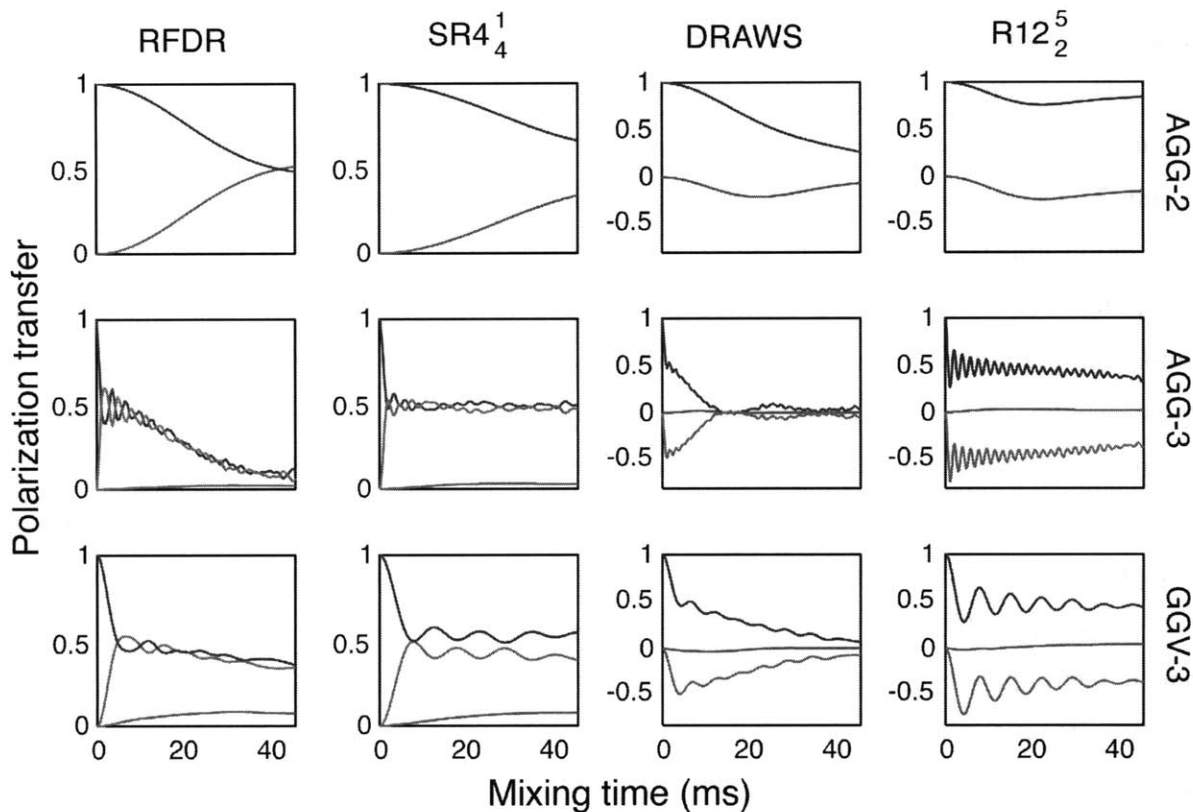
Figure 2. Simulated build-up curves of magnetization transfer from spin 2 (blue) to spin 1 (red) and spin 3 (green) in model spin systems AGG-2, AGG-3 and GGV-3 with various recoupling sequences. Zero-quantum sequences: RFDR, $\omega_r/2\pi = 10$ kHz, $\omega_1 = 50$ kHz; SR4$_4^1$, $\omega_r/2\pi = 30$ kHz, $\omega_1 = 30$ kHz. Double-quantum sequence: R12$_2^5$, $\omega_r/2\pi = 15$ kHz, $\omega_1 = 90$ kHz. Mixed sequence: DRAWS, $\omega_r/2\pi = 10$ kHz, $\omega_1 = 85$ kHz.

Similar results were obtained from simulations of other pulse sequences such as SPC-5 (ref. [11]), POST-C7 (ref. [9]), R16$_2^7$ and R20$_2^9$ (ref. [19]). The pulse sequences presented in Fig. 2 were selected to represent various classes of homonuclear recoupling sequences, namely, zero-quantum π-pulse: RFDR[3,8], zero-quantum symmetry-based: SR4$_4^1$(refs. [19,67]), double-quantum symmetry-based: R12$_2^5$ (refs. [19,68]), and mixed zero- and double-quantum: DRAWS[2,69]. We restricted our investigation to zeroth-order dipolar recoupling sequences that produce a zeroth-order recoupled dipolar Hamiltonian, which are the ones typically used to obtain accurate internuclear distance measurements. We specifically avoided the treatment of proton-driven spin diffusion[70-72], DARR/RAD[73-75], homonuclear TSAR[22,23] techniques and MIRROR[25] recoupling, which perform polarization transfer via higher-order mechanisms that are more tolerant to

124

dipolar-truncation effects in multiple spin systems, as shown previously[22,51], but typically yield ambiguous distance constraints due to the complexity of the mechanisms.

B. Experimental characterization of dipolar truncation in RFDR spectra

We investigated the dipolar truncation effect on our model three-spin systems experimentally using the RFDR pulse sequence[3,8,21]. Polarization transfer efficiencies (defined as cross-peak intensity divided by diagonal intensity at zero mixing time) were measured in a series of broadband 2D correlation spectra with increasing RFDR mixing times following cross-polarization preparation of all ^{13}C nuclei. Experimental data from three representative mixing times are shown in Fig. 3 as cross-sections of 2D spectra through the diagonal peak of spin 2 (the $^{13}C\alpha$). The arrow in Fig. 3 indicates the cross-peak arising from polarization transfer between the weakly coupled spin pair (from spin 2 to spin 1), present in all three samples, while the adjacent cross-peak arises from the coupling that is the source of the dipolar truncation (from spin 2 to spin 3). The cross-sections in Fig. 3 show the sensitivity we were able to obtain with RFDR mixing in these challenging spin systems. Indeed, we attempted similar recoupling experiments and more time-efficient 1D variants with several other pulse sequences often with less satisfactory results. We note that the 4.86 Å internuclear carboxyl/methylene distance of AGG-2 proved challenging for several recoupling sequences due to deleterious effects such as the accumulation of pulse imperfections and magnetization losses due to inadequate heteronuclear decoupling during the long mixing period. In such cases, the additional signal attenuation due to dipolar truncation made the experimental shortcomings only more evident. These technical difficulties, which vary from one recoupling scheme to another, rendered the experimental evaluation and fair comparison of dipolar truncation for several pulse sequences rather difficult. We, therefore, present experimental data for a single recoupling scheme, RFDR, which afforded ample sensitivity to measure long-range polarization transfer and to, therefore, directly assess the dipolar truncation effect.
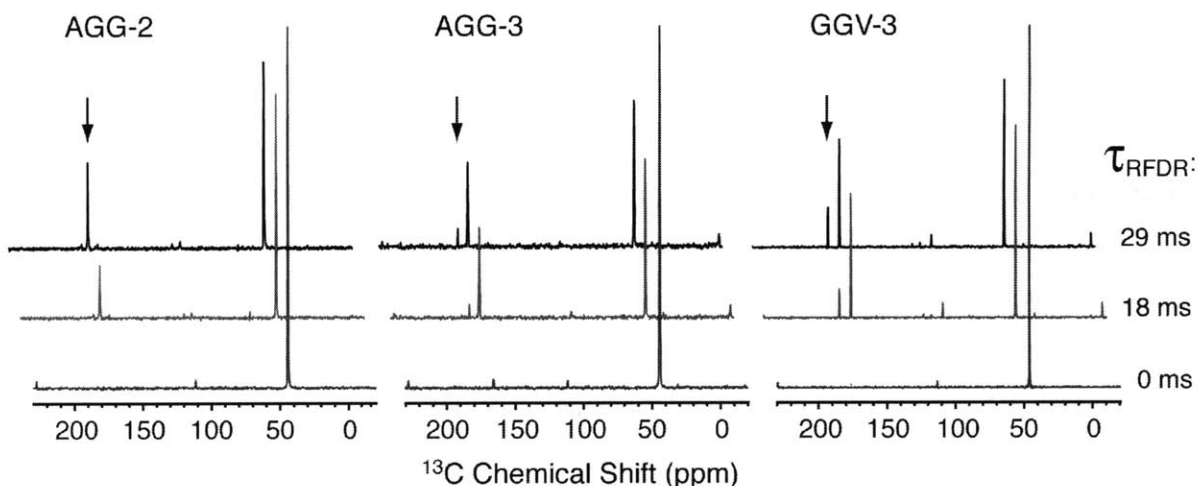
Figure 3. Cross-sections of 2D RFDR correlation spectra through the diagonal peak of spin 2 for three different mixing times. The arrow indicates the crosspeak between spin 2 (a Cα carbon at 44 ppm) and the distant spin 1 (a carboxyl carbon at 175 ppm). The intensity of this crosspeak is diminished in AGG-3 and GGV-3 compared to AGG-2, demonstrating the attenuation of polarization transfer due to the dipolar truncation effect.

Fig. 4 (a-c) presents experimental curves of RFDR polarization transfer in AGG-2, AGG-3 and GGV-3. The data points correspond to integrated peak volumes from 2D spectra for the diagonal ^{13}C nucleus (spin 2) and its crosspeaks to the distant carboxyl nucleus (spin 1) and the nearby carbonyl nucleus (spin 3), normalized to the intensity of the diagonal peak at zero mixing time. Numerical simulations including all parameters used in the experiments, depicted in Fig. 4 (d-f), are in good agreement with our experimental results.

Polarization transfer over the 4.86 Å distance of the two-spin system of AGG-2, shown in Fig. 4(a), was characterized by significant magnetization loss due to insufficient proton decoupling during the mixing period and, to a lesser degree, pulse imperfections. The maximum magnetization transfer observed in this weakly coupled spin pair was approximately 19.0 %, obtained after 29 ms of RFDR mixing, the longest mixing time employed. Simulations for longer mixing periods (to 60 ms) show that the polarization buildup in AGG-2 plateaus after ~30 ms.
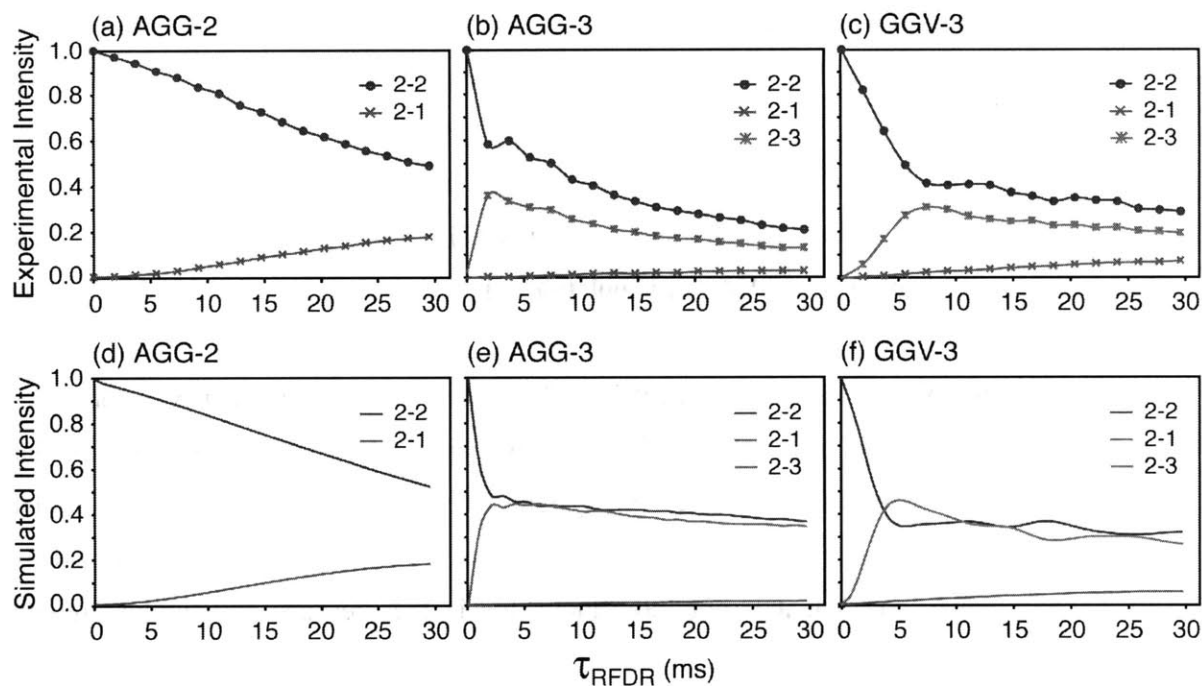
Figure 4. Top: Experimental RFDR buildup curves for AGG-2, AGG-3, and GGV-3. Each data point is the integrated peak intensity of 2D correlation peaks normalized to the diagonal peak at zero mixing. RFDR experiments employed 50 kHz ^{13}C π pulses and 115 kHz and 90 kHz ^{1}H decoupling fields during the mixing period and were performed at a MAS frequency of 8.929 kHz. Bottom: Numerical simulations of polarization transfer in AGG-2, AGG-3, and GGV-3. These simulations included ^{13}C CSA parameters, eight neighboring ^{1}H spins, and experimental parameters identical to those performed in our experiments.

In the three-spin system of AGG-3, polarization transfer over the weakly coupled spin pair is considerably decreased compared to that of AGG-2, experimentally demonstrating the effect of dipolar truncation. The RFDR buildup curve of AGG-3, shown in Fig. 4(b), presents a maximum transfer efficiency of 2.7 % between spins 1 and 2, with a direct dipolar coupling of 66 Hz. There is of course also relay polarization transfer from spin 2 via spin 3 to spin 1 which may be faster than the direct transfer from spin 2 to spin 1. Thus, the strong one-bond dipolar coupling between spins 2 and 3 (2.15 kHz) dominates the RFDR recoupling dynamics and causes a seven-fold reduction in transfer efficiency between the distant spin pair.

Compared to AGG-3, Fig. 4(c) shows that the effect of dipolar truncation is diminished in the GGV-3 spin system, in which the RFDR recoupling dynamics are dominated by a dipolar coupling of moderate magnitude (two-bond coupling, 550 Hz) instead of the one-bond coupling

present in AGG-3. The maximum polarization transfer achieved in the weakly coupled spin pair (spins 1 and 2) of GGV-3 was 7.0 %, almost a factor of 3 smaller than the results for AGG-2 but a considerable improvement in long-range polarization transfer over the AGG-3 case. Even though the weak coupling in GGV-3 is slightly larger than that of AGG-3 (80 and 66 Hz respectively), the increased transfer efficiency is primarily the result of the smaller dipolar coupling effecting truncation in GGV-3, as simulations similar to those in Fig. 2 indicate (data not shown).

However, in spite of the improvement in GGV-3 over AGG-3, the approximately two-fold reduction in long-range recoupling efficiency compared to AGG-2 shows that dipolar truncation is still significant in the GGV-3 spin system. These results demonstrate that dipolar truncation of weak couplings is diminished in the absence of the strong couplings formed by directly bonded nuclei, and yet, it is not fully attenuated in the presence of medium-strength couplings such as those arising from two-bond ^{13}C-^{13}C interactions. Therefore, the majority of homonuclear recoupling schemes can be expected to suffer partial dipolar truncation during long-range ^{13}C polarization transfer in proteins whether they are enriched uniformly or with the alternating scheme, although to different degrees. On the other hand, the attenuation of dipolar truncation by removal of one-bond couplings yields favorable spin systems and allows efficient long-range recoupling, as exemplified by our RFDR results.

4.5 Conclusions

We have demonstrated the effect of dipolar truncation in model three-spin systems and shown its dependence on the magnitude of the dominating dipolar coupling. The samples we used were selected to model spin systems encountered in proteins produced with uniformly and alternating labeling. Our results indicate that the deleterious effect of dipolar truncation in many zeroth-order homonuclear recoupling schemes is moderately alleviated, yet it is still significant, in the alternating labeling scenario. Therefore, while alternating labeling should facilitate the application of schemes that typically show dipolar truncation effects it does not represent a complete solution to the problem.

As our experimental results suggest, the inefficiency of long-range dipolar recoupling with existing recoupling schemes is due in part to the limitations typically associated with implementing lengthy mixing times, such as intense heteronuclear decoupling fields. In this regard, recoupling techniques that function well in the absence of ^{1}H decoupling[16,20,21] may offer an important improvement over other schemes. Beyond practical considerations, dipolar truncation is a fundamental problem in homonuclear dipolar recoupling with a severe deleterious effect, particularly in uniformly labeled systems. Addressing dipolar truncation has led to the design of a variety of interesting experimental schemes[29,46,76-79] and remains an active area of research in our and other laboratories. Finally it should be noted that, in the large spin systems of proteins[80], the multitude of dipolar interactions can be expected to present complexities in addition to dipolar truncation (e.g., relayed polarization transfer) that further complicate the analysis of dipolar recoupling experiments in such systems and will need to be addressed in future work.

References

1. Griffin, R. G. (1998). Dipolar recoupling in MAS spectra of biological solids. *Nature Struct. Biol.* **5**, 508-512.

2. Gregory, D. M., Mitchell, D. J., Stringer, J. A., Kiihne, S., Shiels, J. C., Callahan, J., Mehta, M. A. & Drobny, G. P. (1995). Windowless dipolar recoupling: the detection of weak dipolar couplings between spin 1/2 nuclei with large chemical shift anisotropies. *Chem. Phys. Lett.* **246**, 654-663.

3. Bennett, A. E., Griffin, R. G., Ok, J. H. & Vega, S. (1992). Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys.* **96**, 8624-8627.

4. Nielsen, N. C., Bildsoe, H., Jakobsen, H. J. & Levitt, M. H. (1994). Double-quantum homonuclear rotary resonance: efficient dipolar recovery in magic-angle spinning nuclear magnetic resonance. *J. Chem. Phys.* **101**, 1805-1812.

5. Sun, B., Costa, P. R., Kocisko, D., Lansbury, J. & Griffin, R. G. (1995). Internuclear distance measurements in solid state nuclear magnetic resonance: Dipolar recoupling via rotor synchronized spin locking. *J. Chem. Phys.* **102**, 702-707.

6. Lee, Y. K., Kurur, N. D., Helmle, M., Johannessen, O. G., Nielsen, N. C. & Levitt, M. H. (1995). Efficient dipolar recoupling in the NMR of rotating solids. A sevenfold symmetric radiofrequency pulse sequence. *Chem. Phys. Lett.* **242**, 304-309.

7. Verel, R., Baldus, M., Nijman, M., van Os, J. W. M. & Meier, B. H. (1997). Adiabatic homonuclear polarization transfer in magic-angle-spinning solid-state NMR. *Chem. Phys. Lett.* **280**, 31-39.

8. Bennett, A. E., Rienstra, C. M., Griffiths, J. M., Zhen, W., Lansbury, J. & Griffin, R. G. (1998). Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys.* **108**, 9463-9479.

9. Hohwy, M., Jakobsen, H. J., Eden, M., Levitt, M. H. & Nielsen, N. C. (1998). Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: A compensated C7 pulse sequence. *J. Chem. Phys.* **108**, 2686-2694.

10. Rienstra, C., Hatcher, M., Mueller, L., Sun, B., Fesik, S. & Griffin, R. (1998). Efficient Multispin Homonuclear Double-Quantum Recoupling for Magic-Angle Spinning NMR: 13C-13C Correlation Spectroscopy of U-13C-Erythromycin A. *J. Am. Chem. Soc.* **120**, 10602-10612.

11. Hohwy, M., Rienstra, C. M., Jaroniec, C. P. & Griffin, R. G. (1999). Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy. *J. Chem. Phys.* **110**, 7983-7992.

12. Brinkmann, A., Eden, M. & Levitt, M. H. (2000). Synchronous helical pulse sequences in magic-angle spinning nuclear magnetic resonance: Double quantum recoupling of multiple-spin systems. *J. Chem. Phys.* **112**, 8539-8554.

13. Verel, R., Ernst, M. & Meier, B. H. (2001). Adiabatic Dipolar Recoupling in Solid-State NMR: The DREAM Scheme. *J. Magn. Reson.* **150**, 81-99.

14. Ishii, Y. (2001). 13C-13C dipolar recoupling under very fast magic angle spinning in solid-state nuclear magnetic resonance: Applications to distance measurements, spectral assignments, and high-throughput secondary-structure determination. *J. Chem. Phys.* **114**, 8473-8483.

15. Hohwy, M., Rienstra, C. M. & Griffin, R. G. (2002). Band-selective homonuclear dipolar recoupling in rotating solids. *J. Chem. Phys.* **117**, 4973-4987.

16. De Paëpe, G., Bayro, M. J., Lewandowski, J. & Griffin, R. G. (2006). Broadband homonuclear correlation spectroscopy at high magnetic fields and MAS frequencies. *J. Am. Chem. Soc.* **128**, 1776-7.

17. Bennett, A. E., Griffin, R. G. & Vega, S. (1994). Recoupling of homo- and heteronuclear dipolar interactions in rotating solids. *NMR Basic Principles Progress* **33**, 1-77.

18. Dusold, S. & Sebald, A. (2000). Dipolar recoupling under magic-angle spinning conditions. *Ann. Rep. NMR Spectr.* **41**, 185-264.

19. Levitt, M. H. (2002). Symmetry-based pulse sequences in magic-angle spinning solid-stater NMR. In *Encyclopedia of NMR* (Grant, D. M. & Harris, R. K., eds.), Vol. 9, pp. 165-196. John Wiley & Sons, Ltd., Chichester.

20. De Paëpe, G., Lewandowski, J. R. & Griffin, R. G. (2008). Spin dynamics in the modulation frame: Application to homonuclear recoupling in magic angle spinning solid-state NMR. *J. Chem. Phys.* **128**, 124503.

21. Bayro, M. J., Ramachandran, R., Caporini, M. A., Eddy, M. T. & Griffin, R. G. (2008). Radio frequency-driven recoupling at high magic-angle spinning frequencies: Homonuclear recoupling sans heteronuclear decoupling. *J Chem Phys* **128**, 052321.

22. Lewandowski, J. R., De Paëpe, G. & Griffin, R. G. (2007). Proton assisted insensitive nuclei cross polarization. *J. Am. Chem. Soc* **129**, 728-9.

23. De Paëpe, G., Lewandowski, J. R., Loquet, A., Böckmann, A. & Griffin, R. G. (2008). Proton Assisted Homonuclear Recoupling and Protein Structure Determination. *J. Chem. Phys.* **129**, 245101.

24. Scholz, I., Meier, B. H. & Ernst, M. (2007). Operator-based triple-mode Floquet theory in solid-state NMR. *J Chem Phys* **127**, 204504.

25. Scholz, I., Huber, M., Manolikas, T., Meier, B. H. & Ernst, M. (2008). MIRROR recoupling and its application to spin diffusion under fast magic-angle spinning. *Chem Phys Lett* **460**, 278-283.

26. Creuzet, F., McDermott, A., Gebhard, R., van der Hoef, K., Spijker-Assink, M. B., Herzfeld, J., Lugtenburg, J., Levitt, M. H. & Griffin, R. G. (1991). Determination of membrane protein structure by rotational resonance NMR: bacteriorhodopsin. *Science* **251**, 783-6.

27. Thompson, L. K., McDermott, A. E., Raap, J., van der Wielen, C. M., Lugtenburg, J., Herzfeld, J. & Griffin, R. G. (1992). Rotational resonance NMR study of the active site structure in bacteriorhodopsin: conformation of the Schiff base linkage. *Biochemistry* **31**, 7931-8.

28. McDermott, A. E., Creuzet, F., Gebhard, R., van der Hoef, K., Levitt, M. H., Herzfeld, J., Lugtenburg, J. & Griffin, R. G. (1994). Determination of internuclear distances and the orientation of functional groups by solid-state NMR: rotational resonance study of the conformation of retinal in bacteriorhodopsin. *Biochemistry* **33**, 6129-36.

29.   Costa, P. R., Sun, B. & Griffin, R. G. (1997). Rotational resonance tickling: Accurate internuclear distance measurement in solids. *J. Am. Chem. Soc.*

30.   Verdegem, P. J. E., Helmle, M., Lugtenburg, J. & de Groot, H. (1997). Internuclear Distance Measurements up to 0.44 nm for Retinals in the Solid State with 1-D rotational resonance 13C MAS NMR Spectroscopy. *J. Am. Chem. Soc* **119**, 169-174.

31.   Benzinger, T. L., Gregory, D. M., Burkoth, T. S., Miller-Auer, H., Lynn, D. G., Botto, R. E. & Meredith, S. C. (1998). Propagating structure of Alzheimer's beta-amyloid(10-35) is parallel beta-sheet with residues in exact register. *Proc. Nat. Acad. Sci. USA* **95**, 13407-12.

32.   Feng, X., Verdegem, P. J., Lee, Y. K., Helmle, M., Shekar, S. C., de Groot, H. J., Lugtenburg, J. & Levitt, M. H. (1999). Rotational resonance NMR of 13C2-labelled retinal: quantitative internuclear distance determination. *Solid State Nucl. Magn. Reson.* **14**, 81-90.

33.   Griffiths, J. M., Bennett, A. E., Engelhard, M., Siebert, F., Raap, J., Lugtenburg, J., Herzfeld, J. & Griffin, R. G. (2000). Structural investigation of the active site in bacteriorhodopsin: geometric constraints on the roles of Asp-85 and Asp-212 in the proton-pumping mechanism from solid state NMR. *Biochemistry* **39**, 362-71.

34.   Carravetta, M., Eden, M., Johannessen, O. G., Luthman, H., Verdegem, P. J., Lugtenburg, J., Sebald, A. & Levitt, M. H. (2001). Estimation of carbon-carbon bond lengths and medium-range internuclear distances by solid-state nuclear magnetic resonance. *J. Am. Chem. Soc* **123**, 10628-38.

35.   Balbach, J. J., Petkova, A. T., Oyler, N. A., Antzutkin, O. N., Gordon, D. J., Meredith, S. C. & Tycko, R. (2002). Supramolecular structure in full-length Alzheimer's beta-amyloid fibrils: evidence for a parallel beta-sheet organization from solid-state nuclear magnetic resonance. *Biophys. J.* **83**, 1205-16.

36.   Karlsson, T., Popham, J. M., Long, J. R., Oyler, N. & Drobny, G. P. (2003). A study of homonuclear dipolar recoupling pulse sequences in solid-state nuclear magnetic resonance. *J. Am. Chem. Soc* **125**, 7394-407.

37.   Kristiansen, P. E., Carravetta, M., van Beek, J. D., Lai, W. C. & Levitt, M. H. (2006). Theory and applications of supercycled symmetry-based recoupling sequences in solid-state nuclear magnetic resonance. *J. Chem. Phys.* **124**, 234510.

38.   Raleigh, D. P., Levitt, M. H. & Griffin, R. G. (1988). Rotational resonance in solid state NMR. *Chem. Phys. Lett.*

39.   Raleigh, D. P., Creuzet, F., Das Gupta, S. K., Levitt, M. H. & Griffin, R. G. (1989). Measurement of internuclear distances in polycrystalline solids. Rotationally enhanced transfer of nuclear spin magnetization. *J. Am. Chem. Soc.*

40.   Takegoshi, K., Nomura, K. & Terao, T. (1995). Rotational resonance in the tilted rotating frame. *Chem. Phys. Lett.* **232**, 424-428.

41.   Costa, P. R., Kocisko, D. A., Sun, B. Q., Lansbury, P. T. & Griffin, R. G. (1997). Determination of Peptide Amide Configuration in a Model Amyloid Fibril by Solid-State NMR. *J. Am. Chem. Soc* **43**, 10487-10493.

42.   Costa, P. R., Sun, B. & Griffin, R. G. (2003). Rotational resonance NMR: separation of dipolar coupling and zero quantum relaxation. *J. Magn. Reson.* **164**, 92-103.

43.   Ramachandran, R., Ladizhansky, V., Bajaj, V. S. & Griffin, R. G. (2003). 13C-13C rotational resonance width distance measurements in uniformly 13C-labeled peptides. *J. Am. Chem. Soc* **125**, 15623-9.

44.   Williamson, P. T., Verhoeven, A., Ernst, M. & Meier, B. H. (2003). Determination of internuclear distances in uniformly labeled molecules by rotational-resonance solid-state NMR. *J. Am. Chem. Soc* **125**, 2718-22.

45.   Verhoeven, A., Williamson, P. T., Zimmermann, H., Ernst, M. & Meier, B. H. (2004). Rotational-resonance distance measurements in multi-spin systems. *J. Magn. Reson.* **168**, 314-26.

46.   Ramachandran, R., Lewandowski, J. R., van der Wel, P. C. & Griffin, R. G. (2006). Multipole-multimode Floquet theory of rotational resonance width experiments: 13C-13C distance measurements in uniformly labeled solids. *J. Chem. Phys.* **124**, 214107.

47. Williamson, P. T., Verhoeven, A., Miller, K. W., Meier, B. H. & Watts, A. (2007). The conformation of acetylcholine at its target site in the membrane-embedded nicotinic acetylcholine receptor. *Proc. Nat. Acad. Sci. USA* **104**, 18031-6.

48. Peng, X., Libich, D., Janik, R., Harauz, G. & Ladizhansky, V. (2008). Dipolar chemical shift correlation spectroscopy for homonuclear carbon distance measurements in proteins in the solid state: application to structure determination and refinement. *J. Am. Chem. Soc* **130**, 359-69.

49. Kiihne, S., Mehta, M. A., Stringer, J. A. & Gregory, D. M. (1998). Distance measurements by dipolar recoupling two-dimensional solid-state NMR. *J. Phys. Chem. A*.

50. Hodgkinson, P. & Emsley, L. (1999). The accuracy of distance measurements in solid-state NMR. *J. Magn. Reson.* **139**, 46-59.

51. Grommek, A., Meier, B. H. & Ernst, M. (2006). Distance information from proton-driven spin diffusion under MAS. *Chem. Phys. Lett.* **427**, 404-409.

52. Costa, P. R. (1996). Spins, Peptides, and Alzheimer's Disease: Solid State NMR Investigations of Amyloid Peptide Conformation. PhD, Massachusetts Institute of Technology.

53. LeMaster, D. & Kushlan, D. (1996). Dynamical mapping of E. coli Thioredoxin via 13C NMR relaxation analysis. *J. Am. Chem. Soc.* **118**, 9255-9264.

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

55. Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K. & Oschkinat, H. (2002). Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **420**, 98-102.

56. Zech, S. G., Wand, A. J. & McDermott, A. E. (2005). Protein structure determination by high-resolution solid-state NMR spectroscopy: application to microcrystalline ubiquitin. *J. Am. Chem. Soc* **127**, 8618-26.

57.   Raleigh, D. P., Levitt, M. H. & Griffin, R. G. (1988). Rotational resonance in solid state NMR. *Chem. Phys. Lett.* **146**, 71-76.

58.   Meier, B. H. (1994). Polarization Transfer and Spin Diffusion in Solid-State NMR. *Advances in Magnetic and Optical Resonance* **18**, 1-116.

59.   Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V. & Griffin, R. G. (1995). Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* **103**, 6951-6958.

60.   Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-93.

61.   Subramanian, E. & Lalitha, V. (1983). Crystal structure of a tripeptide, L-alanyl-glycyl-glycine and its relevance to the poly(glycine)-II type of conformation. *Biopolymers* **22**, 833-838.

62.   Lalitha, V., Subramanian, E. & Bordner, J. (1984). Structure and conformation of linear peptides. III. Structure of glycyl-glycyl-L-valine. *Int J Pept Protein Res* **24**, 437-41.

63.   Smith, S. A., Levante, T. O., Meier, B. H. & Ernst, R. R. (1994). Computer Simulations in Magnetic Resonance. An Object-Oriented Programming Approach. *J. Magn. Reson.* **106**, 75-105.

64.   Cheng, V. B., Suzukawa, J. & Wolfsberg, M. (1973). Investigations of a nonrandom numerical method for multidimensional integration. *J. Chem. Phys.* **59**, 3992-3999.

65.   Conroy, H. (1967). Molecular Schrodinger Equation. VIII. A New Method for the Evaluation of Multidimensional Integrals. *J. Chem. Phys.* **47**, 5307-5318.

66.   Bak, M., Schultz, R., Vosegaard, T. & Nielsen, N. C. (2002). Specification and visualization of anisotropic interaction tensors in polypeptides and numerical simulations in biological solid-state NMR. *J. Magn. Reson.* **154**, 28-45.

67.   Veshtort, M. & Griffin, R. G. (2006). SPINEVOLUTION: A powerful tool for the simulation of solid and liquid state NMR experiments. *J. Magn. Reson.* **178**, 248-282.

68.   Brinkmann, A., Schmedt auf der Günne, J. & Levitt, M. H. (2002). Homonuclear zero-quantum recoupling in fast magic-angle spinning nuclear magnetic resonance. *J. Magn. Reson.* **156**, 79-96.

69. Marin-Montesinos, I., Brouwer, D. H., Antonioli, G., Lai, W. C., Brinkmann, A. & Levitt, M. H. (2005). Heteronuclear decoupling interference during symmetry-based homonuclear recoupling in solid-state NMR. *J. Magn. Reson.* **177**, 307-17.

70. Mehta, M. A., Gregory, D. M., Kiihne, S., Mitchell, D. J., Hatcher, M. E., Shiels, J. C. & Drobny, G. P. (1996). Distance measurements in nucleic acids using windowless dipolar recoupling solid state NMR. *Solid State Nucl. Magn. Reson.* **7**, 211-28.

71. Bloembergen, N. (1949). On the interaction of nuclear spins in a crystalline lattice. *Physica* **15**, 386-426.

72. Szeverenyi, N. M., Sullivan, M. J. & Maciel, G. E. (1982). Observation of spin exchange by two-dimensional fourier transform 13C cross polarization-magic-angle spinning. *J. Magn. Reson.* **47**, 462-475.

73. Meier, B. H. (1999). Polarization transfer and spin diffusion in solid-state NMR. In *Advances in Magnetic and Optical Resonance* (Warren, W. S., ed.), Vol. 18, pp. 1. Academic Press, New York.

74. Takegoshi, K., Nakamura, S. & Terao, T. (2001). 13C-1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.* **344**, 631-637.

75. Takegoshi, K., Nakamura, S. & Terao, T. (2003). 13C-1H dipolar-driven 13C-13C recoupling without 13C rf irradiation in nuclear magnetic resonance of rotating solids. *J. Chem. Phys.* **118**, 2325-2341.

76. Morcombe, C. R., Gaponenko, V., Byrd, R. A. & Zilm, K. W. (2004). Diluting abundant spins by isotope edited radio frequency field assisted diffusion. *J. Am. Chem. Soc.* **126**, 7196-7.

77. Marin-Montesinos, I., Mollica, G., Carravetta, M., Gansmuller, A., Pileio, G., Bechmann, M., Sebald, A. & Levitt, M. H. (2006). Truncated dipolar recoupling in solid-state nuclear magnetic resonance. *Chem. Phys. Lett.* **432**, 572-578.

78. Khaneja, N. & Nielsen, N. C. (2008). Triple oscillating field technique for accurate distance measurements by solid-state NMR. *J. Chem. Phys.* **128**, 015103.

79. Pileio, G., Guo, Y., Pham, T. N., Griffin, J. M., Levitt, M. H. & Brown, S. P. (2007). Residual Dipolar Couplings by Off-Magic-Angle Spinning in Solid-State Nuclear Magnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* **129**, 10972-10973.

80. Pileio, G., Mamone, S., Mollica, G., Montesinos, I. M., Gansmuller, A., Carravetta, M., Brown, S. P. & Levitt, M. H. (2008). Estimation of internuclear couplings in the solid-state NMR of multiple-spin systems. Selective spin echoes and off-magic-angle sample spinning. *Chem. Phys. Lett.* **456**, 116-121.

81. Gehman, J. D., Paulson, E. K. & Zilm, K. W. (2003). The influence of internuclear spatial distribution and instrument noise on the precision of distances determined by solid state NMR of isotopically enriched proteins. *J. Biomol. NMR* **27**, 235-59.

Chapter 5. Highly Efficient Homonuclear Dipolar Recoupling in Protein MAS NMR Spectroscopy: The BASE RFDR Scheme

Summary

We describe a low-power, band-selective version of radio frequency-driven recoupling, termed BASE RFDR, that produces highly efficient homonuclear polarization transfer in magic-angle spinning NMR experiments. The pulse sequence was optimized for nuclear resonances with small frequency offsets, taking into consideration favorable finite-pulse effects and effective decoupling of heteronuclear interactions. We show that, when applied to the aliphatic ^{13}C region of proteins produced with the alternating ^{13}C-^{12}C labeling scheme, BASE RFDR yields highly sensitive cross-peaks between distant nuclei via direct reintroduction of ^{13}C dipole-dipole couplings. The applications of this technique include resonance assignment and generation of long-range constraints. In particular, it is a powerful complement to long-range polarization transfer schemes based on indirect, higher-order dipolar interactions.

5.1 Introduction

Solid-state NMR spectroscopy is a powerful technique for the investigation of complex biological systems such as membrane proteins and amyloid fibrils. In magic-angle spinning (MAS) NMR, structural information is obtained via the reintroduction of anisotropic interactions.[1] In particular, a large number of homonuclear dipolar recoupling schemes have been developed and applied to record correlation spectra and measure internuclear distances in peptides and proteins.[1] While many recoupling techniques efficiently transfer polarization between directly bonded ^{13}C nuclei, their effectiveness can be reduced significantly when distant ^{13}C spins with weak dipolar couplings are involved. This limitation is generally imposed both by the experimental constraints that must be maintained during long mixing periods and by the inherent complexities of multiple-spin systems such as dipolar truncation,[2, 3] that is the attenuation of weak dipolar couplings by stronger dipolar couplings in the recoupled dipolar Hamiltonian. As a consequence, spin diffusion techniques,[4-8] which circumvent some of these limitations,[9] have been widely utilized to estimate long-range homonuclear structural constraints in protein solid-state NMR studies.[10, 11] In this chapter we present an experimental approach that provides highly sensitive long-range correlations between aliphatic ^{13}C nuclei through the combination of isotope dilution and efficient polarization transfer via a band selective radio frequency dipolar recoupling (BASE RFDR) scheme. We demonstrate this method with a sample of PI3-SH3 (the SH3 domain of the p85α subunit of phosphatidylinositol 3 kinase) in amyloid fibril form.

5.2 Results and Discussion

The ^{13}C alternating labeling scheme devised by LeMaster and Kushlan,[12] and introduced to solid-state NMR by Hong,[13] which employs 2-^{13}C$_1$ or 1,3-^{13}C$_2$ glycerol as the carbon source, results in amino acids ^{13}C labeled at approximately every other carbon position ("odd/even" labeling). The approach simplifies ^{13}C-^{13}C correlation spectra by reducing the number of spectral lines and narrows the linewidths due to the elimination of J-couplings. It thus offers significant advantages over uniform ^{13}C labeling.[10] From the perspective of the dipolar recoupling dynamics of the spin system, this type of ^{13}C spin dilution provides two additional benefits. First, relayed polarization transfer, which is dominated by rapid diffusion between directly bonded nuclei in uniformly ^{13}C labeled samples, is partially eliminated in odd/even ^{13}C labeled spin systems; magnetization originating from a given site therefore propagates more directly to fewer spins, resulting in correlation spectra that present improved long-range transfer efficiency between structurally interesting pairs of nuclei. Secondly, the elimination of directly bonded ^{13}C nuclei in most amino acid spin systems results in the partial attenuation of dipolar truncation effects, thus enabling the observation of weak dipolar couplings corresponding to long internuclear distances. However, dipolar truncation is most severe in pulse sequences that generate a first-order recoupled dipolar Hamiltonian and is considerably less pronounced in schemes that utilize second order effects to achieve polarization transfer.[9, 14, 15] Therefore, the enhanced observation of long-range correlations in spin diffusion spectra of odd/even labeled proteins (compared to uniformly labeled samples) is likely due to the reduction of relayed polarization transfer and not to the attenuation of dipolar truncation in these dilute spin systems. In contrast, as examined recently,[3] certain zero-quantum (ZQ) recoupling schemes may benefit directly from the attenuation of dipolar truncation in such spin systems.

141
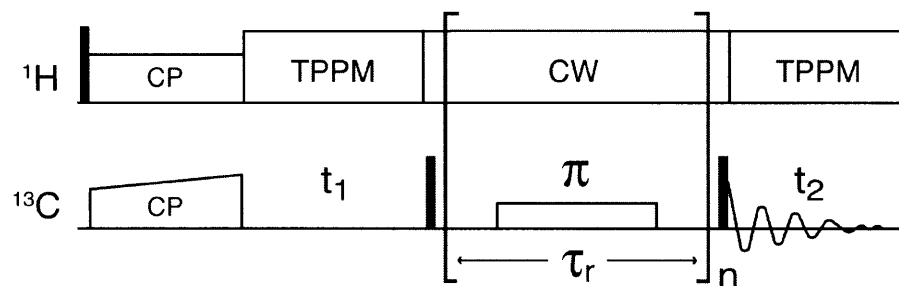
Figure 1. Band-selective radio frequency-driven recoupling (BASE RFDR) pulse sequence for two-dimensional homonuclear correlation spectroscopy. The recoupling π pulses occupy half or more of the rotor period, τ_r, and follow a 32-step phase scheme. The basic building block of RFDR experiments consists of two rotor periods.

In order to explore the advantages of alternating labeling for techniques other than spin diffusion, we investigated the application of efficient recoupling methods to a fibril sample of PI3-SH3 (86 residues) prepared with [2-^{13}C] glycerol as the carbon source (2-PI3-SH3). As may be expected, the elimination of directly bonded nuclei interrupts relayed polarization transfer and attenuates dipolar truncation in broadband radio frequency-driven recoupling[16-20] (RFDR) experiments. The polarization transfer dynamics in this ZQ recoupling approach are heavily dominated by the strongest couplings present, which in the case of 2-^{13}C glycerol labeling are typically the two-bond couplings between sequential ^{13}Cα(i) and ^{13}C'(i-1) resonances. As a consequence, the most prominent cross-peaks in broadband RFDR spectra of 2-PI3-SH3 are medium- and long-range correlations between C' and aliphatic nuclei, while correlations between distant aliphatic nuclei, rich in structural information, are generally too weak to be observed. A representative broadband RFDR spectrum is shown in the supporting information.

In constrast, the application of RFDR using low-power, band selective, π pulses (shown in Figure 1) with the minimal bandwidth necessary to cover the aliphatic ^{13}C spectral region yields intense correlation peaks between distant aliphatic nuclei, as illustrated in Figures 2 and 3. This spectrum was recorded at a spinning frequency ($\omega_r/2\pi$) of 12.5 kHz and employed ^{13}C recoupling π pulses with a 12.5 kHz rf field ($\omega_1/2\pi$) in a manner similar in appearance to conventional RFDR. During the mixing period (τ_{mix} = 17.92 ms), a 80 kHz ^{1}H cw decoupling field was applied and the ^{13}C recoupling π pulses followed a 32-step phase sequence (XY-16, Y$\overline{X}$ -16) optimal for compensation of chemical shift offsets and rf inhomogeneity for weak π pulses that is based on the XY-16 scheme.[21]
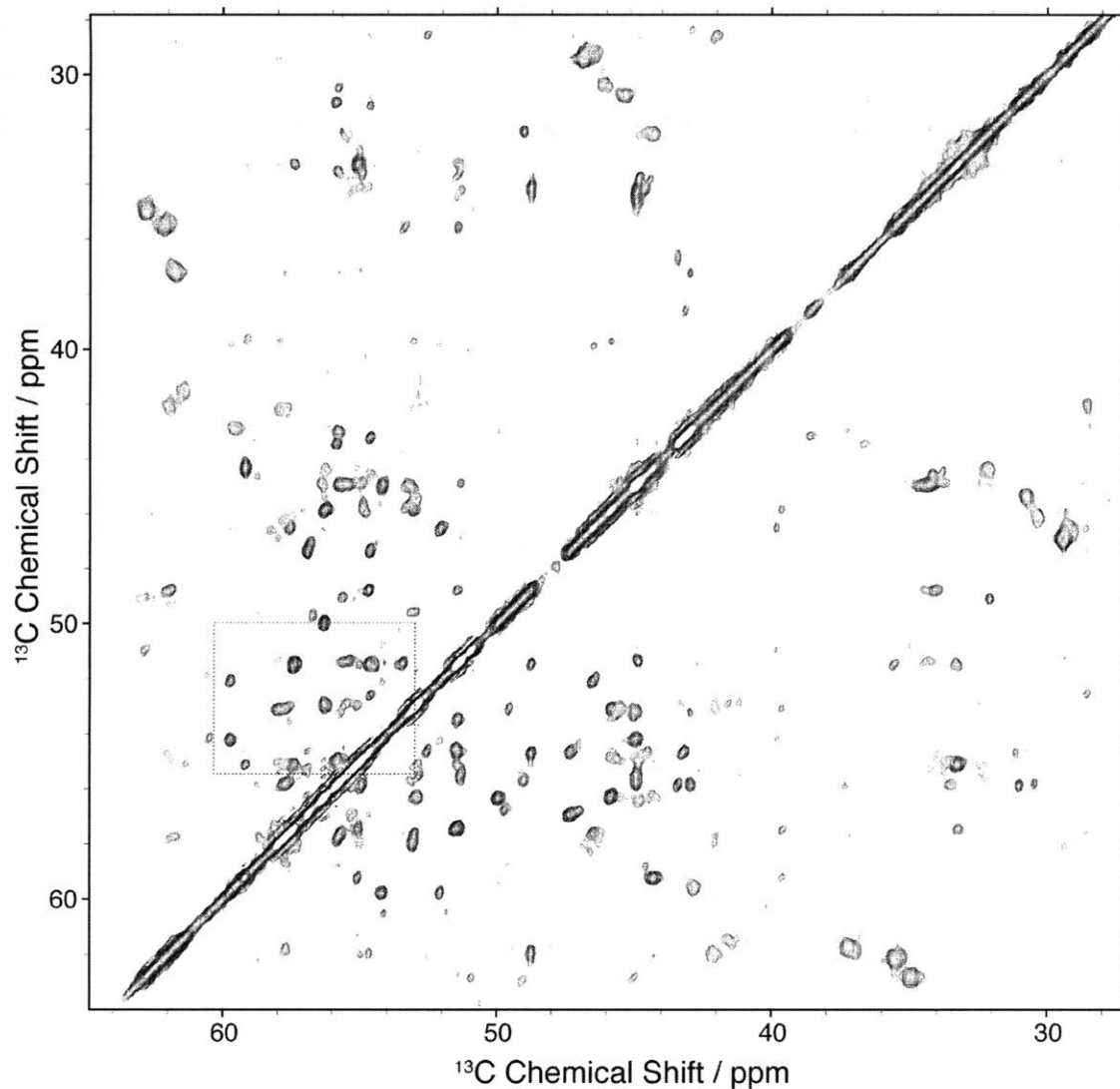
Figure 2. Long-range aliphatic correlation spectrum of 2-PI3-SH3 amyloid fibrils obtained with 17.92 ms of longitudinal mixing via BASE RFDR. The mixing time was optimized for 3.8 Å, typical internuclear distance of sequential Cα-Cα contacts. The spectrum was recorded in 27 hrs at 16.4 T magnetic field and 12.5 kHz MAS frequency, with 12.5 kHz ^{13}C recoupling π pulses centered at 48 ppm. The dotted box indicates peaks shown in Fig. 3.

The long-range polarization transfer efficiency of this BASE RFDR approach is a product of the interplay between several factors in addition to the attenuation of dipolar truncation afforded by alternating labeling. First, the selective bandwidth of the recoupling π pulses centered on the aliphatic region of the ^{13}C spectrum exclude downfield resonances from the

aliphatic-aliphatic recoupling dynamics, resulting in the elimination of relayed polarization transfer via neighboring C' spins, which enhances the direct polarization transfer between distant aliphatic ^{13}C spins. Concurrently, the application of long ^{13}C π pulses occupying a large fraction of the rotor period entails a significant finite-pulse effect, even for moderate spinning frequencies, which facilitates dipolar recoupling between nuclei with small chemical shift offsets such as aliphatic spins.[18, 19] Furthermore, heteronuclear interference conditions[20] are readily avoided by the application of moderate to strong ^{1}H decoupling fields due to the weak effective ^{13}C rf field employed during mixing, leading to minimal polarization losses during the recoupling period. Finally, numerical simulations indicate that while polarization transfer in BASE RFDR is primarily mediated by direct ^{13}C-^{13}C dipolar interactions, third-spin assisted recoupling[14, 15] (TSAR), which relies on higher-order heteronuclear interactions, is active during these (^{13}C/^{1}H) rf field conditions and is compatible with the longitudinal mixing mechanism of BASE RFDR, potentially enhancing the attainable polarization transfer efficiency. Numerical simulations illustrating finite-pulse and TSAR effects are provided in the supplementary information.
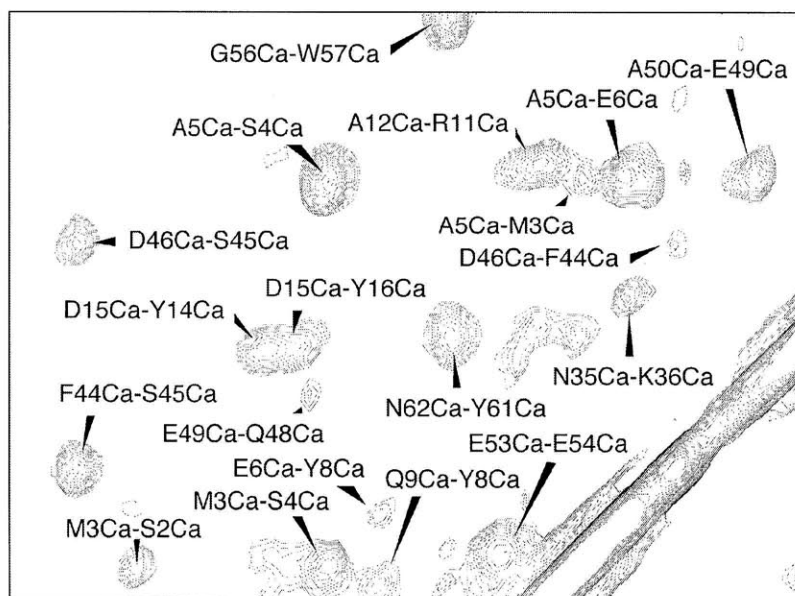
Figure 3. Section of the BASE RFDR spectrum in Figure 2, showing assignments for multiple cross-peaks. At the mixing time employed, 18 ms, intense cross-peaks appear that correspond to sequential Cα- Cα correlations, while cross-peaks between more distant nuclei (> 3.8 Å) appear with lower intenstities.

As may be observed in Figure 4, BASE RFDR leads to improved sensitivity in aliphatic-aliphatic correlation spectra of 2-PI3-SH3 compared to that obtained with proton-driven spin diffusion[5, 6] or DARR[7, 8] for cross-peaks corresponding to internuclear distances of ~3.8 Å and above, arising from sequential inter-residue ^{13}C-^{13}C couplings and longer contacts. However, a few specific exceptions, where cross-peak intensities are attenuated, can be observed most likely as a consequence of strong dipolar truncation effects remaining in certain spin systems in 2-PI3-SH3. We must note that while mixing times in Fig. 2 were optimized for Cα-Cα contacts (~3.8 Å) for direct comparison, in the case of diffusion experiments these cross-peaks build up slowly at such long mixing times (300-500 ms) that longer contacts could potentially be observed.

Optimal experimental conditions for efficient low-power recoupling are determined by the spectral bandwidth required. In the case of aliphatic carbons at 16.4 T (700 MHz ^{1}H Larmor frequency), with a spectral dispersion of ~10 kHz, ^{13}C rf fields of similar magnitude are required for efficient recoupling within the aliphatic region. At the same time, the rotor-synchronized ^{13}C recoupling π pulses must occupy a significant fraction of the rotor period (half or more) in order to produce the most favorable finite-pulse effects. Thus, for 12 kHz ^{13}C π pulses, BASE RFDR can be performed at MAS frequencies ranging from ~12 kHz to 24 kHz, the latter being the windowless recoupling irradiation limit in which the recoupling π pulse occupies the entire rotor period. Conversely, at 12 kHz spinning frequency, BASE RFDR can be applied with ^{13}C rf fields ranging from 6 to 12 kHz.
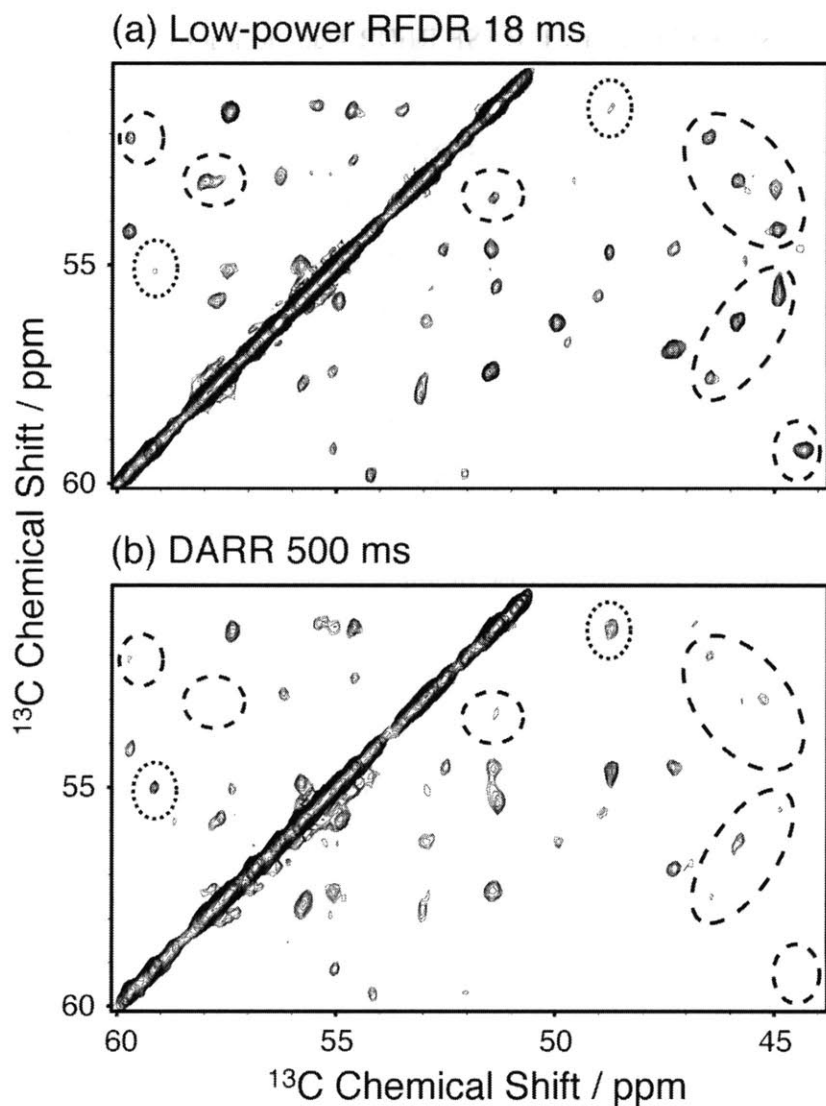
Figure 4. Long-range aliphatic correlations in 2-PI3-SH3 obtained with 17.92 ms of BASE RFDR (a) and 500 ms of DARR (b). Both spectra were recorded in 6.8 hrs under similar experimental conditions to those described for Figure 2. BASE RFDR shows improved overall polarization transfer efficiency, highlighted by dashed ellipses. Dotted ellipses mark a few exceptions, likely the result of remaining dipolar truncation for certain spin systems. See the supporting information for full spectral views.

5.3 Conclusions

The highly efficient long-range aliphatic-aliphatic polarization transfer approach described in this chapter, and demonstrated in studies of PI3-SH3 amyloid fibrils, has potential applications to a variety of aspects of protein MAS NMR investigations, including sequential resonance assignment, generation of constraints between distant nuclei, and accurate distance measurements. The BASE RFDR pulse scheme is an extension of RFDR[16-20] into a regime where the majority of the rotor period during the mixing time is occupied by weak ^{13}C irradiation, at moderate spinning frequencies. While the efficiency of BASE RFDR has been delineated here and exploited to procure highly sensitive long-range aliphatic correlations, further methodological investigation of this dipolar recoupling scheme is currently in progress.

Obtaining efficient homonuclear correlations is a key step in structure determination via MAS NMR spectroscopy. Correlations between distant ^{13}C nuclei are typically achieved by applying mixing schemes that rely on higher-order effects and are not affected severely by dipolar truncation, such as proton-driven spin diffusion, DARR, proton-assisted recoupling[14], and variations thereof. While these techniques are robust, they are largely qualitative because of the multiple orientation-dependent dipolar inteactions that give rise to their mechanism of polarization transfer. Furthermore, at moderate spinning frequencies (between 8 and 15 kHz), used in many biological MAS NMR applications, PDSD and DARR remain the most efficient higher-order techniques. In this chapter we have demonstrated that the BASE RFDR approach may lead to superior polarization transfer than that achieved with DARR and PDSD in protein samples prepared with alternating labeling. We demonstrated this performance with sequential ^{13}Cα-^{13}Cα contacts with an average internuclear distance of ~3.8 Å, in a one-to-one comparison. This constitutes a significant achievement, since polarization transfer in BASE RFDR is mainly driven by zero-order, direct ^{13}C-^{13}C couplings (as shown in the Appendix) and is inherently susceptible to dipolar truncation. The only other cases in which experiments have been shown to perform better than PDSD at moderate spinning frequencies involve highly mobile sites in proteins such as methyl groups. In addition, the homonuclear dipolar coupling dependence of BASE RFDR transfer signifies that it may be possible to use this technique to measure internuclear distances quantitatively, via mixing time build-up experiments, provided that relaxation effects may be accounted for.

The BASE RFDR scheme has been applied successfully to multiple proteins and is utilized in various investigations described in this thesis. First, recognizing that highly sensitive, sequential $^{13}C\alpha$-$^{13}C\alpha$ correlations were easily obtained with BASE RFDR, we developed a resonance assignment protocol for alternatingly labeled proteins using this experiment as the primary source of sequential contacts, supplementing it with three other experiments to elucidate the majority of ^{13}C and ^{15}N backbone resonances, as described in chapter 6. The BASE RFDR approache also proved very powerful in establishing inter-molecular correlations and in the determination of tertiary structure of amyloid fibrils, as discussed in chapter 8. Finally, chapter 11 describes the observation of backbone dynamics in BASE RFDR correlations, which indicates that it may be possible to apply this approach to the measurement of protein relaxation rates and molecular order parameters.

5.4 Experimental Methods

The sample of 2-PI3-SH3 fibrils was prepared as described previously[22] but using [2-^{13}C]glycerol and $NaH^{13}CO_3$ as the sole sources of carbon. The total amount of protein in each packed sample was approximately 8 mg. All experiments were performed using a custom-designed spectrometer operating at 700 MHz ^1H Larmor frequency (courtesy of Dr. David J. Ruben at the Francis Bitter Magnet Laboratory, Massachusetts Institute of Technology) equipped with a 3.2 mm Varian/Chemagnetics probe (Palo Alto, CA). Correlation experiments consisted of ^1H-^{13}C cross-polarization, chemical shift evolution, longitudinal homonuclear mixing, and detection periods. TPPM[23] decoupling (100 kHz ^1H rf field) was applied during the evolution and detection periods. Such high-power decoupling was desirable for improved resolution, but not strictly necessary for experimental performance. Further data acquisition details are provided in the Appendix below.

References

1.   Griffin, R. G. (1998). Dipolar recoupling in MAS spectra of biological solids. *Nature Struc. Biol.* **5 Suppl**, 508-12.

2.	Costa, P. R. (1996). Spins, Peptides, and Alzheimer's Disease: Solid State NMR Investigations of Amyloid Peptide Conformation. PhD, Massachusetts Institute of Technology.

3.	Bayro, M. J., Huber, M., Ramachandran, R., Davenport, T. C., Meier, B. H., Ernst, M. & Griffin, R. G. (2009). Dipolar truncation in magic-angle spinning NMR recoupling experiments. *J. Chem. Phys.* **130**, 114506.

4.	Bloembergen, N. (1949). On the interaction of nuclear spins in a crystalline lattice. *Physica* **15**, 386-426.

5.	Szeverenyi, N. M., Sullivan, M. J. & Maciel, G. E. (1982). Observation of spin exchange by two-dimensional fourier transform 13C cross polarization-magic-angle spinning. *J. Magn. Reson.* **47**, 462-475.

6.	Meier, B. H. (1999). Polarization transfer and spin diffusion in solid-state NMR. In *Advances in Magnetic and Optical Resonance* (Warren, W. S., ed.), Vol. 18, pp. 1.

7.	Takegoshi, K., Nakamura, S. & Terao, T. (2001). 13C-1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.* **344**, 631-637.

8.	Morcombe, C. R., Gaponenko, V., Byrd, R. A. & Zilm, K. W. (2004). Diluting abundant spins by isotope edited radio frequency field assisted diffusion. *J. Am. Chem. Soc.* **126**, 7196-7.

9.	Grommek, A., Meier, B. H. & Ernst, M. (2006). Distance information from proton-driven spin diffusion under MAS. *Chem. Phys. Lett.* **427**, 404-409.

10.	Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K. & Oschkinat, H. (2002). Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **420**, 98-102.

11.	Manolikas, T., Herrmann, T. & Meier, B. H. (2008). Protein structure determination from 13C spin-diffusion solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **130**, 3959-66.

12.	LeMaster, D. & Kushlan, D. (1996). Dynamical Mapping of E. coli Thioredoxin via 13C NMR Relaxation Analysis. *J. Am. Chem. Soc.* **118**, 9255-9264.

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

14. De Paëpe, G., Lewandowski, J. R., Loquet, A., Böckmann, A. & Griffin, R. G. (2008). Proton Assisted Homonuclear Recoupling and Protein Structure Determination. *J. Chem. Phys.* **129**, 245101.

15. Lewandowski, J. R., De Paëpe, G. & Griffin, R. G. (2007). Proton assisted insensitive nuclei cross polarization. *J. Am. Chem. Soc* **129**, 728-9.

16. Bennett, A. E., Griffin, R. G., Ok, J. H. & Vega, S. (1992). Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys.* **96**, 8624-8627.

17. Bennett, A. E., Rienstra, C. M., Griffiths, J. M., Zhen, W., Lansbury, J. & Griffin, R. G. (1998). Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys.* **108**, 9463-9479.

18. Boender, G. J., Vega, S. & de Groot, H. J. M. (2000). Quantized field description of rotor frequency-driven dipolar recoupling. *J. Chem. Phys.* **112**, 1096-1106.

19. Ishii, Y. (2001). 13C-13C dipolar recoupling under very fast magic angle spinning in solid-state nuclear magnetic resonance: Applications to distance measurements, spectral assignments, and high-throughput secondary-structure determination. *J. Chem. Phys.* **114**, 8473-8483.

20. Bayro, M. J., Ramachandran, R., Caporini, M. A., Eddy, M. T. & Griffin, R. G. (2008). Radio frequency-driven recoupling at high magic-angle spinning frequencies: Homonuclear recoupling sans heteronuclear decoupling. *J. Chem. Phys.* **128**, 052321.

21. Gullion, T., Baker, D. B. & Conradi, M. S. (1990). New, compensated Carr-Purcell sequences. *J. Magn. Reson.* **89**, 479-484.

22. Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D. & Dobson, C. M. (1998). Amyloid Fibril Formation by an SH3 Domain. *Proc. Natl. Acad. Sci. USA* **95**, 4224-4228.

23. Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V. & Griffin, R. G. (1995). Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* **103**, 6951-6958.

Appendix to Chapter 5

5A.1 Finite-pulse effects in RFDR-type recoupling pulse sequences

The original theoretical description of radio frequency-driven recoupling (RFDR) considered infinitely short π pulses, yielding a recoupled dipolar Hamiltonian of purely zero-quantum nature with an effective coupling constant that vanishes for small isotropic chemical shift differences between the interacting pair of spins.[1,2] In practice, it was observed that RFDR correlation spectra exhibit intense crosspeaks between nuclei with small chemical shift differences.[1-4]. This effect, also observed in a related experiment[5], was attributed to the finite duration of the recoupling π pulses. The RFDR experiment was analyzed theoretically in great detail by Boender, et al.[6], showing that the experimental observations that RFDR sequences recouple both closely and widely separated resonances were consistent with their treatment based on Floquet theory. In a subsequent study[7], the effect of finite pulses was shown to become increasingly relevant at high spinning frequencies ($\geq$ 20 kHz). In general, the optimal rf settings for RFDR-type experiments thus depend on chemical shift differences, desired bandwidth, spinning frequency, and minimizing heteronuclear interference[8]. While fast spinning ($\geq$ 20 kHz) offers a number of advantages[8], many solid-state NMR studies require flexibility in the range of MAS frequencies employed. We therefore investigated the effect of finite pulses in RFDR at moderate spinning rates (8 to 16 kHz), and found that in the case of small chemical shift dispersion it is possible to achieve efficient polarization transfer by employing long-pulse or band-selective conditions, specifically those in the regime where the recoupling irradiation occupies between half and the full rotor period, as shown in Figure S1. In the present manuscript we refer to this regime, depicted in the pulse sequence in Figure 1 of the main text, as band selective radio frequency-driven recoupling (BASE RFDR). Polarization is efficiently preserved by the recoupling π pulses while heteronuclear interference during the mixing period is easily avoided with a large ratio between ^{1}H and ^{13}C irradiation, making the approach ideally suited for long-range polarization transfer.

151

**Finite pulse effect
at 10 kHz MAS frequency***
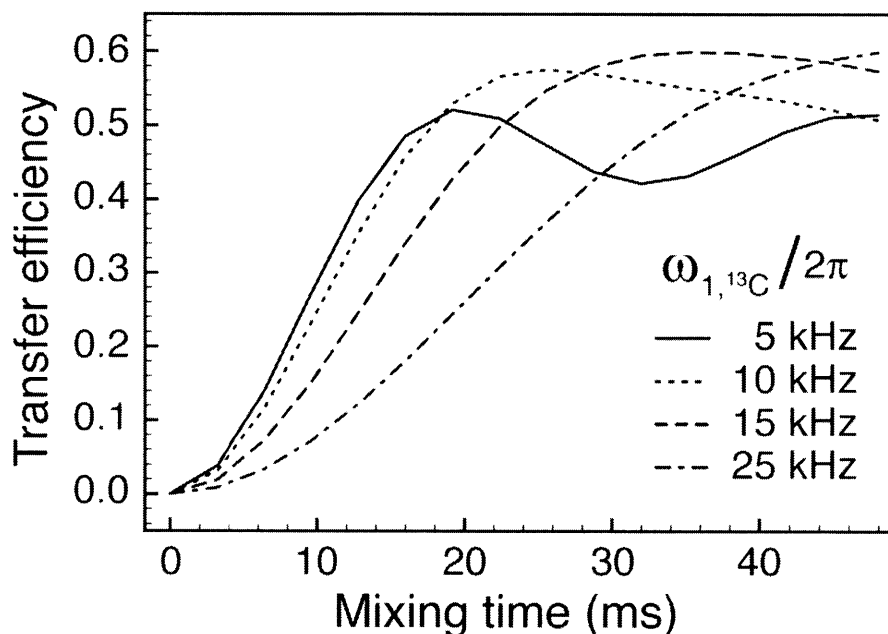
*Spin system: C^{α}--C^{α}, 3.8 Å distance

Figure S1. Numerical simulations of polarization transfer via RFDR-type longitudinal mixing at 10 kHz spinning frequency for various ^{13}C recoupling π-pulse nutation frequencies ($\omega_{1,^{13}C}/2\pi$). The simulated spin system was a ^{13}C-^{13}C pair with an internuclear distance of 3.8 Å (typical of sequential C^{α}- C^{α} nuclei) and a chemical shift separation of 2 kHz at 16.4 T magnetic field. Polarization buildup rates increase with the weakest irradiation levels, and the overall optimal transfer is achieved with ^{13}C rf fields between 5 and 10 kHz (in which the recoupling pulse occupies the full and half the rotor period, respectively), indicating that the finite-pulse mechanism of RFDR-type pulse sequences persists even at moderate spinning frequencies. The efficiency of the finite-pulse effect is the result of the low chemical shift separation between the recoupled resonances, as is found in aliphatic ^{13}C regions of protein NMR spectra.
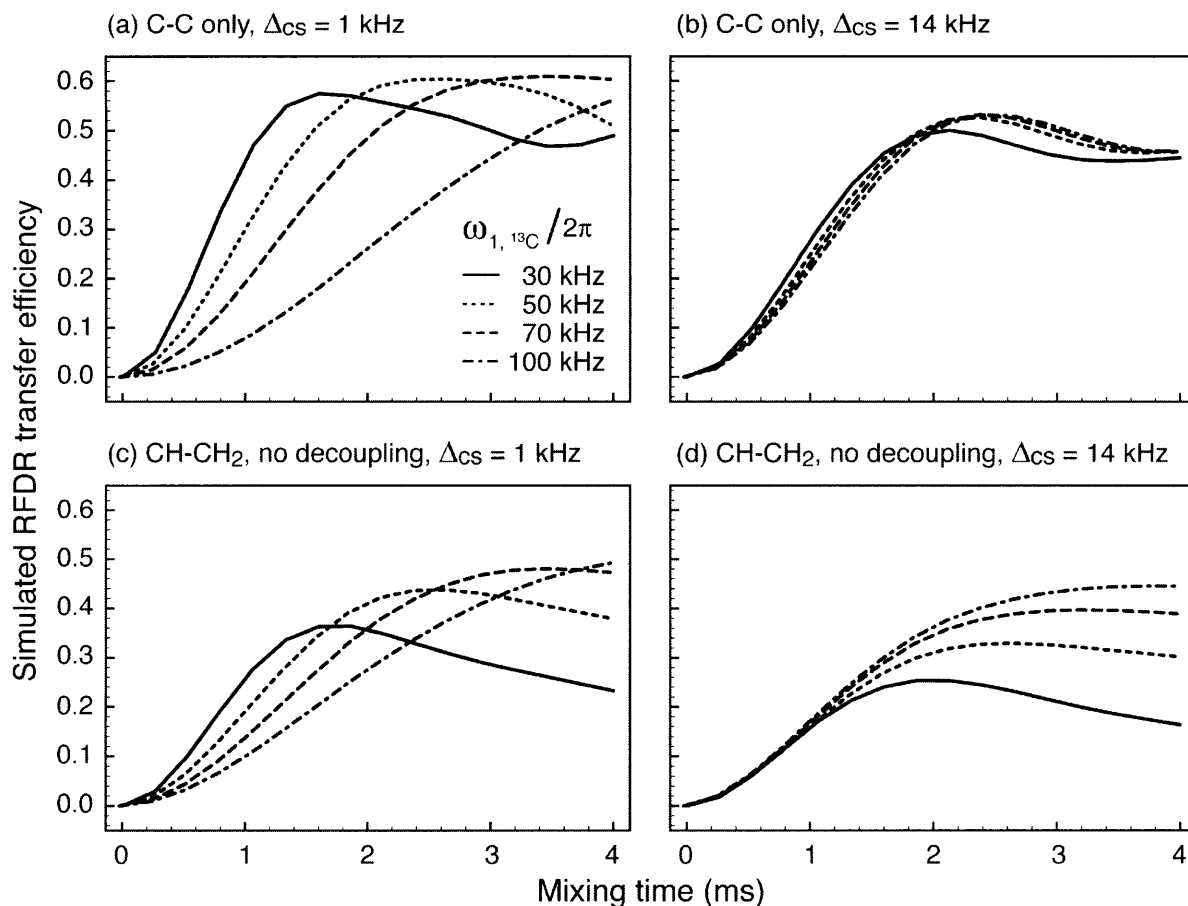
Figure S1B. Numerical simulations comparing of effects of chemical shift offset and ^{1}H decoupling for various ^{13}C rf fields in a RFDR-type sequence at 30 kHz MAS frequency. The ^{13}C-^{13}C pair has a one-bond coupling. In the absence of heteronuclear interactions (or in the limit of very strong ^{1}H decoupling irradiation), shown in (a) and (b), low ^{13}C rf powers are optimal in terms of scaling and efficiency for small chemical shift offsets (a) while the effect is not observed for large offsets (b). When protons are included in the simulations, and in the absence of decoupling, high ^{13}C rf powers are better in overall efficiency.
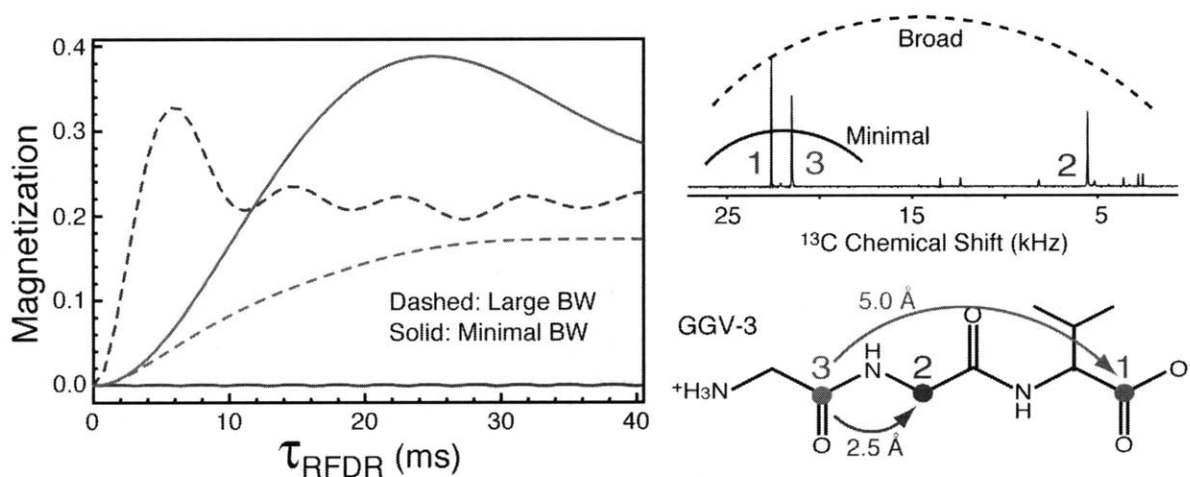
Figure S1C. The BASE RFDR scheme. Recoupling is limited to spins with resonances within the bandwidth of low-power π ^{13}C pulses, effectively reducing the spin system. The approach works best when interfering spins one wishes to remove from the recoupling dynamics (such as spin 2 above) are not directly bonded to the selected spins (spins 1 and 3 in this case). This is the case of many spin systems in proteins produced with alternating labeling. In this chapter, we used the BASE RFDR approach in the aliphatic region, thus removing spins from the carbonyl region, that is, the opposite selection as the one depicted above.
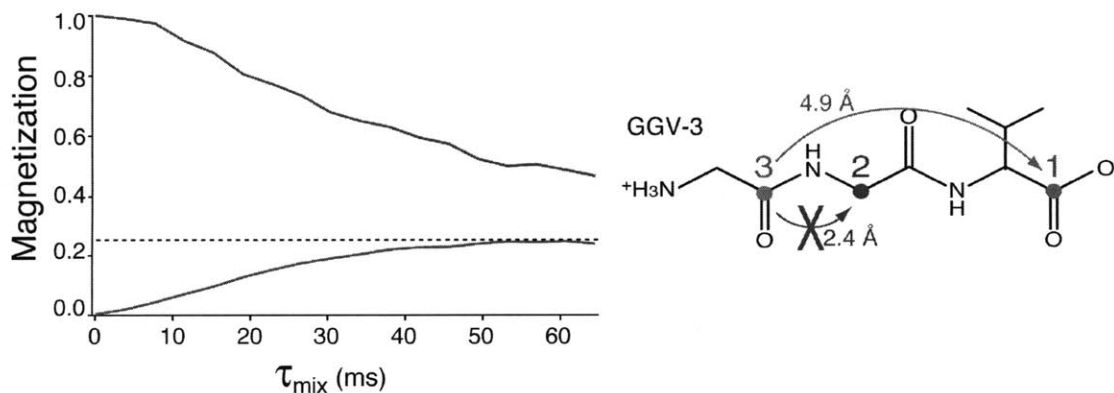
Figure S1D. The BASE RFDR scheme. Experimental demonstration of the case described above. Transfer efficiency reaches 25% over a 4.9 Å distance by diminishing the effect of spin 2. In the broadband case, polarization occurs in a relayed fashion because spin 2 is between spin 1 and spin 3, and thus it is difficult to separate the effect of dipolar truncation from relayed transfer, but the efficiency is lower (~7% for a 4.6 Å distance, as seen in chapter 4).

154

5A.2 The effect of heteronuclear interactions

One of the practical advantages of BASE RFDR is that heteronuclear decoupling can be achieved with moderately strong ^{1}H fields during mixing (70 to 100 kHz) with minimal losses from heteronuclear interference, which is generally of significant concern[8]. Figure S2 compares the influence of various dipolar couplings during a BASE RFDR polarization transfer experiment consisting of 12.5 kHz ^{13}C rf field and 80 kHz ^{1}H rf field at 12.5 kHz spinning frequency via numerical simulations with a CH-CH spin system. As can be observed, the ^{13}C-^{13}C coupling alone leads to highly efficient recoupling over the 3.8 Å internuclear distance. In an all-couplings simulation, the efficiency decreases slightly by an amount that is consistent with depolarization of ^{13}C magnetization by heteronuclear interference in a lone ^{1}H-^{13}C spin pair under similar rf and MAS settings (simulations not shown). Therefore, we can attribute the slight decrease in efficiency to incomplete suppression of first-order heteronuclear interactions. However, we can see that removing the ^{13}C-^{13}C coupling still results in a small extent of polarization transfer, likely due to the effect of higher-order interactions involving heteronuclear dipolar couplings and cross-terms between them, as occurs in third spin assisted recoupling (TSAR)[9]. While the relative TSAR effect in these simulations appears to be minor compared to the direct ^{13}C-^{13}C coupling we must note that (1) TSAR simulations are highly dependent on spin-system geometry, (2) the selected spin system geometry (which is of interest in the present study) is roughly 30% less efficient than the collinear C-H-C geometry that is optimal for TSAR transfer, and (3) in real spin systems the influence of geometry is attenuated by the presence of many other couplings making the TSAR mechanism more general. While the recoupling mechanism in BASE RFDR is clearly distinct from that of TSAR, the latter may be simultaneously active, leading to enhanced recoupling efficiency in a BASE RFDR experiment. Indeed, in BASE RFDR simulations including all couplings, the ^{1}H field can be chosen to maximize the TSAR effect (around 60 kHz in these settings), although at the expense of overall recoupling efficiency due to heteronuclear depolarization. On the other hand, with ^{1}H fields above 100 kHz we observe first-order direct ^{13}C-^{13}C recoupling, as expected. Numerical simulations in Figures S1 and S2 were carried out with the SPINEVOLUTION[110] package.

Dipolar interactions in BASE RFDR*
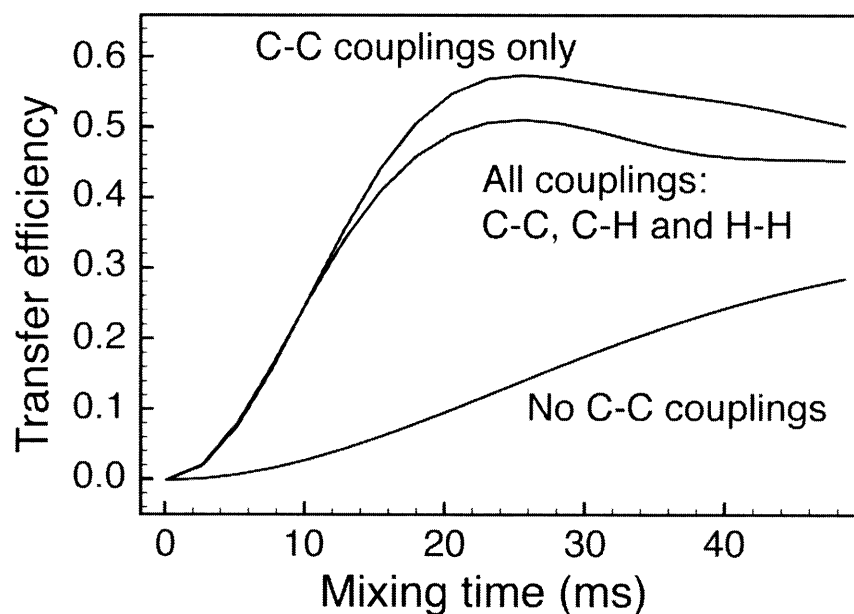
*Spin system: $C^\alpha H^\alpha$--$C^\alpha H^\alpha$. Settings:

$$\frac{\omega_{rotor}}{2\pi} = 12.5 \text{ kHz}, \quad \frac{\omega_{1,^{13}C}}{2\pi} = 12.5 \text{ kHz}, \quad \frac{\omega_{1,^{1}H}}{2\pi} = 80 \text{ kHz}$$

Figure S2. Numerical simulations of band selective RFDR polarization transfer mediated by different dipolar interactions in three cases: ^{13}C-^{13}C only, no ^{13}C-^{13}C couplings (only ^{1}H-^{13}C, ^{1}H-^{1}H, and their cross-terms), and all dipolar couplings. The simulated 4-spin system consisted of two ^{13}C-^{1}H pairs in a geometry corresponding to sequential $C^\alpha H^\alpha$- $C^\alpha H^\alpha$ nuclei (3.8 Å ^{13}C-^{13}C inter-nuclear distance) in a nearly ideal β-strand configuration, with a H-C-C-H dihedral angle of 179.8°, and C-H-C angles of 70° and 74°.

5A.3 Broadband RFDR of 2-PI3-SH3

PI3-SH3 fibrils prepared with protein produced with 2-^{13}C glycerol as the sole carbon source (2-PI3-SH3) were examined in this study. This type of labeling simplifies the spin dynamics as explained in the main text. Figure S3 presents a broadband RFDR[1, 2] spectrum of this sample, acquired with a 24.0 ms mixing period. RFDR mixing is well suited for long mixing periods thanks to its robustness agaisnt rf inhomogeneity and pulse and phase imperfections. The tolerance of the π pulse train with respect to rf imperfections is due in part to XY-16 phase cycling[11]. We found that appending XY-16 with $\overline{YX}$-16 to form a 32-step phase cycle further preserves ^{13}C polarization over long mixing periods, in particular for weak ^{13}C rf fields. The 32-step cycle is thus XYXY YXYX $\overline{XYXY}$ $\overline{YXYX}$ Y$\overline{XY}$X $\overline{X}$Y$\overline{X}$Y $\overline{Y}$X$\overline{Y}$X X$\overline{Y}$X$\overline{Y}$. Despite the robust performance of broadband RFDR, evidenced by many medium and long-range carbonyl-to-aliphatic contacts in Figure S3, few correlations are observed in the aliphatic-aliphatic region of the spectrum. From a technical point of view, this shortcoming is the result of the small chemical shift difference between the aliphatic resonances and also possible interference effects between ^{13}C and ^{1}H rf fields that often lead to incomplete heteronuclear decoupling[8]. Additionally, examining the spin systems generated by alternating labeling (see for example references [12] and [13]) one notices that carbonyl and Cα nuclei, although mostly not labeled in the same residue, are often the closest labeled spins (when sequential), separated by two bonds and leading to considerable dipolar truncation[14]. Other aliphatic nuclei also tend to be in close proximity to labeled carbonyl nuclei and thus carbonyl-aliphatic crosspeaks are the strongest in broadband RFDR spectra of 2-PI3-SH3.
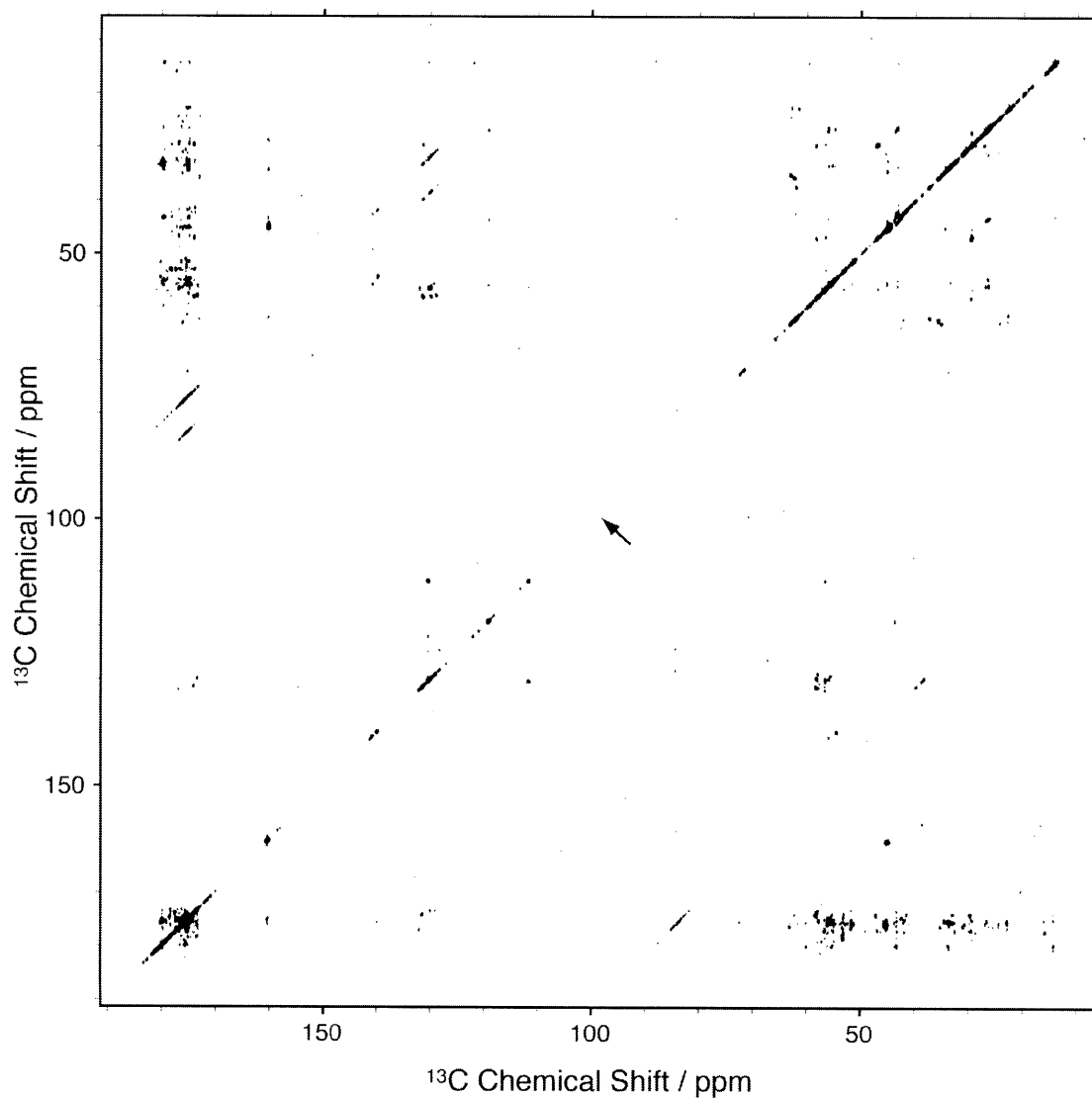
Figure S3. Broadband RFDR spectrum of 2-PI3-SH3 fibrils acquired at 16.0 kHz spinning frequency and 16.4 T (700 MHz ^{1}H Larmor frequency) using 40 kHz ^{13}C recoupling pulses and a 100 kHz ^{1}H decoupling field during a 24.0 ms mixing period. The arrow indicates the carrier position (98 ppm). 2048 and 512 points with dwell times of 12 us and 30 us were acquired in the direct and indirect dimensions, respectively (total t_2=24.576 ms and t_1=15.36 ms), with 32 scans averaged per t_1 increment, for a total acquisition time of 27 hrs. 100 kHz TPPM decoupling was applied during chemical shift evolution periods.

158

5.A.4 Comparison between BASE RFDR and DARR

Figure 2 in the main text demonstrates that BASE RFDR mixing produces many long-range aliphatic-aliphatic correlations in 2-PI3-SH3. The majority of these cross-peaks arise from sequential $C\alpha$-$C\alpha$, sequential $C\alpha$-$C\beta$, and other distant pairs of nuclei. The assignments shown in the inset of Figure 1(b) are part of ongoing work by the same authors (in preparation). The performance of LP RFDR is highlighted in Figure 4 of the main text with a comparison to DARR/RAD[15, 16] mixing, illustrating that long-range aliphatic-aliphatic recoupling via BASE RFDR is generally superior to spin diffusion methods. Figure S4 shows an overlay of the BASE RFDR and DARR spectra from Figure 4, in full view, excluding the empty right side of the spectra. It can be observed that carbonyl-aliphatic correlations are suppressed in BASE RFDR, while long-range aliphatic-aliphatic correlations are enhanced with respect to broadband RFDR. Correlation crosspeaks between distant aliphatic ^{13}C nuclei in BASE RFDR are more generally more sensitive that those obtained with DARR. Exceptions to this trend (some of which are noted in Figure 4 in the main text) can be expected to arise from remaining one-bond dipolar truncation in at least one of the interacting nuclei. Alternating labeling[12] with 2-^{13}C glycerol produces a small number of directly bonded pairs, such as Val $C\alpha$-$C\beta$, Leu $C\beta$-$C\gamma$, and Ile $C\alpha$-$C\beta$.[13] For such spin systems, dipolar truncation can be expected to decrease the efficiency of BASE RFDR transfer to distant spins, while mixing schemes that rely mainly on higher-order interactions, such as TSAR and DARR, are generally robust to such multi-spin effects although they are difficult to quantify precisely due to their complex magnetization transfer mechanisms.
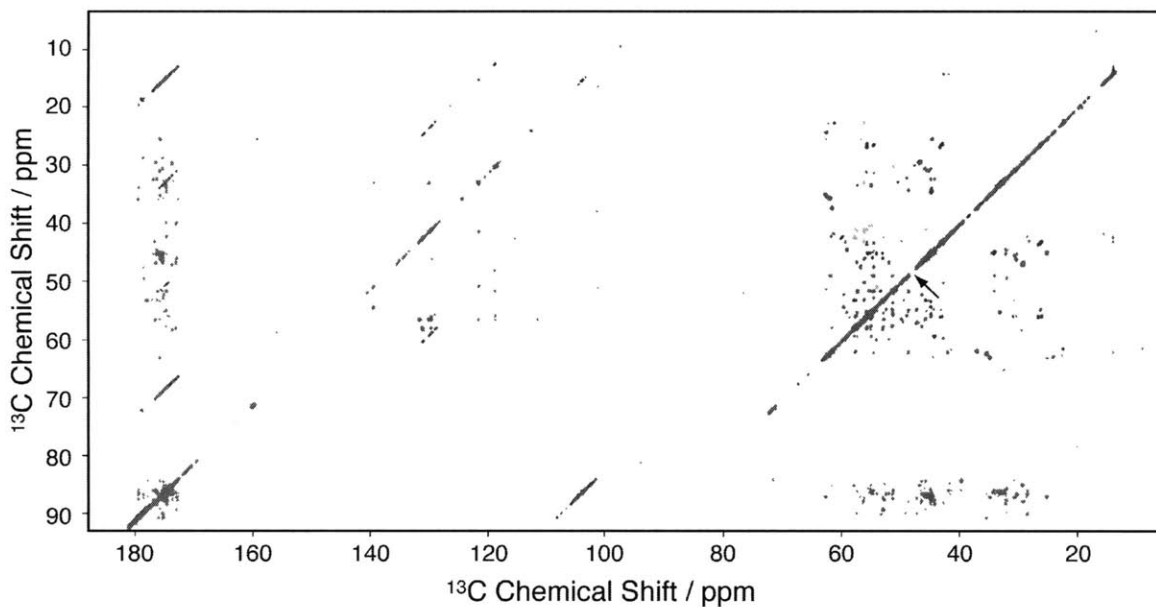
Figure S4. Overlay of BASE RFDR (blue) and DARR (red and orange) spectra of PI3-SH3 fibrils. The spectra are full views of those in Figure 2 in the main text and were acquired at 12.5 kHz spinning frequency and 16.4 T (700 MHz ^{1}H Larmor frequency). The BASE RFDR mixing period consisted of 12.5 kHz ^{13}C rotor-synchronized π pulses with a 32-step phase cycle for total duration of 17.92 ms, while the DARR mixing period consisted of ~12 kHz ^{1}H irradiation with a 500 ms duration. The arrow indicates the carrier position (48 ppm) for both experiments. DARR crosspeaks folded into the aliphatic region are shown in orange contours for clarity. For each 2D correlation experiment 2048 and 256 points with dwell times of 12 us and 60 us were acquired in the direct and indirect dimensions, respectively (total t_2=24.576 ms and t_1=15.36 ms), with 16 scans averaged per t_1 increment. The total experimental time was 6.8 hrs per 2D. 100 kHz TPPM decoupling was applied during chemical shift evolution periods. The BASE RFDR experiment was performed four times and averaged together to yield the spectrum shown in Figure 1(b) of the main text.

5A.5 Windowless BASE RFDR and off-resonance effects

As mentioned in the main text of this chapter, BASE RFDR can be implemented with r.f. pulses with nutation frequencies between approximately the spinning frequency to exactly half the spinning frequency. In the latter case, each π pulse occupies the entire rotor period and the pulse sequence is thus windowless. In windowless BASE RFDR, the 2-to-1 ratio of the MAS frequency ($\omega_r/2\pi$) to ^{13}C r.f. field ($\omega_1/2\pi$) resembles that found in the HORROR experiment, with the important difference that while the HORROR condition with constant-phase irradiation results in double-quantum polarization transfer during a spin-locking irradiation, application of XY-type phase sequences during BASE RFDR enables the transfer of longitudinal magnetization via zero-quantum coherence, as illustrated in Figure S5.
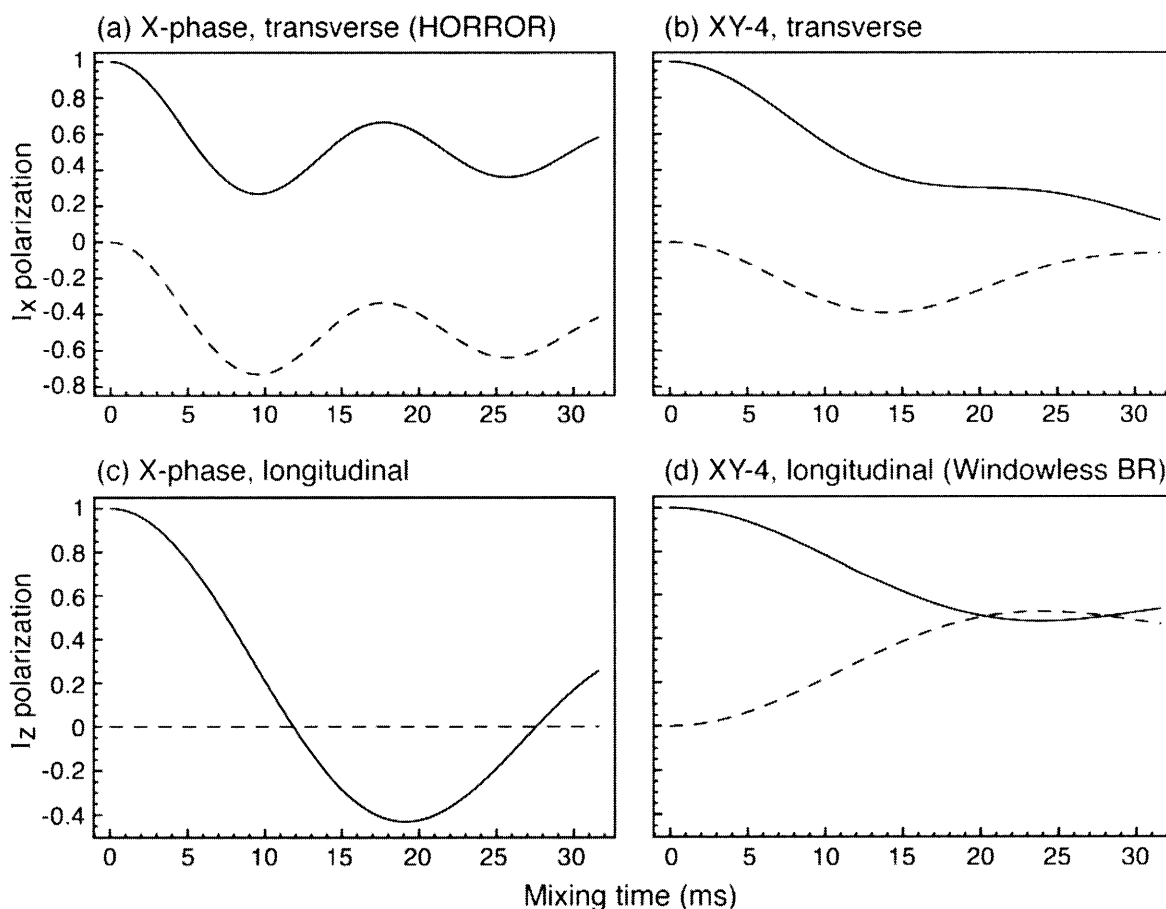
Figure S5. Simulations of transverse (a and b) and longitudinal (c and d) polarization transfer under a HORROR-type condition ($\omega_r/2\pi$ = 10 kHz, $\omega_1/2\pi$ = 5 kHz) using single-phase irradiation (a and c) and an XY phase sequence (b and d). Initial polarization (solid line) from one spin is transferred to a second spin (dashed line), with a 4 Å internuclear distance.

Using the windowless limit of BASE RFDR provides the optimal scaling factor for closely spaced resonances, as shown in the simulations of Figure S1. At the same time, employing such reduced bandwidths may facilitate restricting polarization transfer to a small range of chemical shift offsets. For example, Figure S6 shows a 1D ^{13}C CP spectrum of [U-^{13}C, ^{15}N] N-acetyl-VL (acquired at 500 MHz ^{1}H Larmor frequency and with a 10.6 kHz MAS frequency) that spans ~20 kHz, while the carbonyl resonances span less than 2 kHz. At this spinning frequency, windowless BASE RFDR is applied with $\omega_1/2\pi$ = 5.3 kHz, which is sufficient to cover the carbonyl region. Indeed, after 20 ms of 5.3 kHz windowless BASE RFDR applied in a 1D fashion, approximately 50 % of the carbonyl signal of N-a-VL remains when the phase-shifted XY-32 scheme is used (see chapter 3, section 2.3). The amount of depolarization is significantly low considering the weak r.f. field applied and the large chemical shift anisotropy of these nuclei, both of which factors are compensated for by the phase-shifted XY-32 scheme described in chapter 3. Therefore we recorded windowless BASE RFDR spectra with the objective of recoupling dipolar interactions between the carbonyl nuclei exclusively, placing the carrier frequency near in this region, which is between ~14 and ~18 kHz downfield of the $_{13}$Cα and side-chain resonances, respectively

[U-^{13}C,^{15}N] N-Ac-VL, 10.6 kHz MAS

10.6 kHz ^{13}C rf

5.3 kHz ^{13}C rf

25   15   5

^{13}C Chemical Shift (kHz)
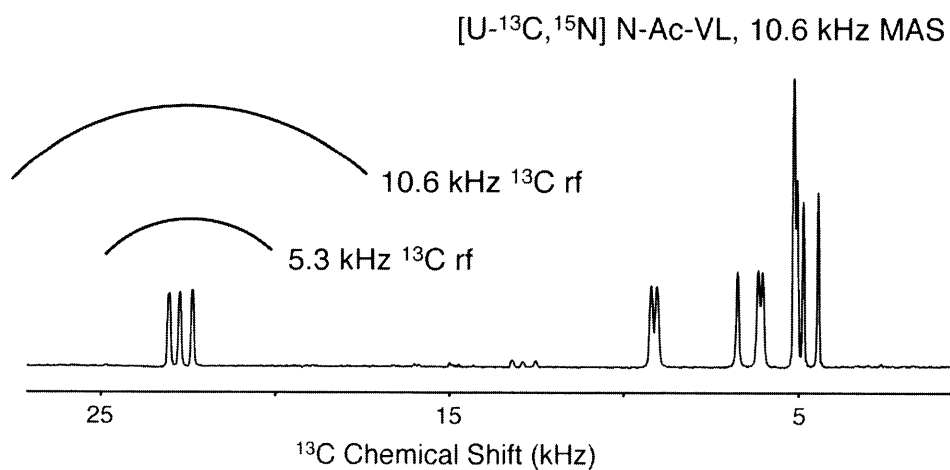
Figure S6. ^{13}C spectrum of [U-^{13}C, ^{15}N] N-a-VL recorded at 500 MHz ^{1}H Larmor frequency, showing chemical shift distribution in kHz, and the approximate band width of BASE RFDR pulses, down to the windowless limit, in which the recoupling irradiation frequency is half the spinning frequency.
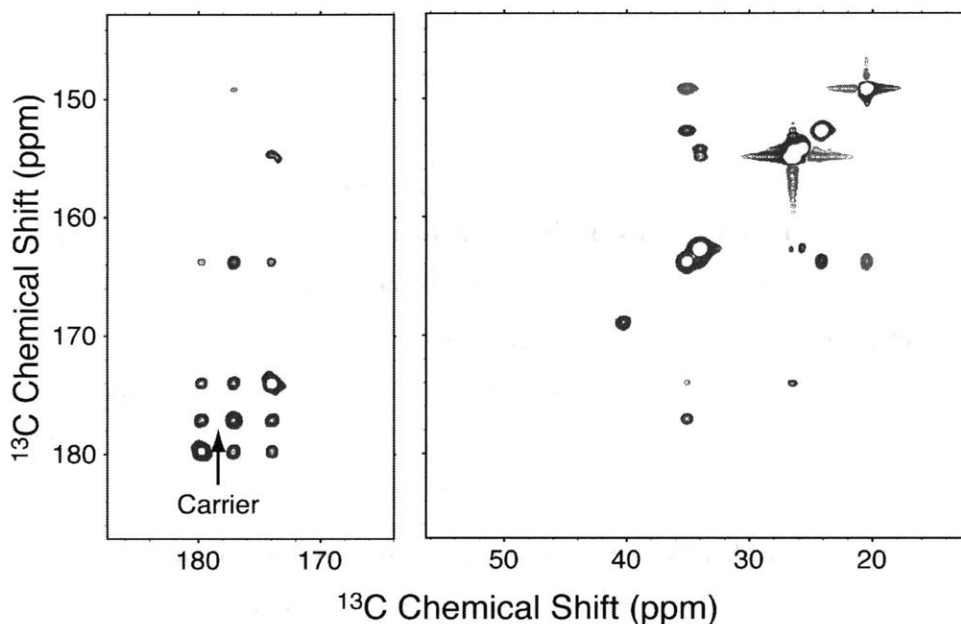
Figure S7. Windowless BASE RFDR spectrum of [U-^{13}C, ^{15}N] N-a-VL recorded with a 24 ms mixing time and the carrier frequency near the carbonyl region. Aliphatic resonances are aliased in the indirect dimension. Strong cross-peaks are observed between the three ^{13}C nuclei in the targeted carbonyl region, despite strong dipolar truncation in this uniformly ^{13}C labeled system. In addition, a number of positive (blue) and negative (red) cross-peaks, corresponding to zero-quantum and double-quantum polarization transfer, respectively, are observed in the aliphatic-aliphatic and aliphatic-carbonyl spectral regions.

A representative ^{13}C-^{13}C windowless BASE RFDR correlation spectrum of the carbonyl region of N-a-VL is illustrated in Figure S7, which was recorded at 500 MHz ^{1}H Larmor frequency, and with $\omega_r/2\pi = 10.6$ kHz, $\omega_1/2\pi = 5.3$ kHz, and a mixing period of 24 ms. Strong cross-peaks are observed between the Leu (downfield), Val (upfield), and acetyl (center) carbonyl nuclei. The internuclear distances between these nuclei, from the crystal structure of N-a-VL, are 3.45 Å between Ac and Val, 3.01 Å between Val and Leu, and 6.39 Å between Ac and Leu. All three interactions give rise to cross-peaks, with the Ac-Leu transfer likely being mediated by relayed polarization transfer through the intervening Val carbonyl nucleus. Even though this peptide is uniformly labeled and thus subject to strong dipolar truncation, windowless BASE RFDR results in efficient polarization transfer between these distant nuclei due to the partial exclusion of aliphatic nuclei from the dipolar recoupling dynamics between carbonyl spins. The exclusion is only partial because of multiple off-resonance effects, that is, the interaction of the applied r.f. field with the chemical shift offset of resonances far from the

163

carrier frequency, result in weak recoupling conditions between carbonyl and aliphatic spins. These conditions are also evident in the aliphatic region of Figure S7, which shows multiple cross-peaks with either positive (blue) or negative (red) signs, corresponding to polarization transfer via zero-quantum or double-quantum coherences, respectively. The possible sources of off-resonance recoupling conditions include rotational resonance in the tilted frame and HORROR conditions, both of which can be established with weak r.f. fields and moderate resonance offsets, and are strongest between directly bonded ^{13}C nuclei.

Since the suppression of polarization transfer from spins outside the recoupling bandwidth is only partial, due to off-resonance effects and strong one-bond dipolar couplings, the efficiency of BASE RFDR in this uniformly ^{13}C labeled system is only approximately 11% for the 3.45 Å internuclear distance. We can compare this result to the case of GGV-3, shown in Figure S1D, in which 25% of the polarization is transferred from one carbonyl spin to another by excluding the intervening $^{13}C\alpha$ spin from the BASE RFDR bandwidth. The $^{13}C\alpha$ spin presents a two-bond coupling to one of the carbonyl spins, instead of the one-bond couplings found in the N-a-VL system, and thus the off-resonance effects are reduced in the GGV-3 system, resulting in better BASE RFDR transfer efficiency even though the internuclear distance is considerably longer (4.9 Å) in the GGV-3 carbonyl pair than between sequential carbonyl nuclei in N-a-VL.

Therefore, we may anticipate that off-resonance effects in BASE RFDR experiments applied to spin systems in proteins produced with alternating ^{13}C-^{12}C labeling should be reduced significantly compared to the case of uniformly ^{13}C samples, which constitutes a further advantage of these dilute spin systems, in addition to presenting weaker dipolar truncation. We recorded windowless BASE RFDR experiments on PI3-SH3 amyloid fibrils formed by protein produced with [2-^{13}C] glycerol as the source of carbon (2-PI3-SH3). The experiments were aimed at recoupling interactions between aliphatic nuclei and excluding carbonyl spins from the recoupling dynamics. Similar to the results shown in the main text of this chapter, multiple long-range cross-peaks were observed in the aliphatic region, but with increased intensities. The spectrum shown in Figure S8 was recorded at 700 MHz ^1H Larmor frequency, with $\omega_r/2\pi = 12$ kHz, $\omega_1/2\pi = 6.0$ kHz, and a mixing period of 24 ms, with the carrier frequency placed at 52 ppm. At this mixing time, a number of new long-range correlations are observed, compared to the 18 ms BASE RFDR spectrum in the main text of this chapter. Despite the fact that low-power ^{13}C irradiation is used, the windowless BASE RFDR bandwidth covers the majority of the

164

^{13}Cα resonances, as can be seen in Figure S8, and leads to highly efficient polarization transfer. Indeed, the low-power ^{13}C pulses, which use the phase-shifted XY-32 sequence, facilitate the application of moderate ^{1}H decoupling fields. An important application of efficient recoupling between ^{13}Cα nuclei is described in chapter 8 of this thesis.
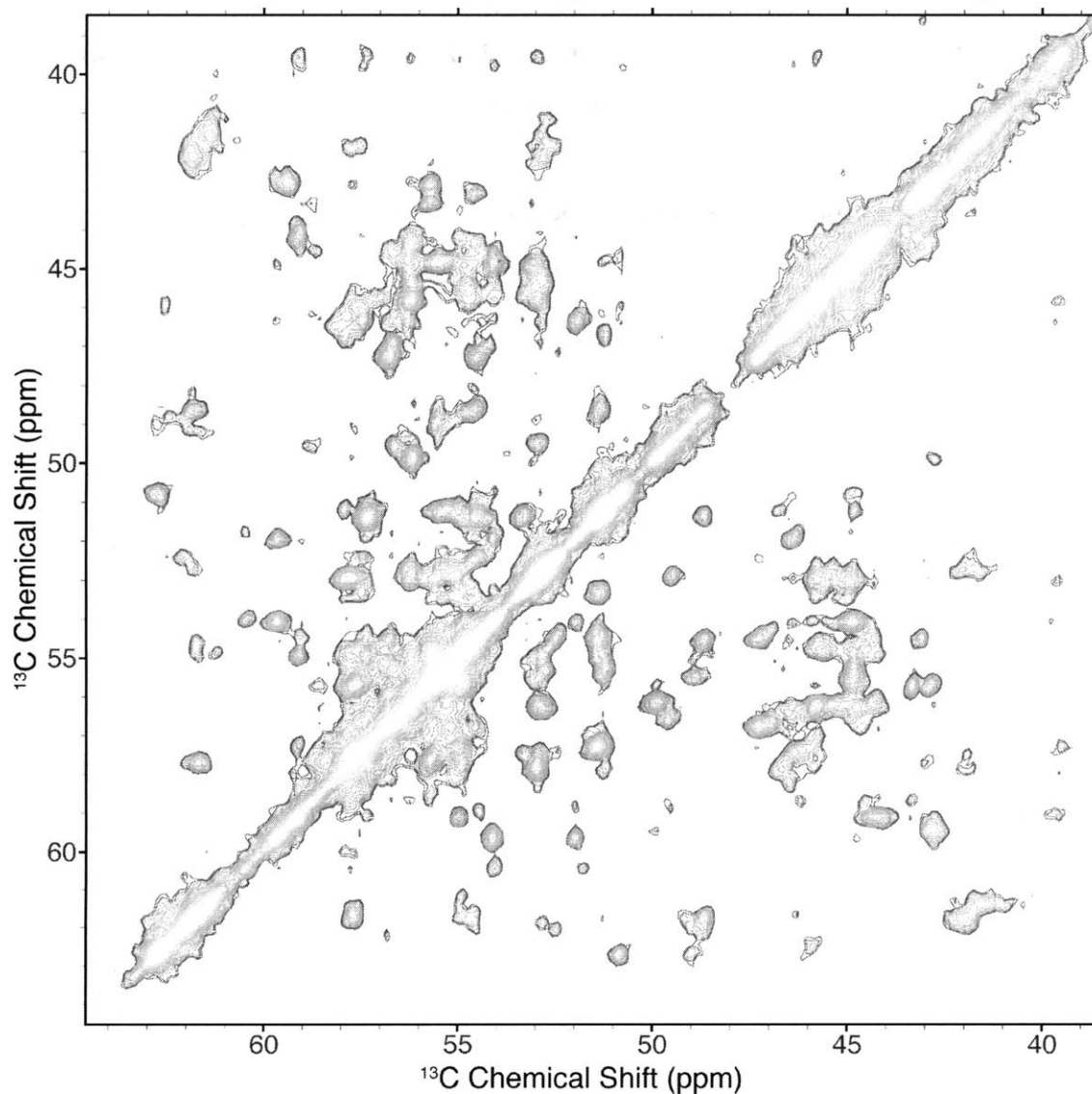
Figure S8. Windowless BASE RFDR spectrum of 2-PI3-SH3 recorded at 700 MHz ^{1}H Larmor frequency, with $\omega_r/2\pi$ = 12 kHz, $\omega_1/2\pi$ = 6.0 kHz, and a mixing period of 24 ms. The carrier frequency was placed in the middle of the ^{13}Cα region.

References for Appendix to Chapter 5

1. Bennett, A. E., Griffin, R. G., Ok, J. H. & Vega, S. (1992). Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys.* **96**, 8624-8627.

2. Bennett, A. E., Rienstra, C. M., Griffiths, J. M., Zhen, W., Lansbury, J. & Griffin, R. G. (1998). Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys.* **108**, 9463-9479.

3. Boender, G. J., Raap, J., Prytulla, S., Oschkinat, H. & de Groot, H. J. M. (1995). MAS NMR structure refinement of uniformly 13C enriched chlorophyll a/water aggregates with 2D dipolar correlation spectroscopy. *Chem. Phys. Lett.* **237**, 502-508.

4. Balaban, T. S., Holzwarth, A. R., Schaffner, K., Boender, G. J. & de Groot, H. J. M. (1995). CP-MAS 13C-NMR Dipolar Correlation Spectroscopy of 13C-Enriched Chlorosomes and Isolated Bacteriochlorophyll c Aggregates of Chlorobium tepidum: The Self-Organization of Pigments Is the Main Structural Feature of Chlorosomes. *Biochemistry* **34**, 15259-15266.

5. Ok, J. H., Spencer, R. G. S., Bennett, A. E. & Griffin, R. G. (1992). Homonuclear correlation spectroscopy in rotating solids. *Chem. Phys. Lett.* **197**, 389-395.

6. Boender, G. J., Vega, S. & deGroot, H. J. M. (2000). Quantized field description of rotor frequency-driven dipolar recoupling. *J. Chem. Phys.* **112**, 1096-1106.

7. Ishii, Y. (2001). 13C-13C dipolar recoupling under very fast magic angle spinning in solid-state nuclear magnetic resonance: Applications to distance measurements, spectral assignments, and high-throughput secondary-structure determination. *J. Chem. Phys.* **114**, 8473-8483.

8. Bayro, M. J., Ramachandran, R., Caporini, M. A., Eddy, M. T. & Griffin, R. G. (2008). Radio frequency-driven recoupling at high magic-angle spinning frequencies: Homonuclear recoupling sans heteronuclear decoupling. *J. Chem. Phys.* **128**, 052321.

9. De Paëpe, G., Lewandowski, J. R., Loquet, A., Böckmann, A. & Griffin, R. G. (2008). Proton Assisted Homonuclear Recoupling and Protein Structure Determination. *J. Chem. Phys.* **129**, 245101.

10. Veshtort, M. & Griffin, R. G. (2006). SPINEVOLUTION: a powerful tool for the simulation of solid and liquid state NMR experiments. *J. Magn. Reson.* **178**, 248-82.

11. Gullion, T., Baker, D. B. & Conradi, M. S. (1990). New, compensated Carr-Purcell sequences. *J. Magn. Reson.* **89**, 479-484.

12. LeMaster, D. & Kushlan, D. (1996). Dynamical Mapping of E. coli Thioredoxin via 13C NMR Relaxation Analysis. *J. Am. Chem. Soc.* **118**, 9255-9264.

13. Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K. & Oschkinat, H. (2002). Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **420**, 98-102.

14. Bayro, M. J., Huber, M., Ramachandran, R., Davenport, T. C., Meier, B. H., Ernst, M. & Griffin, R. G. (2009). Dipolar truncation in magic-angle spinning NMR recoupling experiments. *J. Chem. Phys.* **130**, 114506.

15. Takegoshi, K., Nakamura, S. & Terao, T. (2001). 13C-1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.*

16. Morcombe, C. R., Gaponenko, V., Byrd, R. A. & Zilm, K. W. (2004). Diluting abundant spins by isotope edited radio frequency field assisted diffusion. *J. Am. Chem. Soc.* **126**, 7196-7.

Chapter 6. A Protocol for Rapid Backbone Resonance Assignment of Proteins in the Solid State

Adapted from a manuscript in preparation by Marvin J. Bayro, Matthew T. Eddy, Galia T. Debelouchina, and Robert G. Griffin

Summary

We describe a data-collection and analysis protocol for the rapid and unambiguous assignment of backbone resonances of proteins via solid-state magic-angle spinning NMR spectroscopy. The basic experimental data set consists of four two-dimensional experiments that are optimal for establishing one-, two-, and three-bond sequential correlations in protein samples prepared with [2-^{13}C] glycerol labeling. The experiments exploit both the favorable 13C spin dynamics found in these samples and a series of characteristic ^{13}C-^{13}C and ^{15}N-^{13}C patterns exhibited between neighboring residues in this alternating labeling scheme. The protocol is shown to yield almost complete backbone resonance assignment in a small 56-residue protein, and can be expected to be applicable to proteins with up to 100-150 residues.

6.1 Introduction

A variety of complex biological systems exist for which solid-state NMR is ideally suited as a powerful method to extract high-resolution structural and dynamical information, including membrane proteins, protein fibrils, and protein assemblies. The initial step in NMR structure determination of proteins is the assignment of the individual resonances in the amino acid sequence, which permits the subsequent identification of long-range internuclear constraints that define the overall structure (*1, 2*). However, chemical shifts themselves are highly sensitive to the local chemical environment and provide important insight regarding site-specific secondary structure conformation. Moreover, methodology to derive protein structures based on chemical shifts alone has been developed recently, and new approaches continue to emerge for this purpose (*3-7*). Elucidating site-specific chemical shifts is, therefore, a critical process in structural studies of proteins, for which an extensive variety of resonance assignment protocols have been designed in solution NMR spectroscopy.

In magic-angle spinning (MAS) NMR spectroscopy, sufficient resolution is available typically only in ^{13}C and ^{15}N spectra (for fully protonated protein samples) and thus the resonance assignment process is limited to double-resonance experiments and based on connectivities between directly bonded nuclei. The general resonance assignment scheme in MAS NMR spectroscopy consists of intra-residue CC and NCACX correlations and inter-residue, sequential NCOCX correlations, implemented in two (2D) or three (3D) spectral dimensions, applied on U-^{13}C,^{15}N labeled proteins (*8, 9*). In these experiments, ^{13}C-^{13}C homonuclear correlations within each residue can be acquired with a variety of dipolar recoupling pulse sequences, while ^{15}N-^{13}C heteronuclear correlations are typically obtained with specific cross-polarization between directly bonded ^{15}N-^{13}Cα or ^{15}N-^{13}C' pairs (*10*) and are followed by a ^{13}C-^{13}C mixing period to establish NCACX or NCOCX correlations, respectively. The experimental implementation of this protocol is robust, as it relies on the strong dipolar interactions between directly bonded nuclei, yielding adequate experimental efficiencies for multiple-step polarization transfer experiments. On the other hand, data analysis is severely complicated even by modest levels of resonance overlap, becoming tedious and ambiguous. The resonance assignment process is, therefore, greatly facilitated by performing the experiments in 3D or 4D fashion, which is usually limited by sample sensitivity. Despite these limitations,

resonance assignments have been obtained for various moderately sized proteins in the solid state (11-19). Partial resonance assignments have also been obtained for large protein systems (over 200 residues), although often by means of employing multiple samples with various labeling schemes such as specific or reverse amino acid labeling (20, 21). Nevertheless, resonance assignment remains a significant bottleneck in protein MAS NMR studies even for moderately sized proteins, in particular when samples present intermediate line widths or low sensitivity in general.

An effective method to simplify spectra and reduce ^{13}C line widths is to employ the alternating ^{13}C-^{12}C labeling scheme introduced by LeMaster and Kushlan (22), which incorporates ^{13}C nuclei at approximately every other carbon site, abolishing the majority of one-bond ^{13}C-^{13}C dipolar and J-couplings. Originally applied to solid-state NMR studies as a means of decongesting long-range spin diffusion spectra used for structure determination (23, 24), alternating ^{13}C-^{12}C labeling has been proven also to be of great help in the sequential resonance assignment process. Recently, Oschkinat and co-workers described the benefits of alternating labeling in sequential assignments in terms of narrowing of resonances in medium-range ^{13}C-^{13}C spin diffusion spectra of a protein presenting intermediate line widths (25). Subsequently, an exhaustive approach was implemented with several 3D correlation experiments combining sequential information from three different samples labeled with [U-^{13}C] glucose, [2-^{13}C] glycerol, and [1,3-^{13}C] glycerol, respectively (26). This strategy exploits the complementary labeling patterns of samples in either type of glycerol, an additional benefit of alternating labeling. From a spin dynamics perspective, we have demonstrated that the simplified local spin systems encountered in alternating ^{13}C-^{12}C labeling are amenable to long-range dipolar recoupling via efficient homonuclear recoupling pulse sequences, despite lingering dipolar truncation effects (27). Moreover, we have recently shown that band-selective recoupling can be employed to obtain highly sensitive medium- and long-range ^{13}C-^{13}C correlations between aliphatic nuclei in protein samples produced with 2-^{13}C glycerol (28). This labeling scheme can also enhance the efficiency of medium-range heteronuclear correlation spectra, yielding sequential connectivities that facilitate the assignment process, as recently shown for an amyloid fibril of an SH3 domain, described in chapter 7 of this thesis.

In this chapter we propose a comprehensive set of 2D dipolar correlation experiments designed to establish inter-residue connectivities that can be analyzed in a highly efficient

manner to obtain sequential resonance assignments of protein samples produced with [2-^{13}C] glycerol. Proteins labeled with [2-^{13}C] glycerol present enhanced resolution and most Cα sites are ^{13}C labeled, at least partially, enabling us to perform experiments in two-dimensional fashion with optimal sensitivity and to establish sequential correlations between backbone nuclei. The experiments proposed rely on medium-range homonuclear and heteronuclear dipolar interactions and are tailored to exploit the characteristic spin systems encountered in sequential pairs of amino acids in alternatively ^{13}C-^{12}C labeled proteins, beyond the spectral simplification that spin dilution affords. We show that a methodical examination of this series of experiments results in the elucidation of sequential assignments with a high degree of redundancy, minimizing ambiguities and thus speeding up the assignment process. The experiments we suggest, together with the analysis we describe, constitute an efficient protocol for the rapid resonance assignment of small to moderate-size proteins in the solid state.

6.2 Methods

6.2.1 Protein samples

We tested our resonance assignment protocol with a microcrystalline sample of the beta-1 domain of Immunoglobulin protein G (GB1, 56 residues), for which complete resonance assignments have been previously published and have been used as part of, and corroborated through, structure determination, and a sample of the SH3 domain of phosphatidylinositol-3-kinase (PI3-SH3, 86 residues) in amyloid fibril form, for which almost complete resonance assignments are reported in chapter 7, but have yet to be fully corroborated with structure determination. Two GB1 samples were prepared as previously described (*15*), one of them using ^{13}C glucose (U-GB1) and the other using 2-^{13}C glycerol and ^{13}C bicarbonate as the sole sources of carbon (2-GB1). Two PI3-SH3 samples were prepared as previously described (*29*), one of them using ^{13}C glucose (U-PI3-SH3) and the other using 2-^{13}C glycerol and ^{13}C bicarbonate as the sole sources of carbon (2-PI3-SH3). Incubation in acidic conditions without seeding during a period of two weeks resulted in the generation of PI3-SH3 amyloid-like fibrils that were subsequently spun down in a 60/40 d5-glycerol/water solvent (w/w), which helps preserve hydration, maintain the sample integrity, and improve the spectral resolution. Approximately 18 mg of GB1 and 8 mg of PI3-SH3 were contained in each sample.

6.2.2 NMR spectroscopy

All experiments were performed in a spectrometer operating at 700 MHz ^{1}H Larmor frequency (courtesy of Dr. David J. Ruben, Francis Bitter Magnet Laboratory, Cambridge MA), corresponding to a 16.4 T field, using a triple resonance Varian/Chemagnetics (Palo Alto, CA) probe equipped with a 3.2 mm stator. Experiments were acquired with a 12.5 kHz spinning frequency, unless otherwise noted, and sample temperatures were maintained around 5 °C with a stream of cooled N_2 gas.

Correlations experiments used ^{1}H-^{13}C cross-polarization preparation and high-power TPPM decoupling (*30*) was applied on the ^{1}H channel during the chemical shift evolution and detection periods. Transferred echo double resonance (TEDOR) experiments (*31, 32*), with mixing periods of either 1.6 or 5.76 ms, consisted of 128 real and 128 imaginary t1 points acquired in 100 μs increments, with total acquisition times of 7 hours for 2-GB1 and 28 hours for 2-PI3-SH3. Broadband radio frequency-driven recoupling (RFDR) experiments (*33-35*) with a

7.68 ms mixing period consisted of 640 real and 640 imaginary t1 points acquired in 28 μs increments, with total acquisition times of 8.8 hours for 2-GB1 and 35 hours for 2-PI3-SH3. Band-selective RFDR experiments (28), with a 17.92 ms mixing period, consisted of 320 real and 320 imaginary t1 points acquired in 56 μs increments, with total acquisition times of 8.8 hours for 2-GB1 and 35 hours for 2-PI3-SH3. Additional experiments included homonuclear mixing via cosine-modulated adiabatic recoupling (CMAR) (36) and RFDR on uniformly labeled samples, used in an initial assessment of sample quality and to determine ^{13}C spin systems. Data processing was performed with the NMRPipe (37) suite and analyzed with the Sparky program.

6.3 Resonance Assignment Protocol

Alternating ^{13}C labeling of proteins with [2-^{13}C] glycerol results in a decreased number of directly bonded ^{13}C-^{13}C pairs of nuclei, which has several important consequences. First, J-couplings are removed for most carbon sites, which narrows the resonance line widths and further decongests the spectra of these dilute ^{13}C samples (*22-24, 38*). Of more profound impact for solid-state NMR experiments, dipolar truncation effects (*27*) are considerably attenuated for most amino acid spin systems in alternatively labeled proteins, facilitating the application of efficient homonuclear recoupling pulse sequences to establish medium- and long-range sequential correlations, instead of being limited to proton-driven spin diffusion (PDSD) (*39-41*) and related techniques (*42, 43*) for ^{13}C-^{13}C polarization transfer. Finally, as we describe below, alternating labeling produces amino acid residues with particular spin systems that yield fingerprint patterns in sequential homonuclear and heteronuclear correlation spectra. These sequential patterns can be exploited to simplify the analysis of correlation spectra tailored for alternatively labeled samples.

The spin systems generated by 2-^{13}C glycerol labeling can be classified into three classes depending on how the backbone nuclei are labeled. (1) 100% ^{13}Cα and ^{12}C': These amino acids (including G, A, S, C, H, W, F, Y, and V) are particularly suited for sequential Cα-C(α/β/γ) correlations. Val is special in that the Cβ site is labeled concurrently with the Cα causing a strong dipolar truncation effect. (2) Fractionally ^{13}Cα/^{12}Cα and ^{12}C'/^{13}C': These are the majority of amino acids (including I, K, E, Q, P, R, D, N, M, and T) and contribute to both aliphatic-aliphatic correlations and carbonyl-aliphatic correlations. Interestingly, even though both Cα and C' sites are partially ^{13}C labeled, the ^{13}C isotope is rarely incorporated simultaneously in both positions of the same residue in a given protein molecule. Thus dipolar truncation is still minimized in these cases while providing the versatility of observing either nucleus in sequential correlations. (3) 100% ^{12}Cα and ^{13}C': Only Leu fits in this category, making it a special case in the sequential analysis procedure. In some cases, amino acids with very small ^{13}Cα fractions (E, Q, P, and R) may be regarded in this category. It must be noted that a small fraction of Gly residues present ^{13}C incorporated in the C' position, but they rarely occur with both ^{13}Cα and ^{13}C' labels concurrently. Indeed, these ancillary Gly C' resonances, thanks to their typically upfield chemical shifts, may be of additional help in the spectral analysis process.
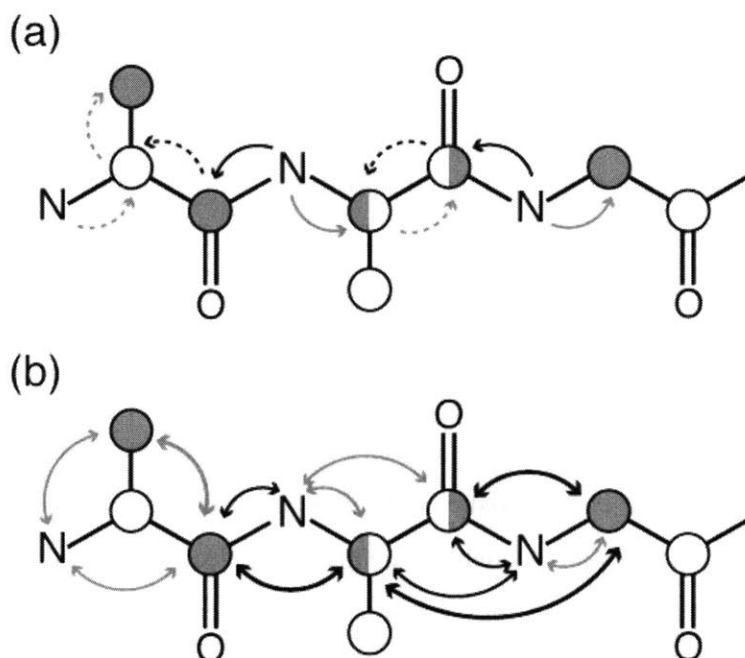
Figure 1. Example of a labeling patterns in a three-residue segment found in alternating ^{13}C-^{12}C labeling with [2-^{13}C] glycerol and possible polarization transfer pathways. (a) NCXCY approach (with short mixing times) and (b) our proposed method. Each residue corresponds to one of the three classes of ^{13}C spin systems described in the text. Fully filled gray circles denote ^{13}C-labeled carbon sites, while empty circles represent sites close to natural abundance. Half-filled circles stand for carbon sites with partial ^{13}C labeling, that is, with ^{13}C labels in a fraction of the total molecules. In alternating labeling, directly bonded carbons may be labeled fractionally but they often are not labeled simultaneously in the same molecule, therefore the adjacent partially labeled sites in the middle residue are depicted with mismatched filled-empty and empty-filled patterns. Gray and black arrows represent intra-residue and inter-residue correlations, respectively. (a) NCACX (gray arrows) and NCOCX (black arrows) correlations based on relayed ^{13}C-^{13}C transfer with short mixing times do not work well due to the alternating labeling pattern. (b) Direct ^{13}C-^{13}C and ^{15}N-^{13}C correlation of short and medium-range intra-residue (gray arrows) and sequential (black arrows) connectivities, established through a set of four experiments described in the text.

An example of a sequence of amino acids including all three kinds of spin systems is presented in Figure 1. Simple inspection of the labeling pattern reveals that sequential correlations of the NCACX/NCOCX type (one-bond ^{15}N-^{13}C cross-polarization (DCP) followed by one- or two-bond ^{13}C-^{13}C mixing) are of limited effectiveness in establishing sequential and intra-residue connectivities in such sequence of amino acids. In the initial ^{15}N-^{13}C DCP step the one-bond heteronuclear transfer is precluded in the complete absence of ^{13}Cα or ^{13}C' nuclei in

NCA or NCO experiments, respectively. However, this constitutes only a minor limitation, since most amino acids are at least fractionally labeled on either backbone carbon position, enabling DCP transfer from an adjacent ^{15}N site. On the other hand, the subsequent ^{13}Cα-^{13}C' (or ^{13}C'-^{13}Cα) homonuclear transfer step is vastly abolished because almost no spin systems are simultaneous labeled on both backbone carbon positions. One possible solution to this issue is simply to perform NCXCY experiments with long ^{13}C-^{13}C mixing times, but in such case, multiple ambiguities may be generated that must be resolved via 3D or 4D experiments or by using multiple types of labeling (26). On the other hand, we propose that an alternative protocol that is better suited for sequential resonance assignment of proteins produced with alternating ^{13}C-^{12}C labeling is one that makes use of not only the favorable spin dynamics in these amino acid spin systems but also the characteristic correlation patterns that occur between them.

Considering the potential sequential contacts between the different classes of spin systems, we can deduce which are the types of homonuclear and heteronuclear correlations expected to arise in alternatively labeled samples, which then allows us to select optimal experiments for the different types of sequential connectivity information we wish to extract. Figure 2 enumerates the possible connectivities between the backbone nuclei of two adjacent residues for the various spin system classes, roughly in decreasing order of occurrence. We bring special attention to four types of sequential (from residue (i-1) to residue (i) or vice versa) backbone-to-backbone contacts: one-bond C'(i-1)-N(i); two-bond Cα(i-1)-N(i); two-bond C'(i-1)-Cα(i); and three-bond Cα(i-1)-Cα(i). We neglect the three-bond C'(i-1)-C'(i) contact because of the small chemical shift dispersion of carbonyl resonances typically obstructs the utility of C'-C' correlations. We must also note that, in addition to backbone nuclei, ^{13}Cβ carbons may also play an important role in establishing sequential contacts, mainly via correlations to adjacent ^{13}Cα nuclei. This is the case of the Leu spin system (class 3), in which the complete absence of ^{13}Cα can be compensated to some degree with a fully labeled ^{13}Cβ site.

In addition to sequential correlations, intra-residue correlations are crucial for the assignment process. However, due to the fractional nature of the ^{13}C labeling, the most efficient method to establish intra-residue correlations in alternatively labeled samples is via direct ^{15}N-^{13}C polarization transfer across one or multiple bonds, instead of the one-bond ^{13}C-^{13}C experiments that outline amino acids spin systems in uniformly ^{13}C labeled samples. A small number of intra-residue side-chain correlations are possible in alternatively labeled samples, but they appear

readily in homonuclear correlation experiments, with characteristics (chemical shifts and intensities) that allow their rapid differentiation. In the procedure we outline below, we focus solely on experiments performed on [2-^{13}C] glycerol labeled samples. However, for practical interpretation of data from such samples, it is best to have amino acid-type classifications for ^{13}Cα and ^{13}Cβ nuclei, even ambiguously, as a starting point. Such information can be obtained from a ^{13}C-^{13}C correlation experiment in a uniformly ^{13}C sample, which is typically recorded at the beginning stages of solid-state NMR studies of proteins.

| Res. i - 1 | Res. i | Sequential backbone contacts | |
Cα – C' – N – Cα – C'		^{13}C$^{(i-1)}$-^{13}C$^{(i)}$	^{13}C$^{(i-1)}$-^{15}N$^{(i)}$
		Cα-Cα; C'-Cα	Cα-N; C'-N
		Cα-Cα	Cα-N
		Cα-Cα	Cα-N
		Cα-Cα; C'-Cα	Cα-N; C'-N
		n/a; *	Cα-N
		C'-Cα; *	C'-N
		n/a; *	Cα-N; C'-N
		Cα-Cα; C'-Cα	C'-N
		n/a	C'-N

Figure 2. Possible combinations of residue pairs for the three classes of alternating ^{13}C-^{12}C labeling spin systems, listed in decreasing frequency of occurrence (partial, ^{13}Cα-^{12}C', and ^{12}Cα-^{13}C'), and specific types of sequential and direct ^{13}C-^{13}C and ^{15}N-^{13}C backbone contacts that are the target of our proposed set of 2D experiments. Gray filling denotes labeling, as describe in the caption to Figure 1. Cases where the possible correlations are outside the scope of our proposed experiments are denoted by n/a, while asterisks indicate cases in which sequential correlations to ^{13}Cβ nuclei are likely to be attainable and potentially useful.

The rationale behind the series of experiments we suggest as part of our sequential assignment protocol arises from the amino acid spin system geometry of alternatively labeled proteins and the types of inter-residue backbone contacts. The set of dipolar recoupling experiments we propose are intended to determine connectivities between: directly bonded N-Cα and N-C'; two-bond inter- and intra-residue N-CX (CX = Cα, Cβ, C'); sequential, two-bond Cα-C'; and sequential, multiple-bond Cα-Cα and Cα-Cβ nuclei. Details of the correlation types and optimal experiments to attain them follow.

6.3.1 One-bond N-C: TEDOR-1

Identification of directly bonded heteronuclear ^{15}N-^{13}C pairs is one of the initial steps in protein resonance assignment. Typically, N-C' and N-Cα correlations can be obtained from more general, multiple transfer-step NCXCY experiments, thanks to the chemical shift dispersion between CX and CY resonances (with the exception of Ser and Thr Cα/Cβ shifts in NCACB). However, in samples with limited sensitivity, it may be prudent to record separate NCX and NCXCY spectra because the subsequent C-C mixing step in the latter may result in the attenuation of ^{15}N-^{13}C' and ^{15}N-^{13}Cα crosspeaks beyond the limit of detection. In the case of alternatively labeled samples, as mentioned above, little information is gained from the short C-C mixing step of NCACX and NCOCX experiments; therefore one-bond ^{15}N-^{13}C' and ^{15}N-^{13}Cα connectivities are best obtained with simple NCA and NCO experiments. Such correlations can be recorded most efficiently with either a broadband ^{15}N-^{13}C double cross-polarization (DCP) experiment or with separate NCA and NCO specific-CP experiments. However, here we suggest TEDOR as the experiment of choice for heteronuclear correlations at moderate spinning frequencies (between 8 and 20 kHz) due to its broadband efficiency and superior robustness with respect to rf imperfections, which may be crucial for long averaging periods. TEDOR can be optimized for a one-bond ^{15}N-^{13}C coupling with a mixing period of about 1.5 ms, yielding primarily ^{15}N-^{13}C' and ^{15}N-^{13}Cα correlations, as well as side-chain ^{15}N-^{13}C cross-peaks. In the following discussion we refer to this correlation experiment as TEDOR-1.

The information content of TEDOR-1 may go further than the sequential connectivity information. Since one-bond TEDOR transfer is fairly uniform for all sites, the relative peak intensity can be used to deduce the ^{13}C spin system class of a given N-Cα (or N-C') correlation and that of the preceding (or subsequent) residue. For example a single N(i)-Cα(i) resonance at a

given ^{15}N chemical shift position without an accompanying N(i)-C'(i-1) peak indicates that residue (i-1) is likely to belong to spin system class 1. However, one must exercise caution when making potential assignments based on relative peak intensity since dynamically disordered residues generate weak signals in MAS spectra. Nevertheless, such residues are typically clustered in separate regions of the protein sequence and can be excluded from the analysis, as was the case in PI3-SH3 and other studies of amyloid proteins. With a sufficient signal-to-noise ratio, and for well-ordered regions of the protein backbone, all backbone ^{13}C-labeled carbon resonances appear in TEDOR-1 spectra (including ancillary Gly C' resonances), which constitute a good starting point for the assignment process in alternatively labeled proteins. Finally, the narrow line width of ^{13}C nuclei (due to the absence of J-couplings) renders the NCA and NCO spectra of TEDOR-1 experiments significantly better resolved, and thus more straightforward to analyze, than those obtained with similar experiments in uniformly ^{13}C labeled samples.

6.3.2 Two-bond, sequential C'-Cα: RFDR-2

The sequential information attainable with TEDOR-1, which consists of pairs of N(i)-Cα(i) and N(i)-C'(i-1) heteronuclear correlations, can be difficult to extract in cases of high degeneracy and resonance overlap, particularly in the amide and carbonyl regions. Such resonance overlap is typically resolved by relying on side-chain resonances. On the other hand, homonuclear correlations between C'(i-1) and Cα(i) nuclei provide a direct way of corroborating sequential TEDOR-1 connectivities and resolving ambiguities from ^{15}N resonance overlap. These homonuclear correlations can be established via dipolar recoupling over the two-bond distance separating the C'(i-1)-Cα(i) pair. Indeed, simultaneously labeled C'(i-1)-Cα(i) pairs are fairly abundant in proteins labeled using 2-^{13}C glycerol. These two-bond dipolar couplings tend to truncate the recoupling of weaker ones, dominating ^{13}C-^{13}C polarization transfer in broadband dipolar recoupling experiments, and thus they appear readily in correlation spectra. A remarkable feature of this type of correlations in alternatively labeled proteins is that the C'-Cα spectral region is virtually free of intra-residue C'(i)-Cα(i) one-bond correlations, granting outstanding resolution to sequential C'(i-1)-Cα(i) correlations.

In order to obtain C'(i-1)-Cα(i) cross-peaks in alternatively labeled proteins we suggest the application of RFDR optimized in a broadband fashion for two-bond polarization transfer with a mixing time of 5 to 8 ms, referred to here as RFDR-2. While many other homonuclear

mixing schemes could be implemented for this purpose, RFDR offers several advantages including excellent broadband efficiency, functionality at a wide range of spinning frequencies, compensation against pulse and phase imperfections, and robustness to heteronuclear depolarization effects, which facilitate its application during extended mixing periods. Of great technical importance, the absence of one-bond dipolar couplings for most ^{13}C spin systems with alternating labeling enables the application of broadband dipolar recoupling schemes such as RFDR to obtain two-bond non-relayed cross-peaks with high transfer efficiency thanks to the attenuate of dipolar truncation effects.

We note that sequential C'(i-1)-Cα(i) pairs derived from congruent correlations to N(i) in TEDOR-1 data should in principle always lead to observable cross-peaks in RFDR-2 spectra. However, the signal intensity with which they appear varies greatly because these cross-peaks depend on the fractional labeling on two carbon sites simultaneously. Particularly weak signals in RFDR-2 spectra can be expected from C'(i-1)-Cα(i) pairs involving E, Q, P, or R in position (i), and K or G in position (i-1). As in the case of TEDOR-1 data, RFDR-2 yields no sequential connectivities for C'(i-1)-Cα(i) pairs involving Leu in position (i) or spin system class 1 residues in position (i-1). Also as in TEDOR-1 spectra, the reduced ^{13}C line widths of alternatively labeled proteins grants high resolution to the C' and Cα spectral regions, facilitating the analysis of RFDR-2 correlation spectra.

6.3.3 Sequential Cα-Cα and Cα-Cβ: BASE RFDR

Homonuclear correlations between Cα nuclei of adjacent residues constitute a potentially valuable approach to obtain sequential assignments in both uniformly and alternatively labeled proteins. In uniformly labeled samples, such correlations can be generated by multiple polarization steps, as in the CA(NCO)CA experiment (44), or via homonuclear mixing experiments, such as PDSD under weak coupling conditions (45). In the first approach, polarization is relayed through two cross-polarization periods followed by homonuclear mixing, resulting in inherently low overall efficiencies for the final transferred polarization, but with the benefit of encoding (i) to (i-1) directionality via the CA-N transfer step. The second approach consists of medium to long homonuclear mixing periods via PDSD or DARR, and relies on being able to identify Cα(i)-Cα(i±1) cross-peaks in typically crowded ^{13}C-^{13}C spectra. While the homonuclear approach is robust, its efficiency and interpretability improve significantly in

alternatively labeled samples, as has been recently described (*25*), due to enhanced resolution in the Cα region of the spectra and the favorable spin dynamics afforded by spin dilution.

We have recently demonstrated that alternatively ^{13}C-^{12}C labeled proteins are amenable to efficient dipolar recoupling schemes and have introduced band-selective RFDR (BASE RFDR) as an effective long-range ^{13}C-^{13}C polarization transfer scheme and an interesting alternative to diffusion-based methods, yielding enhanced Cα-Cα correlation spectra in a protein labeled using 2-^{13}C glycerol (*28*). Furthermore, since this scheme consists of direct zero-order recoupling of ^{13}C-^{13}C interactions, the optimal mixing period is much shorter than in diffusion methods and displays a stronger dependence on the inter-nuclear distance. Thus, a BASE RFDR mixing time can be selected that maximizes polarization transfer through the 3.8 Å distance between Cα(i)-Cα(i±1) pairs, on the order of 15-18 ms, while the equivalent optimal value in PDSD experiments may vary between 300 and 700 ms. We therefore suggest BASE RFDR as the scheme of choice to establish sequential Cα-Cα correlations in protein samples labeled using [2-^{13}C] glycerol, as part of our assignment protocol. In addition, depending on the pattern of neighboring residues, Cα(i)-Cβ(i±1) correlations may be observed with sufficient intensity to assist in the sequential assignment process.

3.4. Two-bond N-C: TEDOR-2

As mentioned above, alternating ^{13}C-^{12}C patterns in the peptide backbone precludes the application of NCOCX experiments with short C-C mixing times that provide sequential correlations in uniformly labeled proteins. However, sequential connectivities can be readily established in alternatively labeled proteins via direct ^{15}N-^{13}C recoupling of medium-range contacts. In particular, $^{15}N(i)$-$^{13}Cα(i-1)$ correlations can be obtained efficiently through TEDOR experiments with a mixing period around 6 ms, optimized for a two-bond internuclear distance, referred to here as TEDOR-2. This type of correlation experiment comprises a powerful way of establishing sequential backbone assignments due to its unambiguous character and sufficient chemical shift dispersion in the Cα region. By comparison to TEDOR-1 spectra, sequential correlations can be quickly identified in TEDOR-2 spectra. While both N(i)-Cα(i) and N(i)-Cα(i-1) cross-peaks are observed simultaneously, they occur with similar intensities, which facilitates spectral analysis, and the narrow $^{13}Cα$ line widths compensates to some degree for the spectral crowding compared to N(CO)CA spectra of uniformly labeled samples.

Figure 3. Complete set of four 2D spectra used in our approach to assign the majority of the 2-GB1 backbone. Top: TEDOR-1. Bottom: TEDOR-2. Middle left: RFDR-2, right: BASE RFDR

6.4 Results

Applying the set of four 2D experiments to 2-GB1, together with the analysis described above, we were able to obtain sequential assignments for the majority of the protein backbone. Each experiment contributed sequential correlations between neighboring residues provided that (1) the target carbon sites were ^{13}C labeled in this sample prepared with [2-^{13}C] glycerol, and (2) the resonances could be at least partially resolved in the two frequency dimensions. The four spectra are shown in Figure 3, including labels for all resonance assignments obtained through their analysis. Correlations between the segment T11-L12-K13-G14 are shown with blue lines and red peak labels. For illustrative purposes, the highlighted correlations in Figure 3 include the trio LKG, which has one residue of each spin system class ($^{12}C\alpha$-$^{13}C'$, partial, and $^{13}C\alpha$-$^{12}C'$, respectively) and has the same pattern as the three-residue segment depicted in Figure 1. Note that multiple lines crossing a single cross-peak indicate multiple correlations for a given resonance, which helps to disambiguate assignment degeneracies.

The combined analysis of all four spectra resulted in the following overall results for 2-GB1. Sequential ^{13}C-^{13}C and/or ^{15}N-^{13}C correlations were established between all 56 residues, arising from contacts found in at least one of the experiments, with most of them having multiple sequential connectivities. The individual breakdown for each spectrum demonstrates the redundancy of information attainable, which aids in the resonance assignment process. TEDOR-2 yielded 43 sequential $^{15}N(i)$-$^{13}C\alpha(i-1)$ contacts. BASE RFDR provided sequential correlations between $^{13}C\alpha$- $^{13}C\alpha$ or $^{13}C\alpha$- $^{13}C\beta$ for 48 of the total 55 possible connectivities. TEDOR-1, which in principle could provide all 55 one-bond $^{15}N(i)$-$^{13}C'(i-1)$ correlations in a uniformly $^{13}C,^{15}N$ labeled samples, showed only 34 of them due to the absence of many ^{13}C carbonyl sites. Those, however are all the $^{13}C'$ resonances, which helps in the interpretation of the RFDR-2 spectra, which yielded 28 sequential $^{13}C\alpha(i)$- $^{13}C'(i-1)$ correlations. Thus for most pairs of residues, multiple sequential correlations, to any of their labeled nuclei, were used to assign their backbone resonanes (^{15}N, $^{13}C'$, $^{13}C\alpha$, and $^{13}C\beta$) in a sequential and unambiguous manner. Finally, we must note that in order to determine the amino acid type to which a given $^{13}C\alpha$ resonance belonged, we compared our spectra to a one-bond ^{13}C-^{13}C correlation spectrum of U-GB1, and the analysis was performed in a blindfold fashion.

6.5 Discussion

The resonance assignment process is most efficient when highly complementary experiments are used, providing distinct pieces of information relating a pair of residues. On the other hand, the assignment process is made increasingly reliable when redundant information is obtained from different experiments. Both of these features are available in the assignment protocol we have introduced. For example, the information extractable from RFDR-2 data is a subset of that from TEDOR-1 data, however they corroborate one another via direct ^{13}C'(i)-^{13}Cα(i+1) correlations and ^{15}N-encoded ^{13}C'(i)-^{15}N(i+1)/^{15}N(i+1)-^{13}Cα(i+1) correlation pairs, respectively. More importantly, they provide a link between often-crowded ^{13}C' and ^{15}N spectral regions to the more dispersed ^{13}Cα region via sequential correlations, which can be connected to TEDOR-2 and BASE RFDR spectra.

A number of other polarization transfer schemes could potentially be applied in place of some of the experiments chosen for our analysis, with both favorable and unfavorable consequences. For example, rotational resonance tickling (R^2T) could yield strong ^{13}C'(i)-^{13}Cα(i+1) correlations, but may result in line broadening and chemical shift distortions. A clear alternative to both RFDR-2 and BASE RFDR is proton-driven spin diffusion (PDSD) or dipolar-assisted rotary resonance (DARR), which are broadband and yield strong ^{13}C'(i)-^{13}Cα(i+1) and ^{13}Cα(i)-^{13}Cα(i±1) correlations. However, the longer mixing times required for aliphatic-aliphatic correlations result in long-range ^{13}C'(i)-^{13}Cα(j) contacts that obscure the sequential ones, and thus the experiments are best applied with two different mixing times, ~50 ms for ^{13}C'(i)-^{13}Cα(i+1) and ~200 ms for ^{13}Cα(i)-^{13}Cα(i±1) correlations. Therefore, little advantages are found in general with PDSD/DARR compared to RFDR-2 and BASE RFDR.

At higher MAS frequencies (≥ 20 kHz), second-order techniques such as PAR, MIRROR, RESORT, PARIS, etc., may be applied effectively, although it is not clear whether they are more efficient than methods used here, for the target two- and three-bond distances. In fact, since some of these schemes provide long-range correlations after relatively short mixing times, they are likely to obscure the sequential contacts necessary for unambiguous assignment. This is specially the case with heteronuclear mixing via NHHC or PAINCP, which are possible alternatives to TEDOR-2, but result in long-range correlations even with short mixing times that interfere with the identification of sequential ^{15}N(i)-^{13}Cα(i-1) contacts.

6.6 Conclusions

We have focused on two-dimensional correlation experiments for the implementation of our sequential resonance assignment strategy. Such implementation results in short, efficient data collection periods and is highly beneficial for cases of limited sensitivity, typically encountered in challenging biological applications of MAS NMR spectroscopy. We anticipate that the resolution afforded by 2D spectra of [2-^{13}C] glycerol labeled samples such as those presented here is sufficient for the study of small to moderately sized proteins (perhaps up to 150 residues). More complicated systems may be tackled by expanding some of our experiments into additional frequency dimensions. For example, RFDR-2 may be preceded by ^{1}H-^{15}N CP, t_1 evolution on ^{15}N, and a ^{15}N-^{13}C double CP step to obtain N(i)-Ca(i)-C'(i-1) or N(i)-C'(i-1)-Ca(i) correlations in a 3D experiment. Note that in such cases the projected F_1-F_3 dimension differs from the corresponding TEDOR-1 2D spectral region in that ^{15}N(F_1)-^{13}C(F_3) cross-peaks only appear if the corresponding intermediate carbon nucleus is ^{13}C labeled to form the ^{13}C'(i-1)-^{15}N(i)-^{13}C(i) triplet. Such experiments thus act as filters for TEDOR-1 correlations, even if implemented in 2D fashion (without the intermediate ^{13}C chemical shift evolution period). A similar approach can be used with BASE RFDR mixing to yield N(i)-Ca(i)-Ca(i±1) 3D correlation spectra with ample resolution and sensitivity. However, 3D implementation benefits from a more general NCXCY approach using short and long PDSD or DARR mixing periods and alternating labeling has been demonstrated recently in various studies. Those studies relied on multiple 3D experiments recorded on both [2-^{13}C] and [1,3-^{13}C] glycerol labeled samples and a uniformly labeled sample, utilizing the differences and similarities among all spectra to resolve multiple ambiguities and establish sequential assignments. Such an exhaustive approach is necessary for very large proteins. In contrast, our assignment strategy relies mainly on a single sample (produced with [2-^{13}C] glycerol) and while it is limited by the resolution available in 2D experiments it results in highly efficient analysis and the rapid elucidation of backbone resonances, and thus provides an important initial step in protein structure determination efforts via MAS NMR spectroscopy.

References

(1)   Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, Wiley Interscience, New York, NY.

(2)   Cavanagh, J., Fairbrother, W. J., Palmer III, A. G., and Skelton, N. J. (1996) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, San Diego, CA.

(3)   Cavalli, A., Salvatella, X., Dobson, C. M., and Vendruscolo, M. (2007) Protein structure determination from NMR chemical shifts. *Proceedings of the National Academy of Sciences of the United States of America 104*, 9615-9620.

(4)   Robustelli, P., Cavalli, A., and Vendruscolo, M. (2008) Determination of protein structures in the solid state from NMR chemical shifts. *Structure 16*, 1764-9.

(5)   Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J. M., Liu, G., Eletsky, A., Wu, Y., Singarapu, K. K., Lemak, A., Ignatchenko, A., Arrowsmith, C. H., Szyperski, T., Montelione, G. T., Baker, D., and Bax, A. (2008) Consistent blind protein structure generation from NMR chemical shift data. *Proceedings of the National Academy of Sciences of the United States of America 105*, 4685-90.

(6)   Robustelli, P., Cavalli, A., Dobson, C. M., Vendruscolo, M., and Salvatella, X. (2009) Folding of small proteins by Monte Carlo simulations with chemical shift restraints without the use of molecular fragment replacement or structural homology. *J. Phys. Chem. B 113*, 7890-6.

(7)   Shen, Y., Vernon, R., Baker, D., and Bax, A. (2009) De novo protein structure generation from incomplete chemical shift assignments. *Journal of biomolecular NMR 43*, 63-78.

(8)   Pauli, J., Baldus, M., van Rossum, B., de Groot, H., and Oschkinat, H. (2001) Backbone and side-chain 13C and 15N signal assignments of the alpha-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 Tesla. *ChemBioChem 2*, 272-81.

(9)   Petkova, A. T., Baldus, M., Belenky, M., Hong, M., Griffin, R. G., and Herzfeld, J. (2003) Backbone and side chain assignment strategies for multiply labeled membrane peptides and proteins in the solid state. *J. Magn. Reson. 160*, 1-12.

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

(11)　McDermott, A., Polenova, T., Bockmann, A., Zilm, K. W., Paulson, E. K., Martin, R. W., Montelione, G. T., and Paulsen, E. K. (2000) Partial NMR assignments for uniformly (13C, 15N)-enriched BPTI in the solid state. *Journal of biomolecular NMR 16*, 209-19.

(12)　Bockmann, A., Lange, A., Galinier, A., Luca, S., Giraud, N., Juy, M., Heise, H., Montserret, R., Penin, F., and Baldus, M. (2003) Solid state NMR sequential resonance assignments and conformational analysis of the 2x10.4 kDa dimeric form of the Bacillus subtilis protein Crh. *J. Biomol. NMR 27*, 323-39.

(13)　Igumenova, T. I., McDermott, A. E., Zilm, K. W., Martin, R. W., Paulson, E. K., and Wand, A. J. (2004) Assignments of carbon NMR resonances for microcrystalline ubiquitin. *Journal of the American Chemical Society 126*, 6720-7.

(14)　Igumenova, T. I., Wand, A. J., and McDermott, A. E. (2004) Assignment of the backbone resonances for microcrystalline ubiquitin. *Journal of the American Chemical Society 126*, 5323-31.

(15)　Franks, W. T., Zhou, D. H., Wylie, B. J., Money, B. G., Graesser, D. T., Frericks, H. L., Sahota, G., and Rienstra, C. M. (2005) Magic-angle spinning solid-state NMR spectroscopy of the beta1 immunoglobulin binding domain of protein G (GB1): 15N and 13C chemical shift assignments and conformational analysis. *Journal of the American Chemical Society 127*, 12291-305.

(16)　Marulanda, D., Tasayco, M. L., Cataldi, M., Arriaran, V., and Polenova, T. (2005) Resonance assignments and secondary structure analysis of E. coli thioredoxin by magic angle spinning solid-state NMR spectroscopy. *J. Phys. Chem. B 109*, 18135-45.

(17)　Kobayashi, M., Matsuki, Y., Yumen, I., Fujiwara, T., and Akutsu, H. (2006) Signal assignment and secondary structure analysis of a uniformly [13C, 15N]-labeled membrane protein, H +-ATP synthase subunit c, by magic-angle spinning solid-state NMR. *Journal of biomolecular NMR 36*, 279-93.

(18) Siemer, A. B., Ritter, C., Steinmetz, M. O., Ernst, M., Riek, R., and Meier, B. H. (2006) 13C, 15N resonance assignment of parts of the HET-s prion protein in its amyloid form. *Journal of biomolecular NMR 34*, 75-87.

(19) Chen, L., Kaiser, J. M., Polenova, T., Yang, J., Rienstra, C. M., and Mueller, L. J. (2007) Backbone assignments in solid-state proteins using J-based 3D heteronuclear correlation spectroscopy. *Journal of the American Chemical Society 129*, 10650-1.

(20) Etzkorn, M., Martell, S., Andronesi, O. C., Seidel, K., Engelhard, M., and Baldus, M. (2007) Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed. 46*, 459-62.

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

(22) LeMaster, D., and Kushlan, D. (1996) Dynamical Mapping of E. coli Thioredoxin via 13C NMR Relaxation Analysis. *Journal of the American Chemical Society 118*, 9255-9264.

(23) Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K., and Oschkinat, H. (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature 420*, 98-102.

(24) Castellani, F., van Rossum, B. J., Diehl, A., Rehbein, K., and Oschkinat, H. (2003) Determination of solid-state NMR structures of proteins by means of three-dimensional 15N-13C-13C dipolar correlation spectroscopy and chemical shift analysis. *Biochemistry 42*, 11476-83.

(25) Becker, J., Ferguson, N., Flinders, J., van Rossum, B. J., Fersht, A. R., and Oschkinat, H. (2008) A sequential assignment procedure for proteins that have intermediate line widths in MAS NMR spectra: amyloid fibrils of human CA150.WW2. *ChemBioChem 9*, 1946-52.

(26) Higman, V. A., Flinders, J., Hiller, M., Jehle, S., Markovic, S., Fiedler, S., van Rossum, B. J., and Oschkinat, H. (2009) Assigning large proteins in the solid state: a MAS NMR

resonance assignment strategy using selectively and extensively 13C-labelled proteins. *Journal of biomolecular NMR 44*, 245-60.

(27)   Bayro, M. J., Huber, M., Ramachandran, R., Davenport, T. C., Meier, B. H., Ernst, M., and Griffin, R. G. (2009) Dipolar truncation in magic-angle spinning NMR recoupling experiments. *J. Chem. Phys. 130*, 114506.

(28)   Bayro, M. J., Maly, T., Birkett, N. R., Dobson, C. M., and Griffin, R. G. (2009) Long-range correlations between aliphatic 13C nuclei in protein MAS NMR spectroscopy. *Angew. Chem. Int. Ed. 48*, 5708-10.

(29)   Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D., and Dobson, C. M. (1998) Amyloid Fibril Formation by an SH3 Domain. *Proc. Natl. Acad. Sci. U.S.A. 95*, 4224-4228.

(30)   Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V., and Griffin, R. G. (1995) Heteronuclear decoupling in rotating solids. *J. Chem. Phys. 103*, 6951-6958.

(31)   Jaroniec, C., Filip, C., and Griffin, R. (2002) 3D TEDOR NMR Experiments for the Simultaneous Measurement of Multiple Carbon-Nitrogen Distances in Uniformly $^{13}C,^{15}N$-Labeled Solids. *Journal of the American Chemical Society 124*, 10728-10742.

(32)   Michal, C. A., and Jelinski, L. W. (1997) REDOR 3D: Heteronuclear distance measurements in uniformly labeled and natural abundance solids. *Journal of the American Chemical Society 119*, 9059-9060.

(33)   Bennett, A. E., Griffin, R. G., Ok, J. H., and Vega, S. (1992) Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys. 96*, 8624-8627.

(34)   Bennett, A. E., Rienstra, C. M., Griffiths, J. M., Zhen, W., Lansbury, J., and Griffin, R. G. (1998) Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys. 108*, 9463-9479.

(35)   Bayro, M. J., Ramachandran, R., Caporini, M. A., Eddy, M. T., and Griffin, R. G. (2008) Radio frequency-driven recoupling at high magic-angle spinning frequencies: Homonuclear recoupling sans heteronuclear decoupling. *J. Chem. Phys. 128*, 052321.

(36)   De Paëpe, G., Bayro, M. J., Lewandowski, J., and Griffin, R. G. (2006) Broadband homonuclear correlation spectroscopy at high magnetic fields and MAS frequencies. *J. Am. Chem. Soc. 128*, 1776-7.

(37)   Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277-93.

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

(39)   Szeverenyi, N. M., Sullivan, M. J., and Maciel, G. E. (1982) Observation of spin exchange by two-dimensional fourier transform 13C cross polarization-magic-angle spinning. *J. Magn. Reson. 47*, 462-475.

(40)   Meier, B. H. (1999) Polarization transfer and spin diffusion in solid-state NMR, in *Advances in Magnetic and Optical Resonance* pp 1.

(41)   Grommek, A., Meier, B. H., and Ernst, M. (2006) Distance information from proton-driven spin diffusion under MAS. *Chem. Phys. Lett. 427*, 404-409.

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

(43)   Morcombe, C. R., Gaponenko, V., Byrd, R. A., and Zilm, K. W. (2004) Diluting abundant spins by isotope edited radio frequency field assisted diffusion. *Journal of the American Chemical Society 126*, 7196-7.

(44)   Franks, W. T., Kloepper, K. D., Wylie, B. J., and Rienstra, C. M. (2007) Four-dimensional heteronuclear correlation experiments for chemical shift assignment of solid proteins. *Journal of biomolecular NMR 39*, 107-31.

(45)   Seidel, K., Lange, A., Becker, S., Hughes, C. E., Heise, H., and Baldus, M. (2004) Protein solid-state NMR resonance assignments from (13C,13C) correlation spectroscopy. *Phys. Chem. Chem. Phys. 6*, 5090-5093.

Part II

Structural Characterization of Amyloid Fibrils and Bacterial Organelles via MAS NMR Spectroscopy

Chapter 7. High-resolution MAS NMR analysis of PI3-SH3 amyloid fibrils: Backbone conformation and implications for protofilament assembly and structure

Adapted from a manuscript in preparation by Marvin J. Bayro, Thorsten Maly, Neil R. Birkett, Christopher M. Dobson, and Robert G. Griffin

Summary

The SH3 domain of the PI3 kinase (PI3-SH3 or PI3K-SH3) readily aggregates into fibrils *in vitro* and has served as an important model system to investigate the molecular properties and mechanism of formation of amyloid fibrils. We describe the molecular conformation of PI3-SH3 in amyloid fibril form as revealed by magic-angle spinning (MAS) solid-state NMR spectroscopy. The MAS NMR spectra of these fibrils display excellent resolution, with narrow ^{13}C and ^{15}N line widths, representing a high degree of structural order and the absence of extensive molecular motion for the majority of the polypeptide chain. We have identified the spin-systems of 82 of the 86 residues in the protein, and obtained sequential resonance assignments for 75 of them. Chemical shift analysis indicates that the protein subunits making up the fibril adopt a compact conformation consisting of four well-defined β-sheet regions and four random-coil elements with varying degrees of local dynamics or disorder. The backbone conformation of PI3-SH3 in fibril form differs significantly from that of the native state of the protein, both in secondary structure and in the location of dynamic/disordered segments. The site-specific MAS NMR analysis of PI3-SH3 fibrils we report here is compared with previously published mechanistic and structural data, resulting in a detailed interpretation of the factors that mediate fibril formation by PI3-SH3 and allowing us to propose a possible model of the core structure of the fibrils. Our results confirm the structural similarities between PI3-SH3 fibrils and amyloids directly related to degenerative or infectious diseases.

7.1 Introduction

Amyloid fibrils are filamentous structures resulting from the spontaneous self-assembly of otherwise soluble peptides and proteins (*1-4*). A large number of human disorders, including Alzheimer's and Parkinson's diseases, type 2 diabetes and a variety of systemic amyloidoses, are associated with the formation of such macromolecular assemblies (*1, 5, 6*). In each of these pathological conditions, a specific peptide or protein, or protein fragment, transforms from its usual soluble native form into insoluble amyloid fibrils that can accumulate in a variety of organs and tissues. It is not yet clear, in what manner the process of amyloid fibril formation leads to the pathogenic behavior that characterizes these diseases; however, in the systemic conditions it is likely that the fibrils themselves contribute very significantly to the process of organ damage (*5*). Furthermore, an increasing number of proteins with no link to deposition diseases have been found to form functional amyloid structures in organisms ranging from bacteria to mammals (*7-10*).

The mechanisms of amyloid fibril formation, and the structures of the amyloid fibrils themselves, are inherently interesting topics and raise important questions from a physical as well as a biological perspective (*11*). To explain the molecular basis of amyloid fibril formation by globular proteins, it has been proposed that a critical early step is the partial unfolding of the protein, resulting in flexible conformers that expose aggregation-prone regions of the sequence, which are largely buried in the native state, to the external environment (*12*). In the example investigated here, fibrils are formed *in vitro* by the SH3 domain of the p85α subunit of bovine phosphatidylinositol 3-kinase (PI3-SH3) under acidic conditions (*13*). The structure of the protein in its native state is well-characterized by X-ray crystallography (*14*) and solution NMR spectroscopy (*15-17*), and studies of the acid unfolded state provide clear evidence that the protein adopts a partially folded conformation prior to fibril formation (*13, 18*). However, to understand in detail the mechanism by which PI3-SH3 fibrils form and to characterize their properties, it is necessary to elucidate the structure of the fibrillar state at atomic resolution. In combination with information available from previous biophysical studies on this system, the characterization of PI3-SH3 fibrils is likely to reveal significant insights into the universal features of amyloid formation and structure (*19-23*).

Proteins that readily aggregate to form amyloid fibrils do not share any obvious sequence identity or structural homology to one another. Prior to their transformation into fibrils,

amyloidogenic proteins can possess a variety of secondary structure elements, but in their fibrillar state, as revealed by X-ray fiber diffraction data, they adopt a cross-β structure, in which arrays of β-strands are oriented approximately perpendicular to the long axis of the fibril. In addition, fibrils from a wide variety of peptides and proteins typically display a long, unbranched and frequently twisted morphology (24), suggesting that the ability to form fibrils is an inherent property of polypeptide chains and not restricted to pathological cases, although the propensity to aggregate and the molecular details of the resulting structures are highly sequence-dependent (1).

The physical properties of amyloid fibrils impede their study by conventional high-resolution structural techniques such as single-crystal X-ray crystallography and solution NMR spectroscopy. Specifically, they do not possess long-range three-dimensional order, do not diffract to high resolution, are typically insoluble, and have high molecular weights. Therefore, the majority of structural information on these species has been obtained through complementary techniques such as transmission electron microscopy, atomic force microscopy, and X-ray fiber diffraction (24-26). In the past decade, however, significant advances in magic angle spinning solid-state NMR (MAS NMR), notably dipolar recoupling methodology, has enabled the *de novo* determination of the structures of complex biological molecules in the solid state. In particular, high-resolution MAS NMR structures have been obtained for short peptides and microcrystalline proteins (27-30) and for an 11-residue fragment of human transthyretin (TTR) in its amyloid fibril form (31, 32). Furthermore, detailed structural information, and in some cases structural models, have been obtained by MAS NMR spectroscopy for several amyloid fibrils and prion proteins such as α-synuclein (33), the amyloid-β (Aβ) peptide (34), the GNNQQNY fragment of Sup35 (35), HET-s (36) and a fragment of the yeast prion protein Ure2p (37). MAS NMR experiments based on dipole-dipole interactions permit the analysis of rigid structural domains, such as the core of amyloid fibrils. In addition, the incorporation of techniques originally developed for liquid-state NMR studies into MAS NMR experiments (38) facilitates the analysis of highly flexible regions of the fibril (33, 39). Solid-state NMR spectroscopy is therefore a highly versatile method for the study of the structure and dynamics of biological macromolecules, even those as challenging to characterize as amyloid fibrils (40).

In this article we describe site-specific structural characteristics of PI3-SH3 in amyloid fibril form determined via MAS NMR spectroscopy. The dipolar correlation spectra of PI3-SH3 exhibit excellent resolution, allowing us to identify 82 spin systems for the 86-residue protein.

Analysis of samples prepared with alternating ^{13}C-^{12}C labeling (28, 41, 42) resulted in site-specific assignments for the majority of the observed ^{13}C and ^{15}N resonances. Spectral and chemical shift analyses suggest that the backbone conformation of PI3-SH3 in fibril form consists of well-defined secondary structure elements interrupted by short segments of less regular structure. Interpreting the results from previous structural and mechanistic studies on PI3-SH3, in light of the molecular conformation of the fibril subunits reported here, provides key insights into the process of fibril formation and overall fibril structure. The amyloid characteristics at the molecular level of PI3-SH3 described here are similar to those of disease-associated protein fibrils, and thus support strongly the idea that amyloid fibrils are commonly accessible structural states (1, 4) and highlight the importance of studying model systems in order to enhance our understanding of the underlying principles of amyloid assembly and structure.

7.2 Materials and Methods

7.2.1 Sample Preparation

The 86-residue, 9.6 kDa PI3-SH3 domain was expressed as a his_6-tag fusion construct in BL21(DE3)pLysS *E. Coli* cells using M9 minimal medium as described previously (*13*). For uniformly ^{13}C and ^{15}N labeled material (U-PI3-SH3), the medium was supplemented with ^{15}N-ammonium chloride and ^{13}C-glucose as the sole nitrogen and carbon sources. A sample in which ^{13}C is incorporated in approximately every other carbon site (2-PI3-SH3) was prepared using [2-^{13}C]glycerol and $NaH_{13}CO_3$ (Cambridge Isotopes, Andover MA) as the sole sources of carbon (*28, 41*). The protein was isolated by nickel affinity chromatography and further purified by size exclusion chromatography. Fibrils were prepared by incubating 1.0 mM monomeric PI3-SH3 in aqueous solution at pH 2.0 for two to three weeks at 25 °C. The fibril morphology was verified periodically by transmission electron microscopy. Upon completion of fibril formation into a long, straight morphology, the fibrils were collected from solution by centrifugation. For MAS NMR experiments, the buffer was changed to a mixture of d_5-glycerol/water (60/40, w/w) via repeated cycles of centrifugation and re-suspension, prior to an ultracentrifugation step for concentration. The wax-like pellet contained approximately 0.5 mg of fibrils per mg of sample material. Finally the pellet was transferred into a MAS rotor by centrifugation.

7.2.2 MAS NMR Spectroscopy and Data Analysis

NMR experiments were performed on custom-designed spectrometers (courtesy of D. J. Ruben, Francis Bitter Magnet Laboratory) operating at ^{1}H Larmor frequencies of 700 and 750 MHz. Experiments at 700 MHz were performed using a Varian 3.2 mm triple-resonance probe (Varian Inc, Palo Alto, CA), while at 750 MHz a Bruker 2.5 mm triple-resonance probe (Bruker BioSpin, Billerica, MA) was used. The sample temperature was maintained by using a stream of nitrogen gas (2 °C). Correlation experiments utilized ramped cross-polarization (CP) and two-pulse phase modulation (TPPM) heteronuclear decoupling (*43*), and were performed at MAS frequencies ($\omega_r/2\pi$) between 10.0 and 28.5 kHz. For homonuclear correlation experiments RFDR (*44-46*), DREAM (*47*), CMAR (*48*), and PDSD (*49*) mixing schemes were used with U-PI3-SH3, while RFDR, PDSD, and BASE RFDR (*50*) were used with 2-PI3-SH3. CMAR experiments were recorded at $\omega_r/2\pi = 28.571$ kHz using a ^{13}C r.f. field of $\omega_1/2\pi = 100.0$ kHz.

Broadband RFDR experiments were recorded either at $\omega_r/2\pi$ = 18.182 kHz using ^{13}C $\omega_1/2\pi$ = 40 kHz and high-power ^{1}H decoupling or at $\omega_r/2\pi$ = 28.571 kHz using ^{13}C $\omega_1/2\pi$ = 120 kHz without ^{1}H decoupling during the mixing time. PDSD experiments were recorded at $\omega_r/2\pi$ = 10 to 12.5 kHz, with mixing times between 50 and 300 ms. DREAM correlations over the ^{13}C aliphatic ^{13}C region were recorded at $\omega_r/2\pi$ = 16.667 and 28.571 kHz. Heteronuclear correlations were obtained using SPECIFIC CP (51) and TEDOR (52, 53) mixing (1.6 ms and 6.0 ms mixing times). ^{13}C and ^{15}N chemical shifts were indirectly referenced to DSS (54) and liquid ammonia (55), respectively. NMR data were processed using NMRPipe (56) and analyzed with Sparky (T.D. Goddard and D.G. Kneller, SPARKY 3, University of California, San Francisco, USA). Secondary structure elements were predicted by calculating the chemical shift deviation from random coil values (57) using the secondary shift values listed by Zhang et al. (58). Backbone torsion angles ϕ and ψ were predicted with the TALOS program (59), version 2007.068.09.07.

7.3 Results

7.3.1 Fibril Homogeneity and Spectral Quality

Amyloid fibrils formed by PI3-SH3 adopt preferentially a single morphology, and are remarkably stable (23). No signals from monomeric PI3-SH3 were observed in MAS NMR spectra. Furthermore, the five samples employed in our analysis show exclusively a single set of chemical shifts and no significant variations in resonance positions were evident between the five different preparations. A second form, uniquely observed in a sixth sample, showed some differences from the first form generated under apparently identical conditions, but was not pursued further in this study, which focuses solely on the first, dominant form.
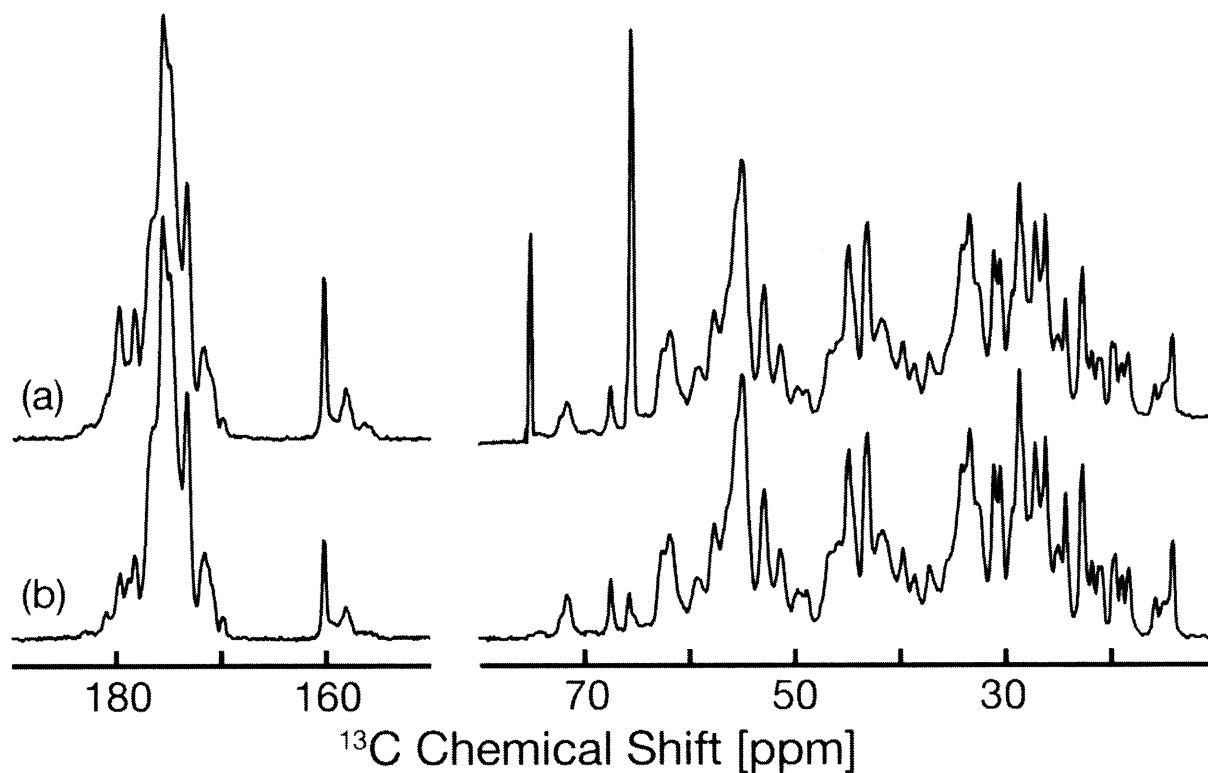
Figure 1. 1D ^{13}C MAS NMR spectra of U-PI3-SH3 fibrils recorded at 750 MHz ^{1}H Larmor frequency and $\omega_r/2\pi$ = 16.67 kHz, at the temperature of 2 °C. (a) Direct ^{13}C spectrum (Bloch decay), scaled by factor of 2.5. (b) ^{13}C cross-polarization spectrum obtained with 1.5 ms contact time.

199

One-dimensional (1D) MAS NMR spectra of PI3-SH3 fibrils are of very high resolution and permit the observation of important structural features. Two spectra of uniformly labeled PI3-SH3 (U-PI3-SH3) are shown in Figure 1, a direct-detected ^{13}C MAS spectrum (top) and a ^{1}H-^{13}C CP MAS spectrum (bottom). These two spectra are remarkably similar to each other in most of their salient features, the only exception being slightly different intensities for a small number of side-chain resonances. This finding indicates that the protein backbone in PI3-SH3 is in a very well defined conformation such that CP enhancement is essentially uniform throughout the polypeptide chain. Indeed, MAS NMR spectra of PI3-SH3 utilizing the INEPT method, aimed at selectively exciting ^{13}C sites in highly flexible regions, yield no protein signals in the temperature range explored in these studies (-10° C to 25° C). CP spectra at various temperatures are show in Figure S1 of the supporting information. The absence of highly mobile carbon sites, and the close similarity between the CP and the direct ^{13}C polarization spectra, indicate that the majority of the PI3-SH3 polypeptide chain adopts a very rigid conformation in its fibrillar state, in contrast to observations for several other fibrillar systems studied recently, such as α-synuclein (33), Het-s (39), and a human prion protein (60), where the constituent protein subunits are observed to have both rigid and mobile segments.

A very high degree of structural homogeneity is clearly evident from the 2D ^{13}C-^{13}C correlation spectra of U-PI3-SH3, as illustrated in Figures 2 and 3. The average ^{13}C linewidth is less than 0.5 ppm (94 Hz at a ^{1}H Larmor frequency of 750 MHz), comparable to that observed in uniformly labeled microcrystalline proteins (28, 29, 61-63). Since the spectral resolution of fibril samples depends on their hydration levels (64), PI3-SH3 fibrils were dispersed in a glycerol/water mixture (60/40, w/w) prior to recording spectra to inhibit dehydration of the samples by exploiting the low vapor pressure and hygroscopic properties of glycerol. Indeed, MAS NMR spectra of PI3-SH3 fibrils in buffer alone (pH 2.0) show considerably larger linewidths under the same experimental conditions than those of fibrils in the glycerol/water mixture described above (Figure S2), while the chemical shifts show only marginal changes.
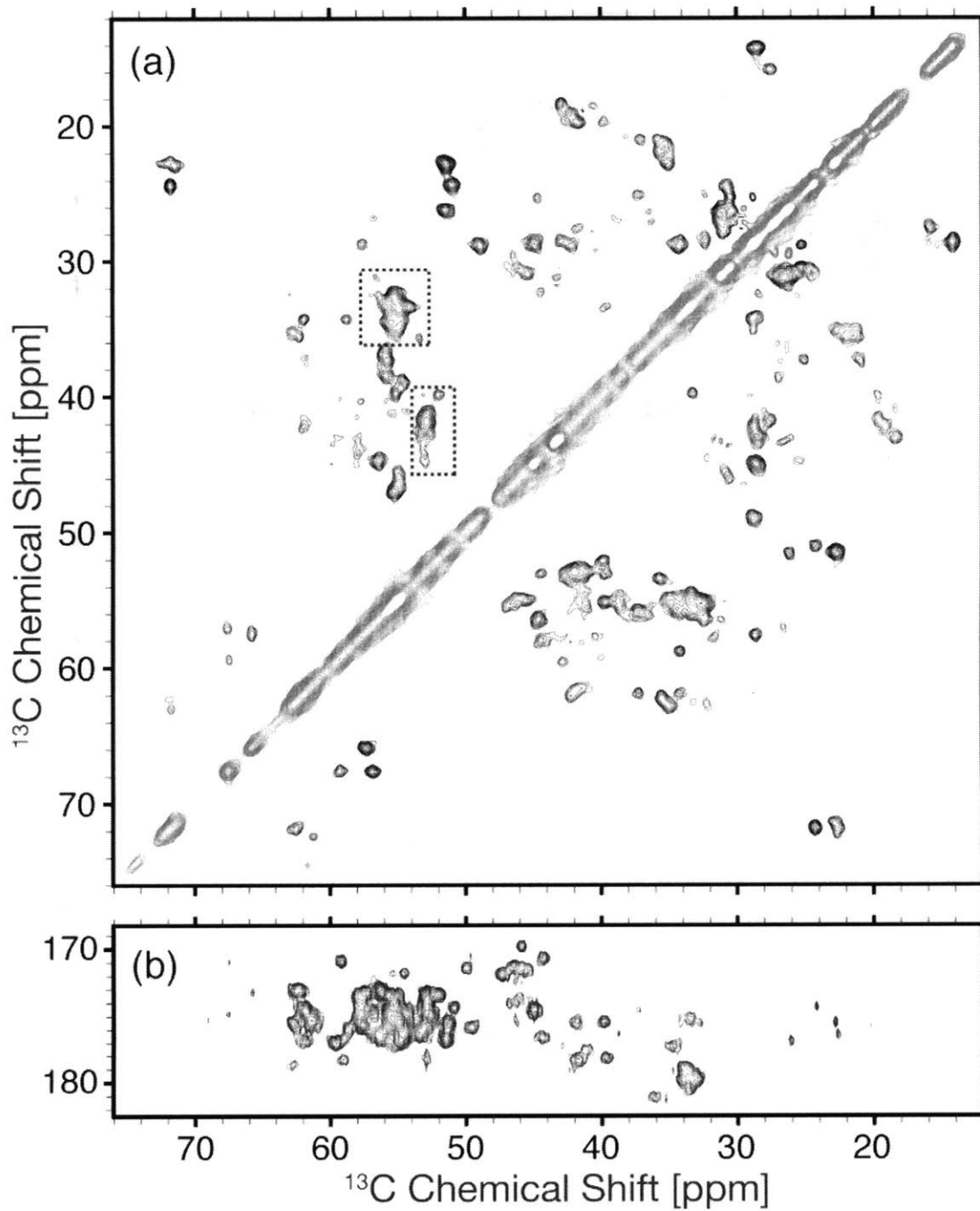
Figure 2. 2D ^{13}C-^{13}C correlation spectra of U-PI3-SH3 (750 MHz ^1H Larmor frequency, T = 2 °C). (a) Aliphatic region of a CMAR spectrum (2 ms mixing time, $\omega_r/2\pi$ = 28.571 kHz). Areas with high spectral overlap are indicated. (b) Carbonyl-aliphatic region of RFDR (1.76 ms mixing time, $\omega_r/2\pi$ = 18.182 kHz).
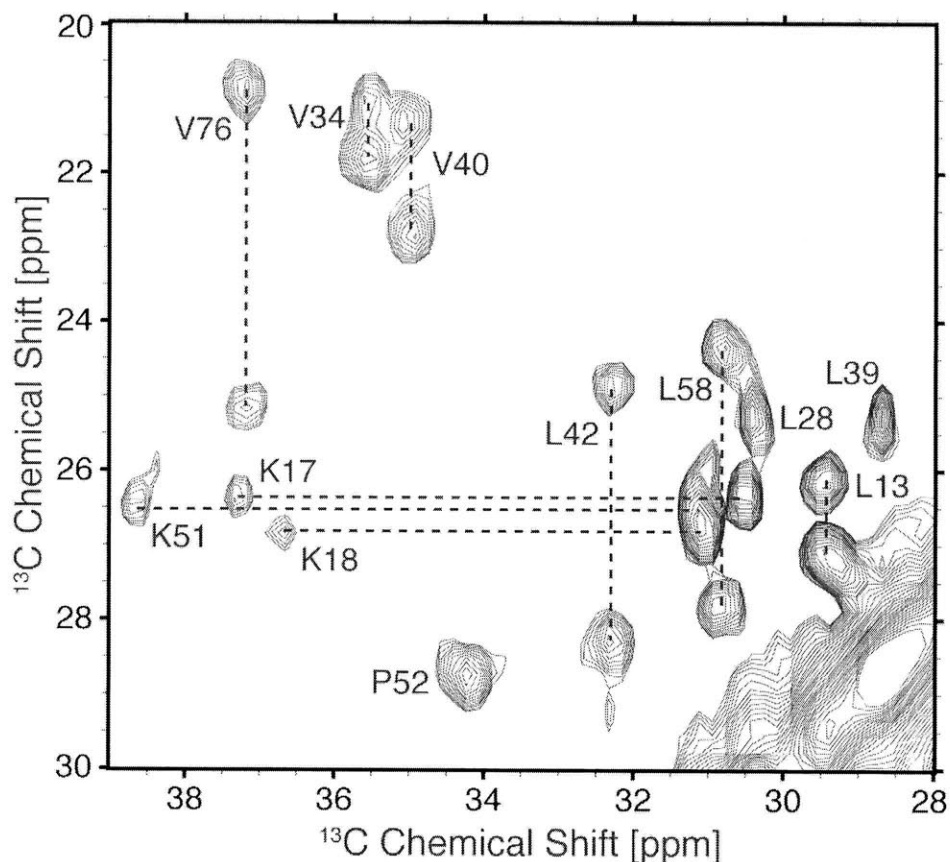
Figure 3. Section of a 2D ^{13}C-^{13}C RFDR (1.76 ms) spectrum of U-PI3-SH3 illustrating aliphatic side-chain correlations, recorded at 750 MHz ^{1}H Larmor frequency, T = 2 °C, and $\omega_r/2\pi$ = 18.182 kHz. Typical line widths are between 80 and 115 Hz (0.4 to 0.6 ppm).

MAS NMR spectra with a comparable resolution to those reported here have been observed for a number of amyloid systems, including the TTR fragment mentioned above (*31*) and the prion protein fragment HET-s (*36*). However, in other cases the observed linewidths for amyloid fibrils observed in MAS NMR spectra typically vary between 1 and 5 ppm (*33-35, 65*). The excellent quality of the dipolar correlation spectra of U-PI3-SH3 is therefore noteworthy, and can be attributed at least in part to careful control of experimental conditions, but is also indicative of the inherent homogeneity of PI3-SH3 fibrils at the molecular level.

Despite the narrow linewidths characteristic of PI3-SH3 fibrils, certain regions of the correlation spectra show a significant degree of resonance overlap consistent with localized structural degeneracy. For example, aliphatic ^{13}C-^{13}C correlation spectra exhibit extensive overlap of $^{13}C\alpha$-$^{13}C\beta$ cross-peaks at chemical shift positions typical of Glu/Gln and Asp/Asn

residues (marked with boxes in Figure 2). A low degree of chemical shift dispersion is a consequence of limited variability in secondary structure, and thus is likely to be a feature of the spectra of amyloid fibrils as a result of their high β-sheet content. However, it could also be a consequence of a small degree of local disorder.

7.3.2 Site-specific Resonance Assignments of PI3-SH3 Amyloid Fibrils

Using a combination of 2D homonuclear dipolar correlation experiments with short mixing periods (Figures 2, 3, and S3), it was possible to identify 82 ^{13}C spin-systems out of a total of 86 residues, and for the majority of spin systems all the side-chain resonances are observed. Sequential connectivities were established with heteronuclear NCACX and NCOCX experiments (51) and with homonuclear ^{13}C-^{13}C correlation experiments under weak coupling conditions (66) using the U-PI3-SH3 sample. In order to complement and expedite spectral analysis, we exploited the improved resolution of the 2-PI3-SH3 sample relative to the uniformly labeled one, as well its characteristic attenuation of dipolar truncation effects (67). Spectra of 2-PI3-SH3 typically exhibit linewidths below 0.2 ppm (~38 Hz, at a ^{1}H frequency of 750 MHz) for positions without an adjacent ^{13}C label, and linewidths below 0.1 ppm (~19 Hz, at a ^{1}H frequency of 750 MHz) for some side-chain resonances. Sequential inter-residue correlations in 2-PI3-SH3 were established using band-selective radio-frequency dipolar recoupling (BASE RFDR) experiments optimized for aliphatic ^{13}C nuclei and RFDR for broadband carbonyl-aliphatic recoupling. BASE RFDR generates highly sensitive sequential correlations between aliphatic nuclei, as described recently (50). In particular, a large number of ^{13}Cα(i)-^{13}Cα(i±1) and ^{13}Cα(i)-^{13}Cβ(i±1) correlations appear readily in BASE RFDR spectra of 2-PI3-SH3, which greatly facilitates resonance assignment. A broadband RFDR experiment, optimized for two-bond ^{13}C'(i)-^{13}Cα(i+1) pairs, can also yield highly informative sequential cross-peaks. Figure 4 compares the resonance assignment information available in broadband RFDR experiments on U-PI3-SH3 and 2-PI3-SH3. Both spectra were recorded under similar conditions (17.6 T and $\omega_r/2\pi$ = 28.5 kHz) using RFDR mixing without ^{1}H decoupling and optimized for one-bond transfer (U-PI3-SH3) and two-bond transfer (2-PI3-SH3). Generation and identification of two-bond ^{13}C'(i)-^{13}Cα(i+1) cross-peaks in U-PI3-SH3 at longer mixing times are precluded by dipolar truncation and by overlap with one-bond intra-residue ^{13}C'(i)-^{13}Cα(i) cross-peaks, respectively. The latter are few in 2-PI3-SH3, and thus numerous sequential correlations may be

identified, provided that both ^{13}C sites are labeled simultaneously by this alternating labeling scheme, in at least a fraction of the protein sample. PDSD experiments with 200-300 ms mixing times, although less efficient, provide additional corroboration of homonuclear sequential connectivities obtained with BASE RFDR and two-bond RFDR. Heteronuclear sequential correlations were obtained with TEDOR experiments on 2-PI3-SH3, optimized for either one-bond or two-bond mixing. The latter provided numerous $^{15}N(i)$-$^{13}C(i-1)$ connectivities that aided the assignment process and corroborated NCXCY correlations from U-PI3-SH3. Figures S4 and S5 in the supporting information show examples of these sequential correlation spectra in 2-PI3-SH3. A detailed description of this resonance assignment scheme and analysis of 2D sequential correlations in samples labeled with [2-^{13}C]glycerol will be the subject of a forthcoming publication.

Site-specific resonance assignments were elucidated for 75 residues, located in two large stretches (G1-K18 and L25-G80) of the sequence, thus leaving two small segments of the sequence unassigned, E19-I24 and the C-terminus, R81-S85 with the exception of P86, assigned unambiguously by elimination. Figure 5a shows the amino acid sequence of PI3-SH3, highlighting the residues that have been assigned at least in part. ^{13}C and ^{15}N chemical shifts for all assigned nuclei are provided in Table 1 of the supplementary material.

For several spin-systems identified in ^{13}C-^{13}C dipolar correlation spectra of U-PI3-SH3, it was not possible to establish inter-residue connections in either U-PI3-SH3 or 2-PI3-SH3 samples due to severe overlap of backbone resonances. In some cases, spin systems were ambiguously identified through their characteristic ^{13}C-^{13}C correlations but did not yield sequential connectivities with sufficient certainty to establish unambiguous assignments.
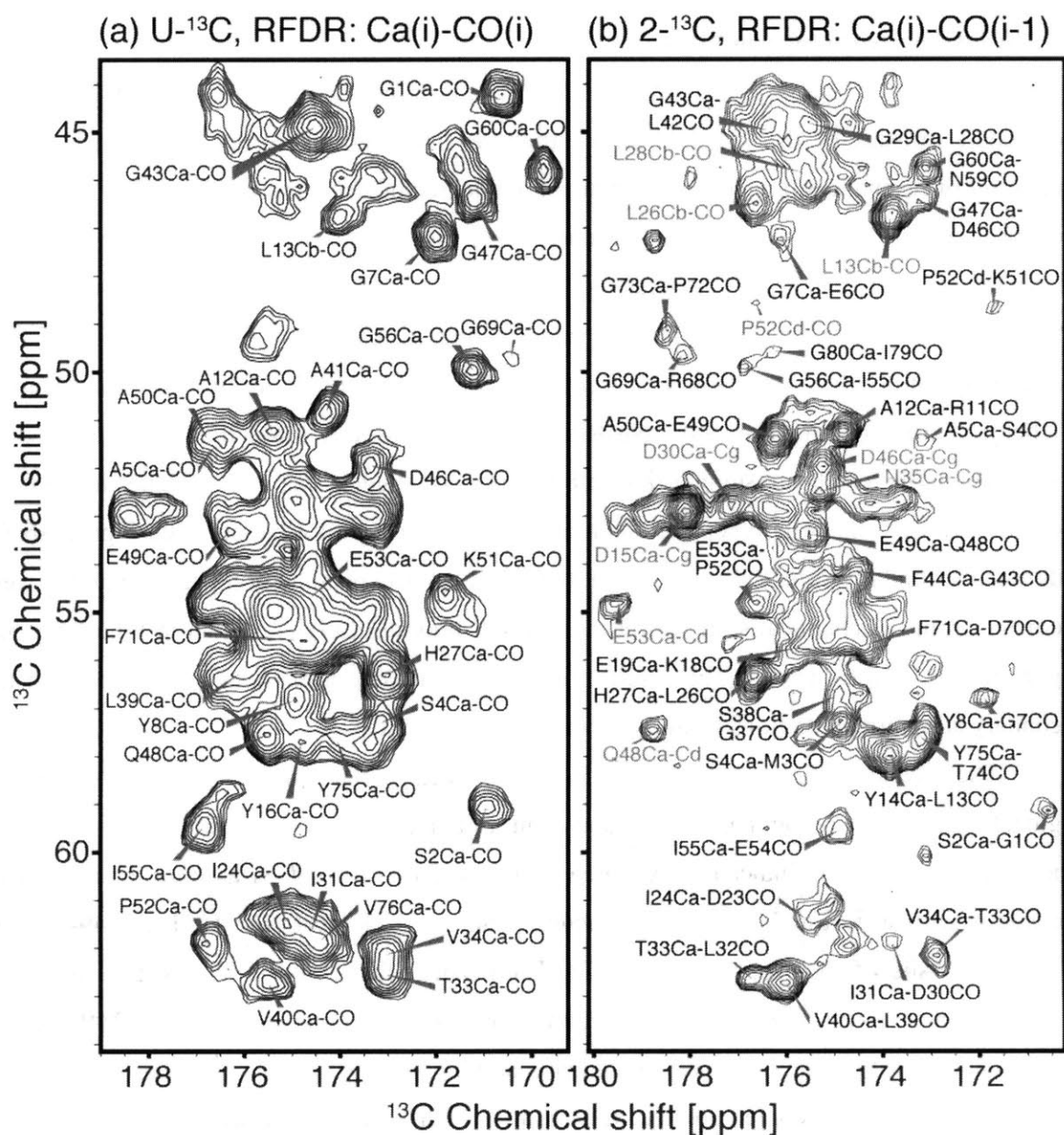
Figure 4. Comparison of cross-peak information attainable in broadband Cα-C' correlation spectra of U-PI3-SH3 and 2-PI3-SH3. (a) Cα-C' region of a U-PI3-SH3 spectrum recorded with RFDR 2.24 ms mixing, displaying intra-residue one-bond correlations, some of which are labeled. (b) Similar region of a 2-PI3-SH3 spectrum recorded with RFDR 6.72 ms mixing, presently multiple inter-residue two-bond correlations, labeled in black, and two-bond intra-residue correlations, labeled in gray. Both spectra were recorded at $\omega_r/2\pi = 28.571$ kHz employing $\omega_1/2\pi = 120$ kHz ^{13}C π pulses and without ^{1}H decoupling during the mixing period.

7.3.3 Secondary Structure and Dynamics

Site-specific analysis of the MAS NMR spectra allows us to elucidate both global and local characteristics of the PI3-SH3 fibril structure. The well-resolved spectra, presenting a unique set of resonances, are consistent with a high degree of structural homogeneity throughout the fibrils, with every residue in the sequence being in a single environment within the fibril. Secondary structure elements were predicted from the difference between the observed ^{13}C and ^{15}N chemical shifts and their random coil values, using the Chemical Shift Index (*57, 58, 68*). Here, the quantity $\Delta\delta = \delta^{13}C\alpha - \delta^{13}C\beta$, from Ref. (*68*), is particularly useful, since the accuracy of chemical shifts obtained by MAS NMR experiments is usually lower than in solution NMR experiments. The results are summarized in Figure 5b and suggest that most residues in the well-ordered segments of the polypeptide chain are in a β-strand conformation. (See Figure S6 for individual secondary shifts.) To supplement these results, backbone torsion angles ϕ and ψ were predicted from the chemical shifts using the TALOS program (*59*), as shown in Figure 5c. See Figure S7 for further details. The results from this database approach agree closely with the secondary structure predictions from the Chemical Shift Index.

The secondary structure of PI3-SH3 in amyloid fibril state consists of four well-defined segments adopting predominantly β-strand conformations, linked by random coil elements. Chemical shift analysis indicates that three long stretches of residues assume torsion angles consistent with a β-strand secondary structure, namely S2-K17, L26-A41, and E49-T64, in addition to a short fourth segment, T74-Y78. The residues that connect the first and second β-strand regions (E19-I24) present weak backbone and side-chain signals in MAS NMR spectra, and thus most of them have not been assigned unambiguously. However, based on such spectral characteristics, we can conjecture that this small region of the sequence is not well defined structurally, presenting either dynamic disorder or conformational heterogeneity. The same conclusions can be reached regarding the last six residues of the sequence. On the other hand, the sequence of residues (L42-Q48) between the second and third β-strand regions appears to be a well-structured loop, based on their intense and narrow MAS NMR signals with random-coil chemical shifts. Another region of random-coil chemicals shifts (T65-G73) is observed between the third and fourth β-strand elements. In summary, the majority of residues giving rise to

narrow lines in the MAS NMR spectra of PI3-SH3 fibrils appear to be located in β-strands or in relatively short and ordered regions that link these secondary structure elements.
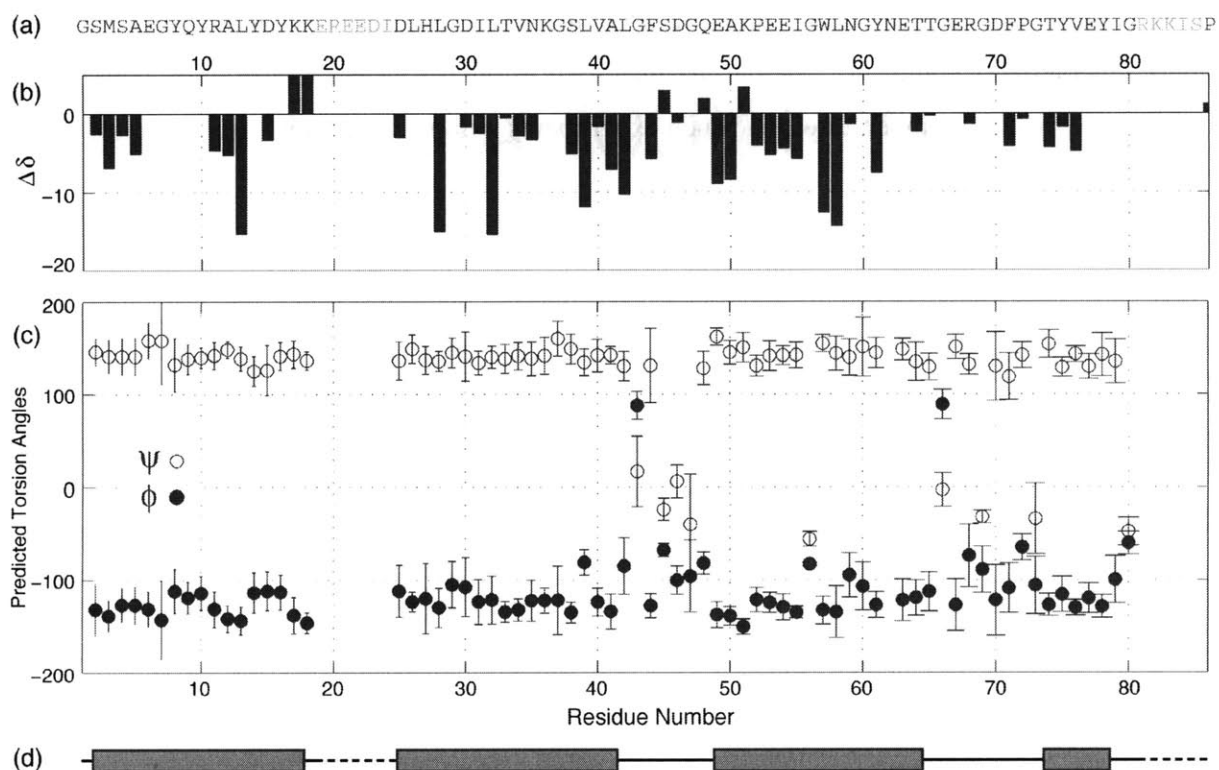
Figure 5: Chemical shift analysis. (a) PI3-SH3 sequence, displaying unambiguously assigned (black) and unassigned (gray) residues. (b) Secondary structure propensity derived from the Chemical Shift Index, where negative numbers indicate a β-sheet conformation. (c) Backbone torsion angles (φ and ψ) predicted from ^{13}C and ^{15}N chemical shift analysis using TALOS. (d) Secondary structure diagram of the PI3-SH3 polypeptide chain in amyloid fibril form. Gray bars indicate regions of high β-sheet content, while a single line indicates a random coil conformation. Dash lines mark the positions of dynamic/disordered residues.

From the absence of any signals detectable by solution-like NMR methods we can conclude that there are no highly flexible regions in the protein backbone of PI3-SH3 in its fibrillar form. In addition to indicating an overall high level of structural homogeneity, the observation of narrow linewidths for the majority of the PI3-SH3 residues when in the fibril state points to the absence of significant dynamic processes on the timescale of the rf irradiation

(decoupling and cross polarization) for these residues, which can lead to line broadening and loss of intensity in MAS NMR spectra (*69, 70*). On the other hand, no resonances could be identified for a few residues within the unassigned segments mentioned above, suggesting that such residues could be in regions of the fibril presenting conformational heterogeneity or undergoing local motions on timescales that interfere with the MAS NMR experiments. Such dynamic interference effects could, in principle, be identified through variable temperature experiments over a wide range of temperatures, and such experiments will be reported in future studies.

7.4 Discussion

7.4.1 Comparison with the Native Fold

The native structure of PI3-SH3 is composed of five β-strands forming a β-barrel and a short helix-like turn (*14-16*). Fiber-diffraction studies suggest that the orientation of the β-strands in PI3-SH3 fibrils is perpendicular to the fibril axis (*13, 71*), a finding inconsistent with the β-barrel architecture of the native fold. Nevertheless, a simple rearrangement of the β-strands in the native PI3-SH3 fold is sufficient to account for electron-density maps from cyro-EM and thus satisfy the above observation. A comparison of the secondary structure elements of the native PI3-SH3 fold, previously obtained with solution NMR and X-ray crystallography, to those derived from our solid-state NMR chemical shift analysis is depicted in Figure 5d, in which bars and cylinders denote β-sheet and helical segments, respectively, and dashed lines highlight flexible or disordered regions. These plots show that the secondary structure elements in the native fold are significantly different from those elements identified for the protein in the fibril state. The MAS NMR data show that none of the assigned residues adopt α-helical or helix-like conformations in the fibrils, a finding that is in accordance with circular dichroism and FT-IR data (*19*). More remarkably, β-strand segments are considerably longer in the fibrils than in the native fold.

The secondary structure elements in the native structure of PI3-SH3, as determined by solution NMR spectroscopy (*15, 16*), consist of five β-strands, which are four to six residues in length, and short helical segments. In the native PI3-SH3 fold, the first six N-terminal residues are disordered and the first β-strand is formed by Q9-A12, which is followed by a long random coil segment that includes the RT loop (L13-D25) and the diverging turn (L26-D30). The second β-strand consists of residues I31-V34 and is followed by a flexible stretch of ~22 residues in random coil conformation (referred to as the N-src loop), including a short helical segment and a solvent-exposed helical turn (K51-E54). A well defined type I β turn (E63-G66) links the third and fourth β-strands, composed of residues W57-N62 and E67-P72, respectively. The fifth β-strand, formed by residues E77-R81A, comes after a short helix-like turn, and is followed by a disordered C-terminus (K82-P86). A very similar native conformation is found in the crystal structure of PI3-SH3, with the major difference being that, as a result of intermolecular packing, the N-src loop and the C-terminus are less flexible and better defined in the crystal structure (*14*). In turn, these intermolecular interactions in the crystal lattice stabilize the secondary structure

elements of native PI3-SH3, whose β-strands are longer by a few residues in the crystal structure than in the solution structure, as shown in Figure 5(d).

Considering the secondary structure elements in the fibrillar state described in the preceding section, we may highlight the following differences (and a few similarities) between PI3-SH3 fibril subunits and the native fold. Three out of the five native β-strands are preserved in the fibril state, as part of longer β-strand elements, while residues that form part of β-strands 4 and 5 in the native state (near the C-terminus) adopt mostly a random-coil conformation in the fibril. The natively disordered N-terminus adopts a rigid β-strand structure in the fibril state, while the native structure's diverging turn goes from a well-defined loop conformation to a β-strand configuration. Similar changes occur for the natively flexible N-src loop region, which encompasses both helical and random-coil residues in the native fold but constitutes the middle of a long, rigid strand-loop-strand region in the fibril state, as shown in Figure 5(d). On the other hand, the C-terminal residues, highly flexible in the solution structure, appear to also be dynamically/structurally disordered in the fibrils. Finally, the natively random-coil RT-loop constitutes, in the fibril state, both part of a rigid β-strand region (on the N-terminal side) and a disordered segment with few observed MAS NMR signals.

Our observations therefore demonstrate that PI3-SH3 subunits in amyloid fibrils adopt a conformation that is vastly different from that of the native structure of the protein. While fibril measurements by X-ray and cryo-EM (20) can be rationalized with changes in the relative location of secondary structure elements, our data indicate that, in the case of PI3-SH3, only a few residues preserve their conformations between the native and the fibril state, while the majority of the sequence presents important differences in secondary structure and/or dynamics between one state and the other.
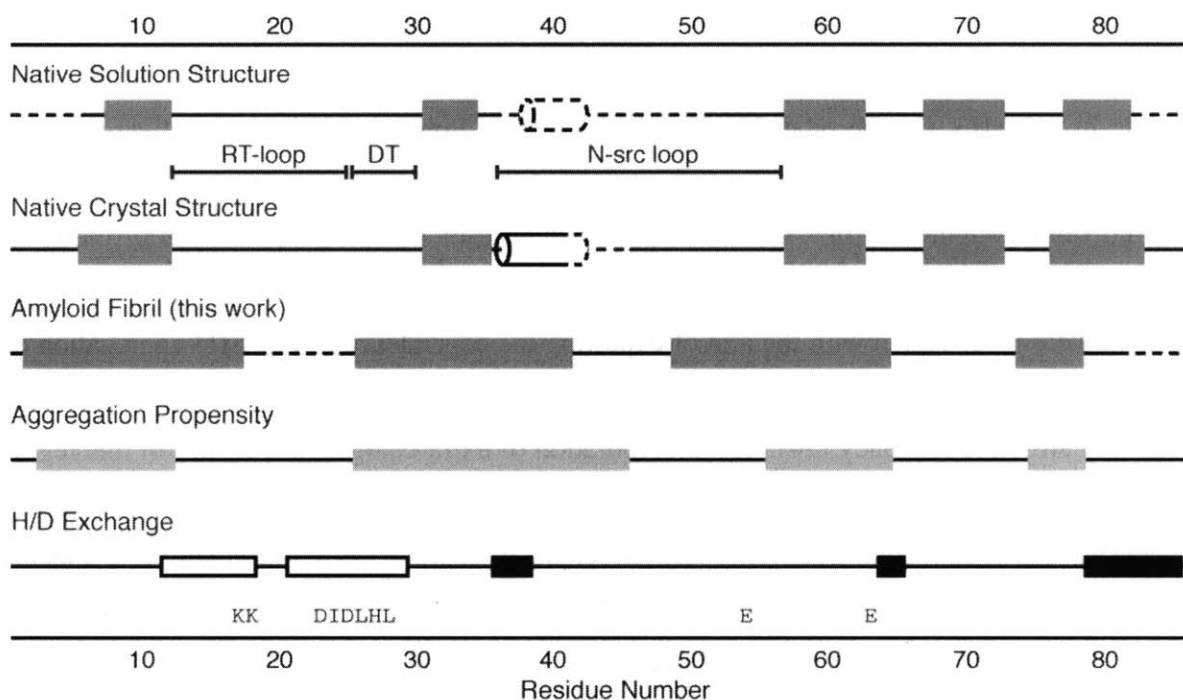
Figure 6: Secondary structure of PI3-SH3 amyloid fibrils in the context of previous structural and mechanistic studies. From top to bottom: Secondary structure of natively folded PI3-SH3 in solution [Refs. 15 and 16] and in crystalline form [Ref. 14], with flexible regions shown with dashed lines, and the RT-loop, DT, and N-src loop, discussed in the text, are marked; secondary structure of PI3-SH3 in fibril form (from Figure 5d); aggregation propensity of the PI3-SH3 amino acid sequence [Ref. 75]; H/D exchange results showing highly protected regions (empty rectangles) and unprotected sites (filled rectangles). A few specific residues discussed in the text are also shown.

7.4.2 Aggregation Propensity and Fibril Conformation

PI3-SH3 has been thoroughly studied as a model for the characterization of amyloid fibril formation by a natively globular protein, starting from the acid-unfolded state. In particular, multiple studies have shed light on this protein's propensity to form fibrils as a function of its amino acid sequence and variants thereof, which has led to the identification of key positions in the sequence in which amino acid substitutions may either disrupt or accelerate the aggregation of monomers into amyloid fibrils, as well as segments that have little effect on fibril formation. For example, preparation of chimeras between PI3-SH3 and spectrin-SH3 (SPC-SH3), a structurally homologous protein that does not form fibrils in all conditions tested, showed that

211

the N-src loop does not have an effect on the aggregation propensity of either protein (72), even though it is the most dissimilar region between the two native structures, being much longer and flexible in PI3-SH3. On the other hand, a short segment from the RT-loop and diverging turn regions of PI3-SH3, D23-L28, was sufficient to confer a SPC-SH3 chimera the ability to form amyloid fibrils, suggesting that this part of the sequence of PI3-SH3 has a specific role in its aggregation ability (73). Additional sets of mutations containing at least one residue in this segment were investigated with the aim of exploring the dependence of PI3-SH3 aggregation on three key physico-chemical characteristics of polypeptides, charge, hydrophobicity, and secondary structure tendency (74). The incorporation of extra charges (at pH 2.0) in this region effectively precluded the formation of aggregates, while increasing the polypeptide's helical propensity had a moderate but noticeable effect on disrupting fibril formation (73). Placing charges in regions of the sequence other than the RT-loop/DT segment can also have dramatic effects. As reported recently (75), E54K and E63K single-mutants have elongation rates several orders of magnitude slower than wild-type PI3-SH3, starting with pre-formed fibrils. These two sites, as well as the RT-loop/DT segment, are located near residue stretches that are found to be prone to aggregation, as calculated by the Zyggregator algorithm (76, 77), to which their strong influence has been attributed. Moreover, mutation of a charged residue to a neutral one at pH 2.0 (K18Q) resulted in a two-fold increase in the rate of elongation of pre-formed fibrils (75). On the other hand, inserting multiple charges at the N-terminus (via a His-tag) did not interfere significantly with PI3-SH3 aggregation (73), further illustrating the sequence dependence of the factors influencing fibril formation by proteins.

Knowing the secondary structure that the PI3-SH3 polypeptide adopts in fibril form, we can discuss the previously published results described above in the context of both aggregation propensity of the amino acid sequence and the final fibril conformation. The RT-loop/DT segment D23-L28 is at the edge of a β-sheet in the fibril conformation that coincides well with a predicted aggregation-prone region in the unfolded state of PI3-SH3 at pH 2.0 (75), which may explain its fundamental role in fibril formation. Conversely, mutated residues near the aggregation-prone segments, such as L28K and D25R (73), may act as "gatekeepers" that prevent intermolecular association and formation of ordered β-sheets via electrostatic charges (78, 79). Similarly, residues E54 and E63 flank a predicted aggregation-prone region and, moreover, adopt a β-sheet conformation in the final fibril state, as part of a β-strand spanning

residues 49 to 64. The relative location of these residues within the fibril β-strand, with E63 being more peripheral, may account for the slower elongation rate of the E54K mutant compared to E63K fibrils (75), while both are orders of magnitude slower than the wild-type case. A different situation is observed for residues K17 and K18, which are distant from calculated aggregation-prone regions and are not effective gatekeepers, even though they end up being adjacent to a β-sheet segment in the fibril state. These observations support the hypothesis that gatekeeper residues can be found in close proximity to calculated aggregation-prone regions in native or unfolded states and are not restricted to positions within β-sheet segments in the mature fibril state. Indeed, several charged residues are incorporated in the highly ordered β-sheets of the fibril conformation of wild-type PI3-SH3 described in Figure 5. Therefore, the ability of charged residues (and additional charges) to disrupt the aggregation process is highly dependent on their position in the sequence, and not only on the conformation they adopt in the fibril state, which may reflect the specific role they play in the assembly mechanism.

7.4.3 On the Mechanism of PI3-SH3 Fibril Formation

It is interesting to compare our results to published work that addresses directly the mechanism of fibril formation by PI3-SH3. A recent study (80) has shown, via a pulse-labeling hydrogen-deuterium (H/D) exchange technique, that the segments A12-K18 and E21-G29 are the most protected from exchange in mature fibrils formed at pH 2.0, which we have analyzed here. A similar region, Y14-I24, is the most protected in the pre-fibrillar aggregates observed at pH 1.5, a condition found to favor the stabilization of PI3-SH3 intermediates, while other regions have dissimilar levels of exchange for the two species. Furthermore, as the aggregates mature into fibrils at pH 1.5, the H/D profile changes into one that similar to those of mature fibrils at pH 2.0, pointing to a common fibril assembly process in which pre-fibrillar intermediates are formed first and subsequently rearrange into the final fibril structure. This observation is consistent with the nucleated conformational conversion (NCC) mechanism (81), in which monomers in solution coalesce into amorphous oligomers that later undergo reorganizations that produce ordered oligomers and finally amyloid fibrils rich in β-sheets. Indeed, the formation of amyloidogenic oligomers in the process of fibril formation at pH 2.0 has been detected and quantified by a single-molecule fluorescence study (82). The partial protection from H/D exchange in PI3-SH3 intermediates indicates a significant degree of intermolecular organization,

which is preserved and extended in the mature fibrils to most residues in the stretch A12-G29. Notably, these protected residues include the mutation-sensitive RT-loop/DT segment D23-L28 as well as part of the N-terminal β-strand of the fibril state. However, while the latter segment (A12-K18) is well ordered according to our MAS NMR data, several residues within the protected segment present weak NMR signals consistent with local structural disorder, namely residues E19 to I24, as described above. Therefore, while residues 12 to 29 are part of a highly persistent structure in the oligomeric intermediates that drive the initial steps of fibril formation, and remain within the core of mature fibrils, our MAS NMR analysis shows that only some of them adopt a highly ordered β-sheet conformation in the final fibril state.

These observations provide further insight into the different functions that distinct segments of the PI3-SH3 sequence may assume during the course of aggregation and fibril elongation. Residues 23 to 28, DIDLHL, present a binary hydrophilic-hydrophobic pattern that has been shown to promote aggregation in various proteins (*83, 84*) and may be responsible for the coalescence of partially ordered oligomers starting from unfolded PI3-SH3 monomers. The few preceding residues (17 to 22, KKEREE), on the other hand, are hydrophilic and positively charged at pH 2.0, yet they do not interfere with aggregation and are partially protected from H/D exchange in both oligomers and fibrils. Finally, residues on both sides of the protected region form β-sheets in the fibril structure (β-strands 1 and 2 in Figure 5), including the segment Y14-K17, which is already protected in the oligomeric intermediates. While there are several aggregation-prone regions in the PI3-SH3 sequence, one in particular (near the DT) seems to be sufficient to cause aggregation and form oligomers and eventually fibrils, in a process consistent with the NCC mechanism. These initial interactions are then likely superseded by their reorganization and the formation of well-ordered β-sheets throughout the remainder of the sequence, as revealed by the MAS NMR measurements presented here.

7.4.4 Implications for a Structural Model

The MAS NMR analysis described above confirms that PI3SH3 fibrils contain extensive regions of β-structure, as do fibrils associated with degenerative or infectious diseases. In early studies, analysis of FT-IR (*19*) and cryo-EM data (*20*) have indicated that approximately ~40% of the PI3-SH3 protein sequence is contained within the fibril β-sheet core, while the remainder of the protein sequence was proposed to form connecting loops and turns within the fibril

structure (*19*). This picture is broadly consistent with our results that indicate that up to 55% of the PI3-SH3 peptide chain is likely to be in well structured, predominantly β-strand conformations, although the fraction that adopts the cross-β structure characteristic of amyloid fibrils could be somewhat lower.

Our site-specific secondary structure assignments for PI3-SH3 fibrils are in good agreement with previous H/D exchange experiments, which show that most of the amide hydrogen atoms in the fibril structure are significantly protected from solvent exchange (*23*). Although certain regions of the PI3-SH3 polypeptide chain in the fibril such as charged side-chain residues show evidence of limited dynamics, the absence of segments with liquid-like dynamics is consistent with a compact structure, and with the finding that none of the sequence is susceptible to ready degradation by proteolytic enzymes in the fibril state (*22*). Thus we can infer that, while the well-structured regions of PI3-SH3 form the core amyloid structure and rigid loops, the remaining segments of the peptide backbone are also packed tightly within the quaternary structure of the fibril.

These results, in combination with structural information for PI3-SH3 previously obtained by cryo-EM, allow us to propose a structural *model* for PI3-SH3 that describes a possible arrangement of the subunit strands within the fibril core. However, we wish to emphasize that this model requires validation, and needs to be refined with additional structural constraints, specifically, interatomic distances and torsion angles that are presently in progress. Nevertheless, we can suggest a possible arrangement of the protein subunits within the fibril architecture given the following facts. (1) Cryo-EM data show that the width of the protofilaments composing the fibrils is ~20 Å, sufficient to accommodate two β-sheets, while X-ray diffraction data suggest that the sheet spacing is 9.4 Å (*13, 20*). (2) No peak doubling is observed in our NMR spectra, indicating that all subunits must reside in identical environments, with the lowest symmetry axis for two sheets being a C_2 axis. (3) H/D exchange measurements show that the C-terminus has a slightly higher exchange rate compared to the rest of the sequence in the fibrils (*80*). (4) A tandem repeat of PI3-SH3 has been observed to form fibrils with a macroscopic morphology similar to those formed by the single-domain protein, but apparently composed of two laterally aligned protofilaments (*85*).
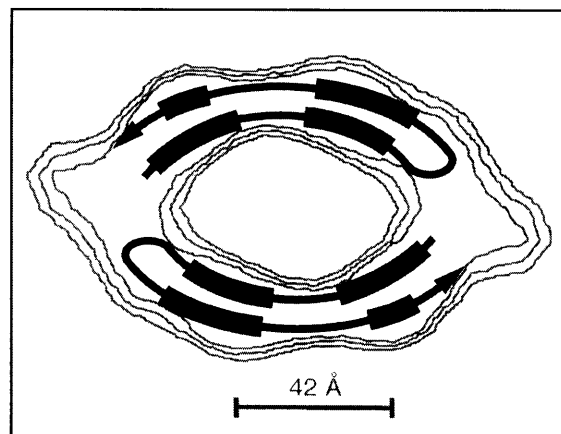
Figure 7: Possible model for PI3-SH3 amyloid fibril architecture. The polypeptide chains are shown as single lines, with the thick segments denoting β-sheets running into the plane of the page, and are superimposed onto the cryo-EM map (adapted from Ref. 20) so that each protein subunit occupies one half of the fibril cross section. An arrowhead indicates the C-terminus.

Analyzing the structural data currently available from models of amyloid fibrils (*32, 86, 87*) we estimated different averages for the length-per-residue of the β-strands, between ~3.2 Å/residue for straight β-strands and ~2.2 Å/residue for curved β-strands (see supporting information). For PI3-SH3 fibrils three sections of β-strands composed of ~16 residues are predicted by our chemical shift analysis. We note that these lengths are greater than most β-strands which typically contain up to ~10 residues in soluble proteins. With a lower and upper length limit of 2.2 to 3.2 Å/residue the individual β-strands would be between 35 and 51 Å long. However, since chemical shift analysis is only approximate, it would not be surprising if addition of structural constraints resulted in shorter β-strands in a calculated NMR structure. Cryo-EM data show a distance of ~42 Å between the regions of highest electron density in the fibril cross-section, which may correspond to the location of tightly packed β-sheets. This leads to the suggestion that in the case of PI3-SH3 fibrils the β-strands in the fibril core may be arranged in long arches across the fibril cross-section. These qualitative constraints then permit only a few possible arrangements, with one possibility illustrated in Figure 7. In this structural *model*, four β-sheets form the core of each individual protofilament, while lateral interactions between subunits in adjacent protofilaments bring them together to form the fibrils. Further experimental data will be necessary to test this hypothetical model.

216

7.5 Conclusions

The success of our MAS NMR experiments in achieving almost complete assignment of ^{13}C and ^{15}N resonances of PI3-SH3 in amyloid fibril state, along with the identification of the majority of backbone torsion angles, suggest that a complete solid-state NMR structure of PI3-SH3 amyloid fibrils is within reach. As an alternative route, combination of our initial results with further analysis of chemical shifts (*88*) and cryo-EM data may lead to a detailed molecular description of the structure of a natively stable protein after conversion into the generic amyloid state. But even prior to a full structure, the present data have provided key information about the local conformation adopted by PI3-SH3 in amyloid fibril form. Solid-state NMR measurements have revealed a high degree of molecular organization in the structure and shown how the secondary structure elements in the amyloid fibril differ from those in the native structure. Our observations in PI3-SH3 reinforce the idea that the main chain preferences for a given type of secondary structure in the native state are not sufficient to determine those of the fibrillar state, but instead, intermolecular and quaternary interactions must guide the conformation of a protein as it is incorporated in amyloid fibrils. The identification of the backbone conformation of PI3-SH3 in the fibril state has allowed us to interpret the results of previous mechanistic studies in terms of site-specific molecular structure and propose a model of protofilament assembly, thus contributing to the further understanding of the complex mechanism of fibril formation by a natively folded protein.

References

(1)   Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem. 75*, 333-66.

(2)   Sipe, J. D. (1992) Amyloidosis. *Annu. Rev. Biochem. 61*, 947-75.

(3)   Sunde, M., and Blake, C. C. (1998) From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Q. Rev. Biophys. 31*, 1-39.

(4)   Chiti, F., and Dobson, C. M. (2009) Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol. 5*, 15-22.

(5)   Sacchettini, J. C., and Kelly, J. W. (2002) Therapeutic Strategies for Human Amyloid Diseases. *Nat. Rev. Drug Discovery 1*, 267-275.

(6)   Dobson, C. M. (2003) Protein folding and misfolding. *Nature 426*, 884-90.

(7)   Si, K., Lindquist, S., and Kandel, E. R. (2003) A neuronal isoform of the aplysia CPEB has prion-like properties. *Cell 115*, 879-91.

(8)   Coustou, V., Deleu, C., Saupe, S., and Begueret, J. (1997) The protein product of the het-s heterokaryon incompatibility gene of the fungus Podospora anserina behaves as a prion analog. *Proc. Natl. Acad. Sci. U.S.A. 94*, 9773-8.

(9)   Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., Normark, S., and Hultgren, S. J. (2002) Role of Escherichia coli curli operons in directing amyloid fiber formation. *Science 295*, 851-5.

(10)   Berson, J. F., Theos, A. C., Harper, D. C., Tenza, D., Raposo, G., and Marks, M. S. (2003) Proprotein convertase cleavage liberates a fibrillogenic fragment of a resident glycoprotein to initiate melanosome biogenesis. *J. Cell Biol. 161*, 521-33.

(11)   Knowles, T. P., Fitzpatrick, A. W., Meehan, S., Mott, H. R., Vendruscolo, M., Dobson, C. M., and Welland, M. E. (2007) Role of intermolecular forces in defining material properties of protein nanofibrils. *Science 318*, 1900-3.

(12)   Uversky, V. N., and Fink, A. L. (2004) Conformational constraints for amyloid fibrillation: the importance of being unfolded. *Biochim. Biophys. Acta 1698*, 131-53.

(13) Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D., and Dobson, C. M. (1998) Amyloid Fibril Formation by an SH3 Domain. *Proc. Natl. Acad. Sci. U.S.A. 95*, 4224-4228.

(14) Liang, J., Chen, J. K., Schreiber, S. T., and Clardy, J. (1996) Crystal structure of PI3K SH3 domain at 2.0 angstroms resolution. *J. Mol. Biol. 257*, 632-43.

(15) Booker, G. W., Gout, I., Downing, A. K., Driscoll, P. C., Boyd, J., Waterfield, M. D., and Campbell, I. D. (1993) Solution structure and ligand-binding site of the SH3 domain of the p85 alpha subunit of phosphatidylinositol 3-kinase. *Cell 73*, 813-22.

(16) Koyama, S., Yu, H., Dalgarno, D. C., Shin, T. B., Zydowsky, L. D., and Schreiber, S. L. (1993) Structure of the PI3K SH3 domain and analysis of the SH3 family. *Cell 72*, 945-52.

(17) Koyama, S., Yu, H., Dalgarno, D., Shin, T., Zydowsky, L., and Schreiber, S. (1993) 1H and 15N assignments and secondary structure of the PI3K SH3 domain. *FEBS Lett 324*, 93-8.

(18) Ahn, H. C., Le, Y. T., Nagchowdhuri, P. S., Derose, E. F., Putnam-Evans, C., London, R. E., Markley, J. L., and Lim, K. H. (2006) NMR characterizations of an amyloidogenic conformational ensemble of the PI3K SH3 domain. *Protein Sci. 15*, 2552-7.

(19) Zurdo, J., Guijarro, J. I., and Dobson, C. M. (2001) Preparation and characterization of purified amyloid fibrils. *J. Am. Chem. Soc. 123*, 8141-2.

(20) Jimenez, J., Guijarro, J., Orlova, E., Zurdo, J., Dobson, C., Sunde, M., and Saibil, H. (1999) Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. *EMBO J. 18*, 815-21.

(21) Zurdo, J., Guijarro, J. I., Jimenez, J. L., Saibil, H. R., and Dobson, C. M. (2001) Dependence on solution conditions of aggregation and amyloid formation by an SH3 domain. *J. Mol. Biol. 311*, 325-40.

(22) Polverino de Laureto, P., Taddei, N., Frare, E., Capanni, C., Costantini, S., Zurdo, J., Chiti, F., Dobson, C. M., and Fontana, A. (2003) Protein aggregation and amyloid fibril formation by an SH3 domain probed by limited proteolysis. *J. Mol. Biol. 334*, 129-41.

(23)   Carulla, N., Caddy, G. L., Hall, D. R., Zurdo, J., Gairi, M., Feliz, M., Giralt, E., Robinson, C. V., and Dobson, C. M. (2005) Molecular recycling within amyloid fibrils. *Nature 436*, 554-8.

(24)   Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol. 273*, 729-39.

(25)   Serpell, L. C., Sunde, M., Benson, M. D., Tennent, G. A., Pepys, M. B., and Fraser, P. E. (2000) The protofilament substructure of amyloid fibrils. *J. Mol. Biol. 300*, 1033-9.

(26)   Harper, J. D., Lieber, C. M., and Lansbury Jr, P. T. (1997) Atomic force microscopic imaging of seeded fibril formation and fibril branching by the Alzheimer's disease amyloid-beta protein. *Chem. Biol. 4*, 951-9.

(27)   Rienstra, C. M., Tucker-Kellogg, L., Jaroniec, C. P., Hohwy, M., Reif, B., McMahon, M. T., Tidor, B., Lozano-Perez, T., and Griffin, R. G. (2002) De novo determination of peptide structure with solid-state magic-angle spinning NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A. 99*, 10260-5.

(28)   Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K., and Oschkinat, H. (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature 420*, 98-102.

(29)   Zech, S. G., Wand, A. J., and McDermott, A. E. (2005) Protein structure determination by high-resolution solid-state NMR spectroscopy: application to microcrystalline ubiquitin. *J. Am. Chem. Soc. 127*, 8618-26.

(30)   Loquet, A., Bardiaux, B., Gardiennet, C., Blanchet, C., Baldus, M., Nilges, M., Malliavin, T., and Bockmann, A. (2008) 3D Structure Determination of the Crh Protein from Highly Ambiguous Solid-State NMR Restraints. *J. Am. Chem. Soc. 130*, 3579-3589.

(31)   Jaroniec, C. P., MacPhee, C. E., Astrof, N. S., Dobson, C. M., and Griffin, R. G. (2002) Molecular conformation of a peptide fragment of transthyretin in an amyloid fibril. *Proc. Natl. Acad. Sci. U.S.A. 99*, 16748-53.

(32)   Jaroniec, C. P., MacPhee, C. E., Bajaj, V. S., McMahon, M. T., Dobson, C. M., and Griffin, R. G. (2004) High-resolution molecular structure of a peptide in an amyloid fibril

determined by magic angle spinning NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* *101*, 711-716.

(33) Heise, H., Hoyer, W., Becker, S., Andronesi, O. C., Riedel, D., and Baldus, M. (2005) Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A. 102*, 15871-6.

(34) Balbach, J. J., Ishii, Y., Antzutkin, O. N., Leapman, R. D., Rizzo, N. W., Dyda, F., Reed, J., and Tycko, R. (2000) Amyloid fibril formation by A beta 16-22, a seven-residue fragment of the Alzheimer's beta-amyloid peptide, and structural characterization by solid state NMR. *Biochemistry 39*, 13748-59.

(35) van der Wel, P. C. A., Lewandowski, J. R., and Griffin, R. G. (2007) Solid-State NMR Study of Amyloid Nanocrystals and Fibrils Formed by the Peptide GNNQQNY from Yeast Prion Protein Sup35p. *J. Am. Chem. Soc. 129*, 5117-30.

(36) Siemer, A. B., Ritter, C., Ernst, M., Riek, R., and Meier, B. H. (2005) High-resolution solid-state NMR spectroscopy of the prion protein HET-s in its amyloid conformation. *Angew. Chem. Int. Ed. 44*, 2441-4.

(37) Chan, J. C., Oyler, N. A., Yau, W. M., and Tycko, R. (2005) Parallel beta-sheets and polar zippers in amyloid fibrils formed by residues 10-39 of the yeast prion protein Ure2p. *Biochemistry 44*, 10669-80.

(38) Andronesi, O. C., Becker, S., Seidel, K., Heise, H., Young, H. S., and Baldus, M. (2005) Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy. *J. Am. Chem. Soc. 127*, 12965-74.

(39) Siemer, A. B., Arnold, A. A., Ritter, C., Westfeld, T., Ernst, M., Riek, R., and Meier, B. H. (2006) Observation of highly flexible residues in amyloid fibrils of the HET-s prion. *J. Am. Chem. Soc. 128*, 13224-8.

(40) Tycko, R. (2006) Molecular structure of amyloid fibrils: insights from solid-state NMR. *Q. Rev. Biophys. 39*, 1-55.

(41) LeMaster, D. M., and Kushlan, D. M. (1996) Dynamical mapping of E. coli thioredoxin via 13C NMR relaxation analysis. *J. Am. Chem. Soc. 118*, 9255-9264.

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

(43) Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V., and Griffin, R. G. (1995) Heteronuclear decoupling in rotating solids. *J. Chem. Phys. 103*, 6951-6958.

(44) Bennett, A. E., Griffin, R. G., Ok, J. H., and Vega, S. (1992) Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys. 96*, 8624-8627.

(45) Bennett, A. E., Rienstra, C. M., Griffiths, J. M., Zhen, W., Lansbury Jr, P. T., and Griffin, R. G. (1998) Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys. 108*, 9463-9479.

(46) Bayro, M. J., Ramachandran, R., Caporini, M. A., Eddy, M. T., and Griffin, R. G. (2008) Radio frequency-driven recoupling at high magic-angle spinning frequencies: Homonuclear recoupling sans heteronuclear decoupling. *J. Chem. Phys. 128*, 052321.

(47) Verel, R., Ernst, M., and Meier, B. H. (2001) Adiabatic dipolar recoupling in solid-state NMR: the DREAM scheme. *J. Magn. Reson. 150*, 81-99.

(48) De Paëpe, G., Bayro, M. J., Lewandowski, J., and Griffin, R. G. (2006) Broadband homonuclear correlation spectroscopy at high magnetic fields and MAS frequencies. *J. Am. Chem. Soc. 128*, 1776-7.

(49) Szeverenyi, N. M., Sullivan, M. J., and Maciel, G. E. (1982) Observation of spin exchange by two-dimensional fourier transform 13C cross polarization-magic-angle spinning. *J. Magn. Reson. 47*, 462-475.

(50) Bayro, M. J., Maly, T., Birkett, N. R., Dobson, C. M., and Griffin, R. G. (2009) Long-Range Correlations between Aliphatic 13C Nuclei in Protein MAS NMR Spectroscopy. *Angew. Chem. Int. Ed. 48*, 5708-5710.

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

(52) Michal, C. A., and Jelinski, L. W. (1997) REDOR 3D: Heteronuclear distance measurements in uniformly labeled and natural abundance solids. *J. Am. Chem. Soc. 119*, 9059-9060.

(53)   Jaroniec, C., Filip, C., and Griffin, R. (2002) 3D TEDOR NMR Experiments for the Simultaneous Measurement of Multiple Carbon-Nitrogen Distances in Uniformly 13C,15N-Labeled Solids. *J. Am. Chem. Soc. 124*, 10728-10742.

(54)   Morcombe, C. R., and Zilm, K. W. (2003) Chemical shift referencing in MAS solid state NMR. *J. Magn. Reson. 162*, 479-86.

(55)   Harris, R. K., Becker, E. D., Cabral de Menezes, S. M., Goodfellow, R., and Granger, P. (2001) NMR nomenclature. Nuclear spin properties and conventions for chemical shifts (IUPAC Recommendations 2001). *Pure Appl. Chem. 73*, 1795-1818.

(56)   Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277-93.

(57)   Wishart, D. S., and Sykes, B. D. (1994) The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J. Biomol. NMR 4*, 171-80.

(58)   Zhang, H., Neal, S., and Wishart, D. S. (2003) RefDB: a database of uniformly referenced protein chemical shifts. *J. Biomol. NMR 25*, 173-95.

(59)   Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR 13*, 289-302.

(60)   Helmus, J. J., Surewicz, K., Nadaud, P. S., Surewicz, W. K., and Jaroniec, C. P. (2008) Molecular conformation and dynamics of the Y145Stop variant of human prion protein in amyloid fibrils. *Proc. Natl. Acad. Sci. U.S.A. 105*, 6284-6289.

(61)   Igumenova, T. I., McDermott, A. E., Zilm, K. W., Martin, R. W., Paulson, E. K., and Wand, A. J. (2004) Assignments of carbon NMR resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc. 126*, 6720-7.

(62)   Franks, W. T., Zhou, D. H., Wylie, B. J., Money, B. G., Graesser, D. T., Frericks, H. L., Sahota, G., and Rienstra, C. M. (2005) Magic-angle spinning solid-state NMR spectroscopy of the beta1 immunoglobulin binding domain of protein G (GB1): 15N and 13C chemical shift assignments and conformational analysis. *J. Am. Chem. Soc. 127*, 12291-305.

(63) Bockmann, A., Lange, A., Galinier, A., Luca, S., Giraud, N., Juy, M., Heise, H., Montserret, R., Penin, F., and Baldus, M. (2003) Solid state NMR sequential resonance assignments and conformational analysis of the 2x10.4 kDa dimeric form of the Bacillus subtilis protein Crh. *J. Biomol. NMR 27*, 323-39.

(64) Siemer, A. B., Ritter, C., Steinmetz, M. O., Ernst, M., Riek, R., and Meier, B. H. (2006) 13C, 15N resonance assignment of parts of the HET-s prion protein in its amyloid form. *J. Biomol. NMR 34*, 75-87.

(65) Petkova, A. T., Ishii, Y., Balbach, J. J., Antzutkin, O. N., Leapman, R. D., Delaglio, F., and Tycko, R. (2002) A structural model for Alzheimer's beta -amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci. U.S.A. 99*, 16742-7.

(66) Seidel, K., Lange, A., Becker, S., Hughes, C. E., Heise, H., and Baldus, M. (2004) Protein solid-state NMR resonance assignments from (C-13, C-13) correlation spectroscopy. *Phys. Chem. Chem. Phys. 6*, 5090-5093.

(67) Bayro, M. J., Huber, M., Ramachandran, R., Davenport, T. C., Meier, B. H., Ernst, M., and Griffin, R. G. (2009) Dipolar truncation in magic-angle spinning NMR recoupling experiments. *J. Chem. Phys. 130*, 114506.

(68) Luca, S., Filippov, D. V., van Boom, J. H., Oschkinat, H., de Groot, H. J., and Baldus, M. (2001) Secondary chemical shifts in immobilized peptides and proteins: a qualitative basis for structure refinement under magic angle spinning. *J. Biomol. NMR 20*, 325-31.

(69) Long, J. R., Sun, B. Q., Bowen, A., and Griffin, R. G. (1994) Molecular Dynamics and Magic Angle Spinning NMR. *J. Am. Chem. Soc. 116*, 11950-11956.

(70) Maus, D. C., Copie, V., Sun, B., Griffiths, J. M., Griffin, R. G., Luo, S., Schrock, R. R., Liu, A. H., Seidel, S. W., Davis, W. M., and Grohmann, A. (1996) A Solid-State NMR Study of Tungsten Methyl Group Dynamics in $[W(\eta^5-C_5Me_5)Me_4][PF_6]$. *J. Am. Chem. Soc. 118*, 5665-5671.

(71) Blake, C., and Serpell, L. (1996) Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous [beta]-sheet helix. *Structure 4*, 989-998.

(72) Ventura, S., Lacroix, E., and Serrano, L. (2002) Insights into the origin of the tendency of the PI3-SH3 domain to form amyloid fibrils. *J. Mol. Biol. 322*, 1147-58.

(73) Ventura, S., Zurdo, J., Narayanan, S., Parreño, M., Mangues, R., Reif, B., Chiti, F., Giannoni, E., Dobson, C. M., Aviles, F. X., and Serrano, L. (2004) Short amino acid stretches can mediate amyloid formation in globular proteins: the Src homology 3 SH3 case. *Proc. Natl. Acad. Sci. U.S.A. 101*, 7258-63.

(74) Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C. M. (2003) Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature 424*, 805-8.

(75) Buell, A. K., Tartaglia, G. G., Birkett, N. R., Waudby, C. A., Vendruscolo, M., Salvatella, X., Welland, M. E., Dobson, C. M., and Knowles, T. P. (2009) Position-dependent electrostatic protection against protein aggregation. *Chembiochem 10*, 1309-12.

(76) Tartaglia, G. G., Pawar, A. P., Campioni, S., Dobson, C. M., Chiti, F., and Vendruscolo, M. (2008) Prediction of aggregation-prone regions in structured proteins. *J. Mol. Biol. 380*, 425-36.

(77) Tartaglia, G. G., and Vendruscolo, M. (2008) The Zyggregator method for predicting protein aggregation propensities. *Chem. Soc. Rev. 37*, 1395-401.

(78) Otzen, D. E., and Oliveberg, M. (1999) Salt-induced detour through compact regions of the protein folding landscape. *Proc. Natl. Acad. Sci. U.S.A. 96*, 11746-11751.

(79) Matysiak, S., and Clementi, C. (2006) Minimalist protein model as a diagnostic tool for misfolding and aggregation. *J. Mol. Biol. 363*, 297-308.

(80) Carulla, N., Zhou, M., Arimon, M., Gairí, M., Giralt, E., Robinson, C. V., and Dobson, C. M. (2009) Experimental characterization of disordered and ordered aggregates populated during the process of amyloid fibril formation. *Proc. Natl. Acad. Sci. U.S.A. 106*, 7828-33.

(81) Serio, T. R., Cashikar, A. G., Kowal, A. S., Sawicki, G. J., Moslehi, J. J., Serpell, L., Arnsdorf, M. F., and Lindquist, S. L. (2000) Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science 289*, 1317-1321.

(82) Orte, A., Birkett, N. R., Clarke, R. W., Devlin, G. L., Dobson, C. M., and Klenerman, D. (2008) Direct characterization of amyloidogenic oligomers by single-molecule fluorescence. *Proc. Natl. Acad. Sci. U.S.A. 105*, 14424-9.

(83) West, M. W., Wang, W. X., Patterson, J., Mancias, J. D., Beasley, J. R., and Hecht, M. H. (1999) De novo amyloid proteins from designed combinatorial libraries. *Proc. Natl. Acad. Sci. U.S.A. 96*, 11211-11216.

(84) Broome, B. M., and Hecht, M. H. (2000) Nature disfavors sequences of alternating polar and non-polar amino acids: Implications for amyloidogenesis. *J. Mol. Biol. 296*, 961-968.

(85) Bader, R., Bamford, R., Zurdo, J., Luisi, B. F., and Dobson, C. M. (2006) Probing the mechanism of amyloidogenesis through a tandem repeat of the PI3-SH3 domain suggests a generic model for protein aggregation and fibril formation. *J. Mol. Biol. 356*, 189-208.

(86) Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A. B., Riek, R., and Meier, B. H. (2008) Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science 319*, 1523-6.

(87) Iwata, K., Fujiwara, T., Matsuki, Y., Akutsu, H., Takahashi, S., Naiki, H., and Goto, Y. (2006) 3D structure of amyloid protofilaments of beta-2-microglobulin fragment probed by solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A. 103*, 18119-18124.

(88) Robustelli, P., Cavalli, A., and Vendruscolo, M. (2008) Determination of protein structures in the solid state from NMR chemical shifts. *Structure 16*, 1764-9.

Supplementary Figures for Chapter 7
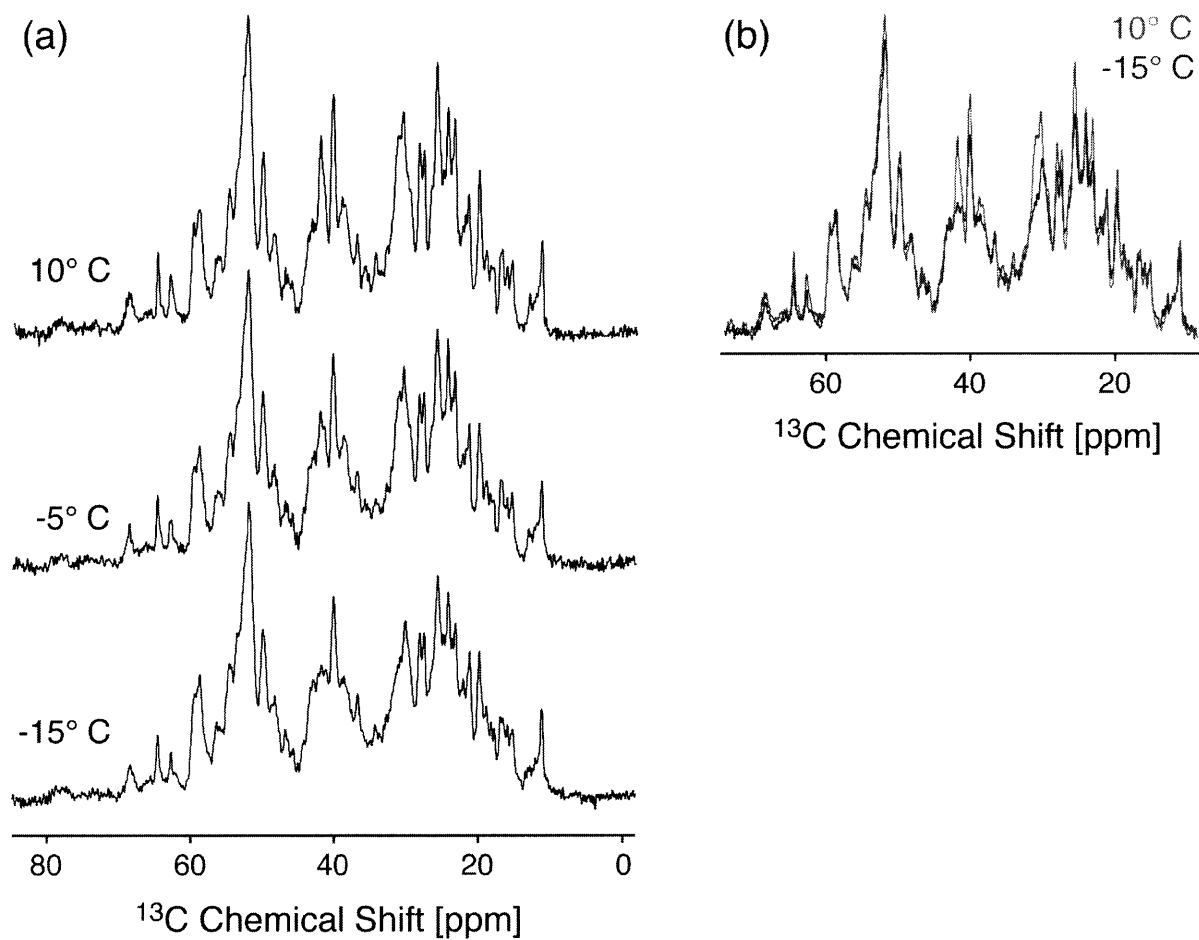
Figure S1. Temperature dependence of ^{13}C cross-polarization MAS NMR spectra of U-PI3-SH3.
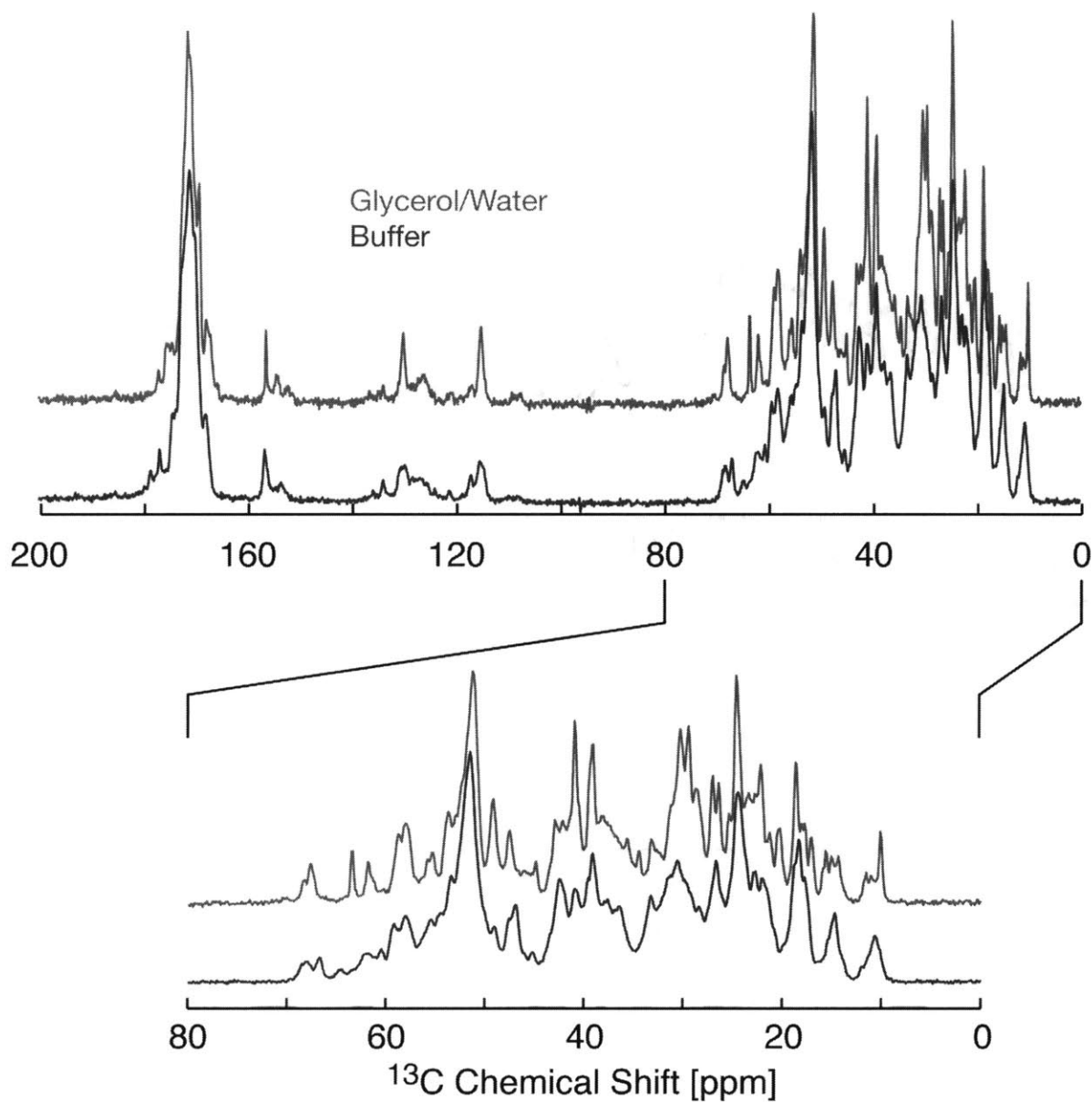
Figure S2. Comparison of the spectral quality of ^{13}C-CPMAS spectra of two different sample preparations. Red: U-PI3-SH3 fibrils prepared in buffer and put into a glycerol/water mixture. Blue: U-PI3-SH3 fibrils prepared in buffer and left in buffer solution. Both spectra are recorded under identical experimental conditions (900 MHz ^{1}H Larmor frequency)
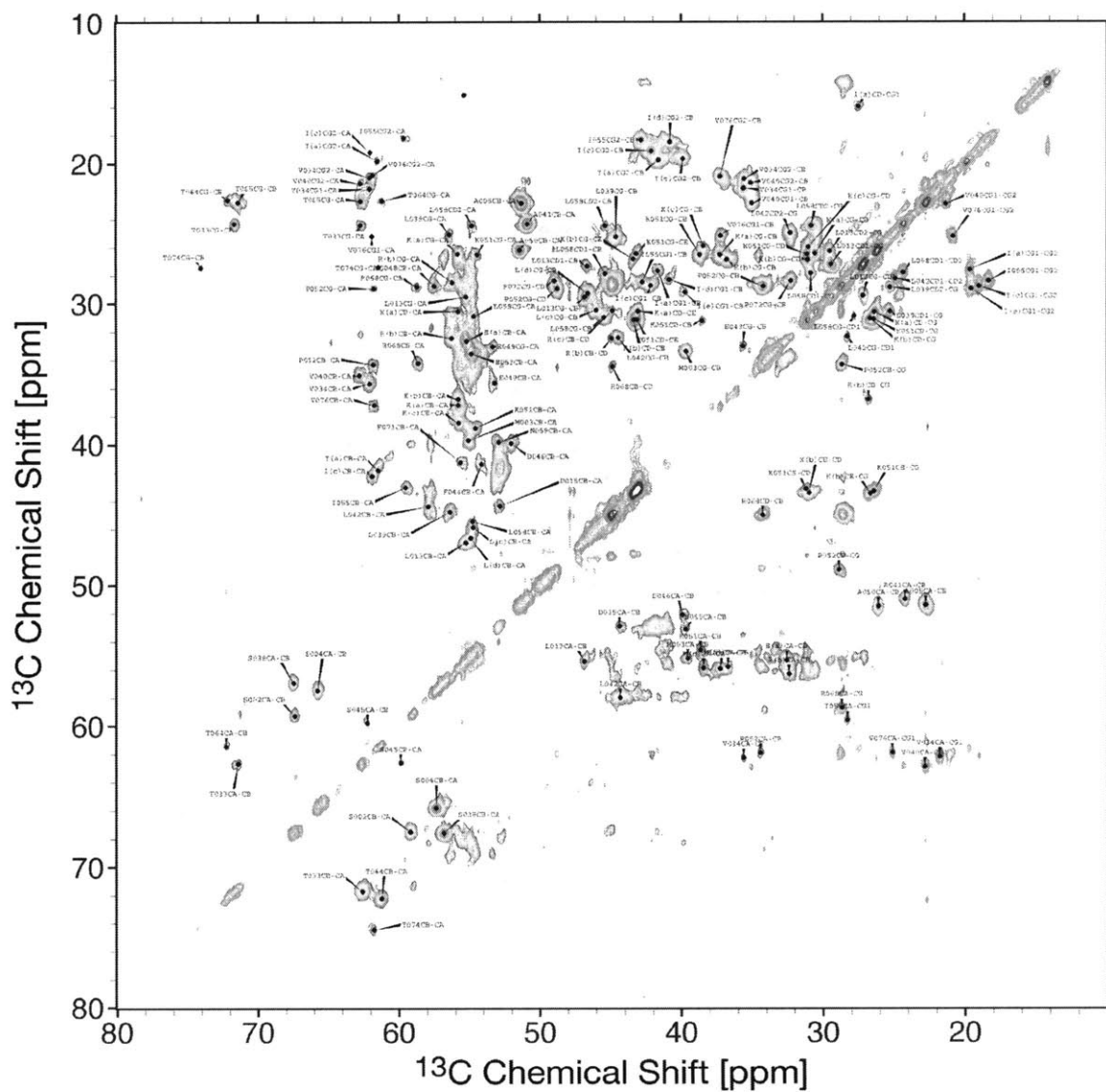
Figure S3. 2D ^{13}C-^{13}C correlation spectrum (DREAM mixing scheme) of U-PI3-SH3 recorded at 2°C and $\omega_r/2\pi$ = 16.67 kHz, 750 MHz (^1H) spectrometer.
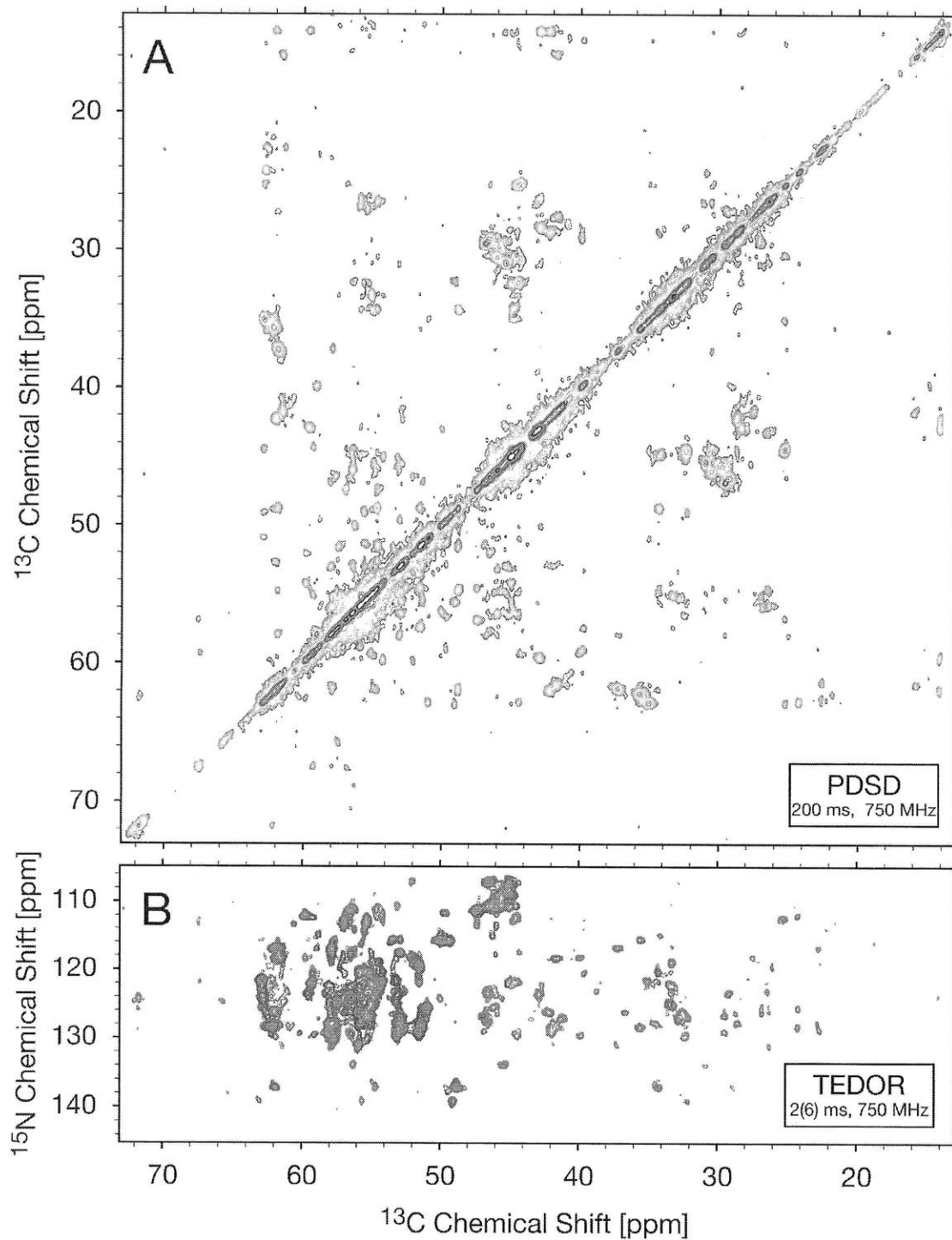
Figure S4. Dipolar correlation spectra of 2-PI3-SH3, recorded at 2°C and $\omega_r/2\pi = 10.0$ kHz. A: PDSD ^{13}C-^{13}C correlation spectrum using a 200 ms mixing time, recorded on a 750 MHz (^{1}H) spectrometer. B: ^{15}N-^{13}C correlation spectrum using TEDOR (1.6 ms mixing time shown in green-blue, 6 ms mixing time shown in grey), recorded on a 700 MHz (^{1}H) spectrometer.
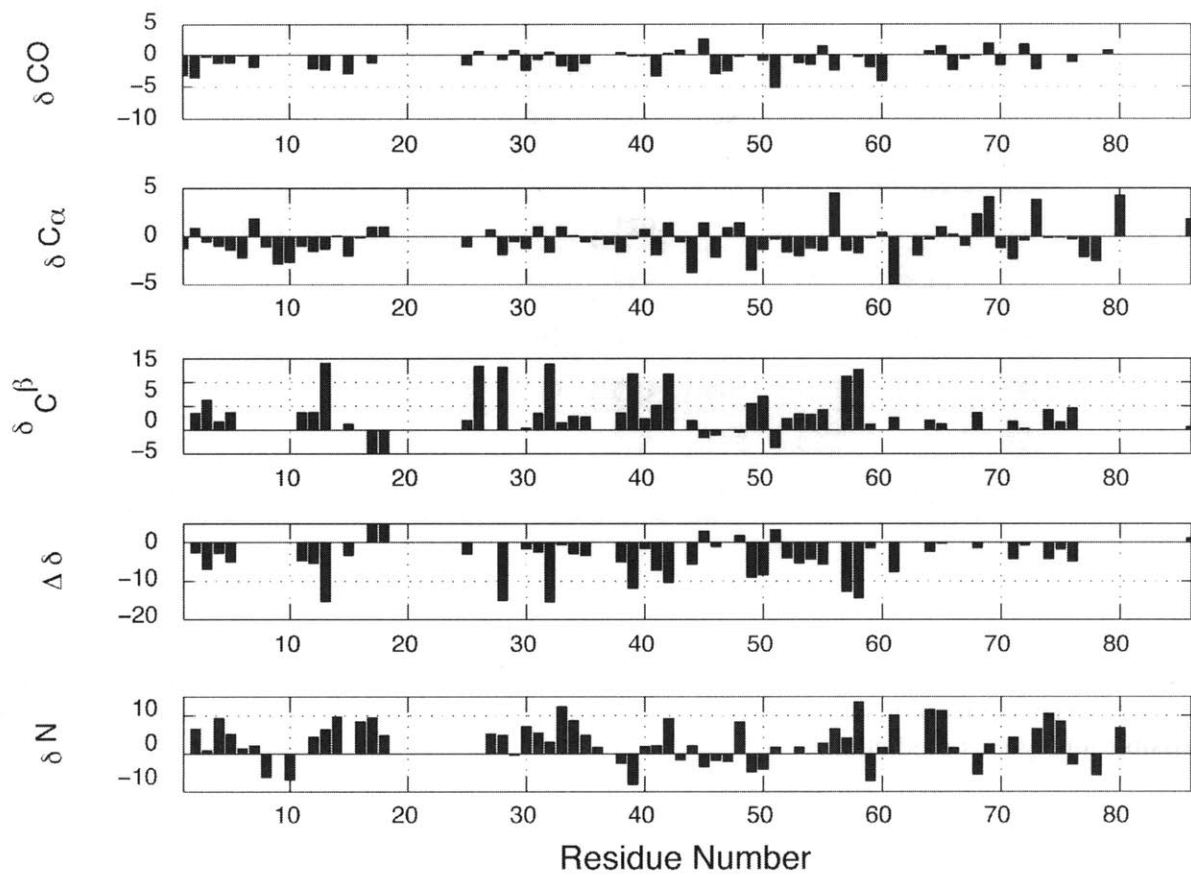
Figure S5. Full plot of the chemical shift analysis of PI3SH3. The chemical shift differences are calculated is described in the original paper.

Figure S6. Accuracy of the predicted torsion angles by the TALOS software package. The colors refer to the standard color scheme used in a TALOS analysis. Green: good prediction, yellow: possible ambiguous prediction.
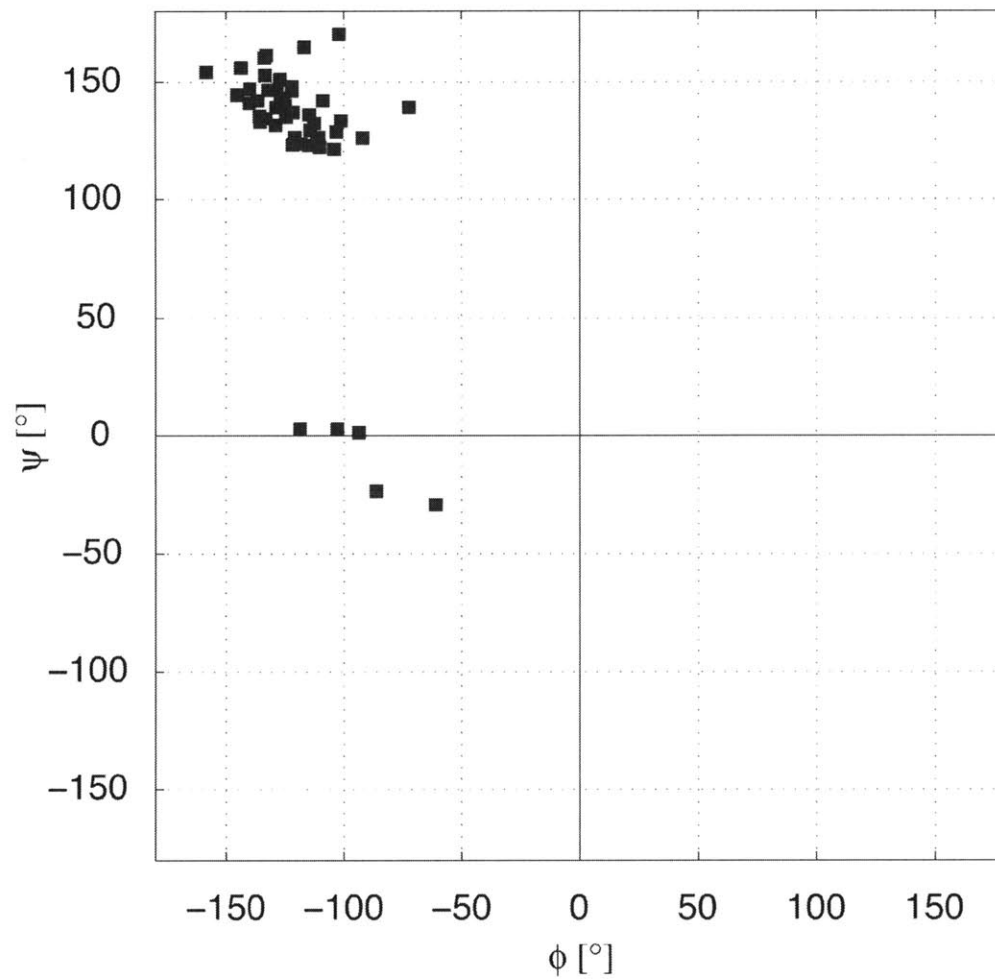
Figure S7. Plot of the ψ vs. φ torsion angles in the backbone of PI3-SH3 in amyloid fibril form, as determined by TALOS. Most well defined ("good") predictions fall into a region of β-strand conformation.

Table 1. MAS NMR ^{13}C and ^{15}N chemicals shifts of PI3-SH3 amyloid fibrils (in ppm)

Residue	C	Cα	Cβ	Cγ (Cγ1/Cγ2)	Cδ (Cδ1/Cδ2)	N
G001	170.7	44.25	/	/	/	32.21
S002	170.9	59.24	67.43	/	/	121.9
M003	175.0	55.07	39.63	33.26	19.66	120.5
S004	173.2	57.40	65.74	/	/	124.8
A005	176.5	51.42	22.73	/	/	128.6
E006	176.3	54.63	/	178.9	/	121.8
G007	172.0	47.37	/	/	/	11.2
Y008	/	56.91	/	131.1	/	113.3
Q009	/	55.31	/		/	/
Y010	/	55.32	/	128.2	/	112.5
R011		55.43	34.32	/	44.95	/
A012	175.5	51.32	22.79	/	/	128.0
L013	174.0	55.28	46.81	29.47	(27.16/26.16)	126.8
Y014	/	58.03	/	131.3	/	129.1
D015	173.4	53.14	42.12	178.3	/	/
Y016	/	57.84	/	129.0	/	127.9
K017	174.5	55.85	37.36	26.32	30.58/43.01	130.9
K018	175.7	55.89	37.00	26.71	30.84/43.47	126.2
E019	/	/	/	/	/	/
R020	/	/	/	/	/	/
E021	/	/	/	/	/	/
E022	/	/	/	/	/	/
D023	/	/	/	/	/	/
I024	/	/	/	/	/	/

D025	174.8	53.14	42.85	177.6	/	/
L026	176.9	/	46.13	29.6	/	/
H027	/	56.54	/	131.9	/	123.9
L028	175.6	54.69	/	30.54	/	125.3
G029	174.6	45.0	/	/	/	108.8
D030	173.9	52.96	41.25	177.4	/	127.1
I031	174.9	62.02	42.10	28.65	14.08	126.3
L032	176.8	54.94	46.58	29.24	/	123.4
T033	173.0	62.64	71.67	24.37	/	125.7
V034	173.2	62.15	35.58	(21.78/21.06)	/	128.5
N035	173.8	52.63	41.24	175.5	/	123.1
K036	/	54.63	/	/	/	123.1
G037	/	44.72	/	/	/	109.1
S038	174.9	56.81	37.53	/	/	112.9
L039	176.2	56.41	44.53	25.25	(30.41/28.71)	112.3
V040	175.5	62.81	35.04	(22.76/21.32)	/	121.6
A041	174.3	50.90	24.24	/	/	125.7
L042	176.6	57.97	44.45	32.33	(28.24/24.89)	129.6
G043	174.6	44.97	/	/	/	107.5
F044	/	54.21	41.38	139.7	/	121.8
S045	177.1	59.77	62.47	/	/	112.0
D046	173.3	52.03	39.78	179.7	/	118.2
G047	171.3	46.43	/	/	/	107.0
Q048	175.6	57.53	28.63	178.8	/	127.8
E049	176.64	53.37	35.64	32.98	179.8	115.6
A050	176.8	51.48	26.10	/	/	119.4
K051	171.8	54.61	38.70	31.15	(43.21/26.47)	123.1

P052	176.8	61.83	34.29	48.79	28.79	137.2
E053	175.2	54.85	33.44	179.7	/	122.2
E054	175.0	55.64	33.37	/	/	120.5
I055	177.0	59.56	42.87	14.15	(28.36/18.28)	123.6
G056	171.5	49.99	/	/	/	115.7
W057	/	56.32	40.87	111.6	/	124.2
L058	176.1	54.85	45.41	30.88	(27.78/24.37)	133.9
N059	173.2	53.06	36.69	/	/	111.0
G060	169.8	45.92	/	/	/	110.7
Y061	/	56.37	38.30	129.9	/	111.1
N062	/	53.02	41.54	/	/	129.6
E063	/	54.92	/	/	/	/
T064	175.3	61.35	72.12	22.65	/	124.8
T065	176.1	62.64	71.35	22.74	/	124.6
G066	171.6	45.73	/	/	/	110.6
E067	175.8	55.92	/	/	/	/
R068	175.9	58.69	34.24	40.85	28.75	114.9
G069	175.8	49.59	/	/	/	111.6
D070	174.7	53.01	/	/	/	/
F071	/	55.62	41.25	140.7	/	123.9
P072	178.6	63.06	32.24	49.04	28.39	139.2
G073	171.7	49.30	/	/	/	115.6
T074	/	61.54	74.31	27.41	/	123.8
Y075		57.89	40.58	129.7	/	128.0
V076	174.6	61.79	37.22	(25.11/20.85)	/	116.9
E077	/	54.75	/	/	178.9	/
Y078	/	55.44	/	129.2	/	113.8

I079	176.3	/	/	/	/
G080	/	49.75	/	/	/	115.8
R081	/	/	/	/	/	/
K082	/	/	/	/	/	/
K083	/	/	/	/	/	/
I084	/	/	/	/	/	/
S085	/	/	/	/	/	/
P086	/	65.22	35.52	49.41	27.62	137.9
E(a)	/	55.21	32.61	/	/	/
I(a)	175.4	61.50	41.66	15.77	(27.51/19.61)	118.6
I(d)	/	60.65	40.76	/	(28.24/18.39)	/
I(e)	/	59.04	39.85	/	(29.98/19.52)	/
K(c)	/	55.83	38.41	31.05	(43.00/25.97)	/
R(b)	/	56.23	32.38	36.63	28.44	/
R(c)	/	/	30.57	45.05	/	/

Chapter 8. Determination of molecular organization in amyloid fibrils by MAS NMR spectroscopy: The parallel, in-register tertiary structure of PI3-SH3 fibrils

Summary

This chapter reports further efforts to characterize the of amyloid fibrils formed by the SH3 domain of the PI3 kinase (PI3-SH3) via magic-angle spinning NMR spectroscopy. We describe two novel approaches for the determination of tertiary structure of amyloid fibrils, with particular emphasis on the identification of parallel, in-register β-sheets, a conformation that is generally difficult to characterize in protein fibrils. The methods we introduce are based on the examination of long-range ^{13}C-^{13}C correlation spectra of samples prepared with alternating ^{13}C-^{12}C labeling and lead to the determination of fibril tertiary structure by the detection of either indirect or direct inter-molecular constraints for multiple sites in the polypeptide chain. We demonstrate that spectral degeneracy can be exploited to identify a parallel, in-register inter-molecular conformation, and that, furthermore, this degeneracy can be circumvented by alternating labeling, resulting in the positive observation of inter-molecular correlations. The methods we present are well suited for the analysis of amyloid fibrils formed by moderately sized proteins, in which multiple site-specific constraints are necessary to characterize the fibril architecture.

238

8.1 Introduction

One of the appealing characteristics of solid-state nuclear magnetic resonance (NMR) spectroscopy is the possibility to study highly intricate molecular systems such as large proteins, membrane proteins, and protein assemblies, since the technique is not restricted by the molecular weight, solubility, or large-scale order of the sample examined. Amyloid fibrils are high molecular weight aggregates formed by peptides and proteins with a characteristic cross-β structure in which β-sheets run parallel to the fibril axis.(*1-3*) Several debilitating pathologies are related to amyloid fibrils and their oligomeric precursors, which are the target of therapeutic approaches.(*4*) In addition, non-pathological and functional types of amyloid have been recognized,(*5*) and the observation of fibril formation by globular proteins unrelated to disease indicates that amyloid fibrils may be generally accessible states of polypeptide chains.(*2, 6*) For these reasons, there is a great amount of interest to decipher the molecular structure of amyloid fibrils and their oligomeric precursors, from both biomedical and biophysical perspectives. Amyloid fibrils are highly amenable to structural analysis via solid-state magic-angle spinning (MAS) NMR spectroscopy because, even though the are not crystalline, they tend to adopt a well defined conformation, at least for part of the polypeptide chain, and thus yield high-quality MAS NMR spectra, as shown in various studies.(*7-13*) In addition, peptides and proteins used for *in vitro* studies can be isotopically labeled with considerable flexibility and produced in sufficient quantities, which allows MAS NMR measurements to provide structural information with a combination of sensitivity and site-specific resolution that surpasses that available by other biophysical techniques.(*14-16*)

The structures of proteins in amyloid fibril form differ conceptually from those of natively folded monomers. While the tertiary structure in the latter results from intra-molecular forces, in the case of amyloid fibrils the tertiary fold is typically determined by inter-molecular interactions that give rise to the core β-sheet structure, such as in $A\beta_{1-40}$ (*17-20*), or by a combination of inter- and intra-molecular contacts as in the β-solenoid of the HET-$s_{218-289}$ prion (*21, 22*). The β-sheets can be formed by parallel or antiparallel strands (or possibly a combination of both) and with residues in or out of register between neighboring strands, or possibly with mixed registry (*9, 18, 23*). The quaternary structure of amyloid fibrils may then be defined as the relative positions and orientations of the different β-sheets that make up the fibril,

and the protofilaments within it (if any), as suggested by Tycko and co-workers (24). If the polypeptide chain is long enough to form multiple layers of β-sheet structure, then the quaternary contacts between β-sheets can be either internal or external, depending on whether the interacting β-sheets belong to the same layer of molecules or to two adjacent layers (24), or there may be both types of contacts defining the quaternary structure of the fibril or protofilaments.

Despite the complexity of the molecular architecture of amyloid fibrils, MAS NMR studies have resulted in the elucidation of structural information of amyloid peptides and proteins at the secondary structure level, via resonance assignment and chemical shift analysis (7, 9, 12, 21, 25-27) or torsion angle measurements (8), and at the tertiary and quaternary level, via precise distance measurements or generation of multiple approximate distance constraints (20, 22, 24, 28-30). In the case of amyloid fibrils formed by peptides sufficiently short that they can be produced via solid-phase synthesis, the samples can be prepared with the incorporation of ^{13}C or ^{15}N labels positioned at specific sites depending on the hypothesis one wishes to test. The experiment thus consists in measuring specific distances, with the possibility of obtaining a null result if the inter-nuclear distance is too long to be able to measure via dipolar recoupling (typically > 7Å). A common hypothesis is a parallel, in-register arrangement of the β-sheets, which is typically tested by incorporation of a single ^{13}C label (usually a carbonyl due to the favorable relaxation characteristics of their NMR signals) and measurement of signal dephasing via the reintroduction of the homonuclear [13]C-^{13}C dipolar coupling. The rate of dephasing can be fit to an internuclear distance and the experiment is typically performed for various residues with separate samples. If the distances are close to the inter-strand distance for all residues tested, the result is consistent with a parallel in-register conformation. Indeed, several studies of amyloid-forming peptides have utilized this and similar approaches, such as inserting pairs of [13]C-^{13}C or [13]C-^{15}N nuclei, to determine tertiary fibril structure with success (17-20, 23, 31-35).

However, this specific labeling approach is limited in biosynthetically produced proteins, in which case isotopes are incorporated via singly ^{13}C-labeled amino acids, resulting in the labeling of all residues of a given type throughout the polypeptide sequence and thus loss of resolution. Although some conclusions may be drawn from such studies (29, 30), an approach that yields multiple site-specific data is more preferable in the study of proteins. An example of a multiple-site approach is preparing fibrils with a mixture of [^{13}C, ^{14}N] and [^{12}C, ^{15}N] labeled molecules and then obtaining tertiary structure information in the form of [15]N-^{13}C contacts

between adjacent β-strands in heteronuclear correlation spectra. This approach has been applied successfully (22), but typically suffers from low sensitivity due to spin dilution and the long internuclear distances involved, which aggravates the technical difficulties of multidimensional heteronuclear experiments.

Here we present two distinct approaches to the determination of tertiary structure in amyloid fibrils via MAS NMR spectroscopy, with particular emphasis on the identification of parallel, in-register β-sheets in protein fibrils. We investigate the tertiary structure of amyloid fibrils formed by the SH3 domain of PI3 kinase (PI3-SH3), an 86-residue protein that has been thoroughly characterized as a model for fibril formation (36-47). The secondary structure of PI3-SH3 fibrils is described in chapter 7, and here we set out to determine its tertiary structure via novel MAS NMR schemes. The methods we describe are based on the examination of long-range ^{13}C-^{13}C correlation spectra of samples prepared with alternating ^{13}C-^{12}C labeling (48-50) and lead to the determination of fibril tertiary structure by the observation of indirect or direct inter-molecular constraints for multiple sites simultaneously. Before presenting our analysis of PI3-SH3 fibril spectra, we demonstrate that inter-molecular correlations can be obtained with excellent sensitivity through the application of the band-selective radio frequency-driven recoupling (BASE RFDR) scheme (13) in a microcrystalline protein. However, the identification of such inter-molecular cross-peaks would generally be precluded in the case of the parallel, in-register β-sheets often found in amyloid fibrils due to the degeneracy of the spin system, since cross-peaks between neighboring strands would fully overlap with cross-peaks within each strand. Therefore, our first approach makes use of this expected degeneracy to infer that ^{13}Cα resonances presenting sensitive BASE RFDR cross-peaks to spins *distant in space* within their own strand but not to ^{13}Cα nuclei that are *distant in sequence* may be indicative of a parallel, in-register inter-molecular configuration. The second approch exploits the partial ^{13}C occupancy observed in several amino acid types of proteins produced with alternating labeling. In residues presenting fractional labeling, ^{13}Cα-^{13}Cβ pairs within each residue are usually not labeled simultaneously in a given protein molecule (50). Therefore we compare proton-driven spin diffusion (PDSD) spectra with short and long mixing times to recognize that certain ^{13}Cα-^{13}Cβ contacts can be attributed to arising from contacts between adjacent β-strands. Finally, we corroborate our findings via heteronuclear experiments on a mixed ^{15}N/^{13}C sample performed in bulk, one-dimensional fashion.

8.2 Results and Discussion

8.2.1 Inter-molecular BASE RFDR correlations

The BASE RFDR scheme in combination with alternating ^{13}C-^{12}C labeling using [2-^{13}C] glycerol has been shown to generate highly efficient correlation spectra between aliphatic ^{13}C nuclei *distant in space* (*13*), such as sequential $^{13}C\alpha(i)$-$^{13}C\alpha(i\pm1)$ and $^{13}C\alpha(i)$-$^{13}C\beta(i\pm1)$ contacts. The origin of the efficiency found in such an approach is the robustness of the pulse sequence with respect to experimental imperfections, the favorable recoupling effect of finite pulses, the elimination of unwanted, sequential $^{13}C\alpha$-^{13}C' polarization transfer by restricting effective recoupling to occur within the aliphatic bandwidth, and the inherent attenuation of dipolar truncation effects (*51*) observed for many spin systems in alternatingly labeled proteins. We therefore tested this approach in a microcrystalline sample of GB1 prepared with [2-^{13}C] glycerol (2-GB1), which resulted in sensitive cross-peaks between nuclei distant in sequence, with inter-nuclear distances up to ~6.5 Å, according to the crystal structure of GB1. Indeed, several long-range cross-peaks for most $^{13}C\alpha$ sites where observed in BASE RFDR spectra of GB1 with mixing times of 20 ms and above. One of these sites, Y45Cα, is shown in Figure 1 as a strip plot and depicts a representative example. In this figure, the asterisks mark cross-peaks to residues in a neighboring molecule in the crystal lattice.
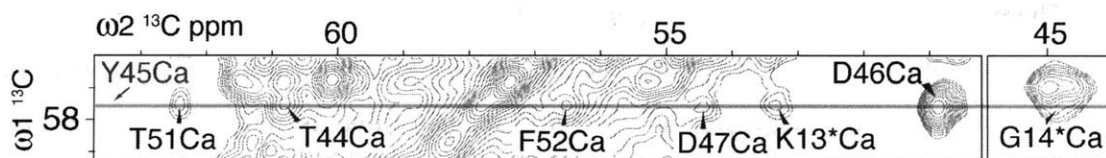
Figure 1. Subsection of a BASE RFDR spectrum of 2-GB1 recorded with a mixing time τ_{mix} = 24 ms, 12.5 kHz ^{13}C π pulses, and 80 kHz ^1H decoupling, at 700 MHz ^1H Larmor frequency and spinning frequency $\omega_r/2\pi$ = 12.5 kHz. The Y45Cα resonance is shown in the indirect dimension with a horizontal line. Various cross-peaks are observed to nuclei distant in space and in sequence, with the asteriks denoting residues in an adjacent protein molecule in the crystal.

We can then compare these sequential and long-range correlations to the crystal structure of GB1, shown in Figure 2. Residue Y45 is located in one of the outer β-strands that is part of an

antiparallel β-sheet. Therefore several of the backbone-backbone correlations seen in BASE RFDR spectra can be attributed to interactions with neighboring strands, both within the molecule and with an adjacent molecule, as seen in Figure 2.
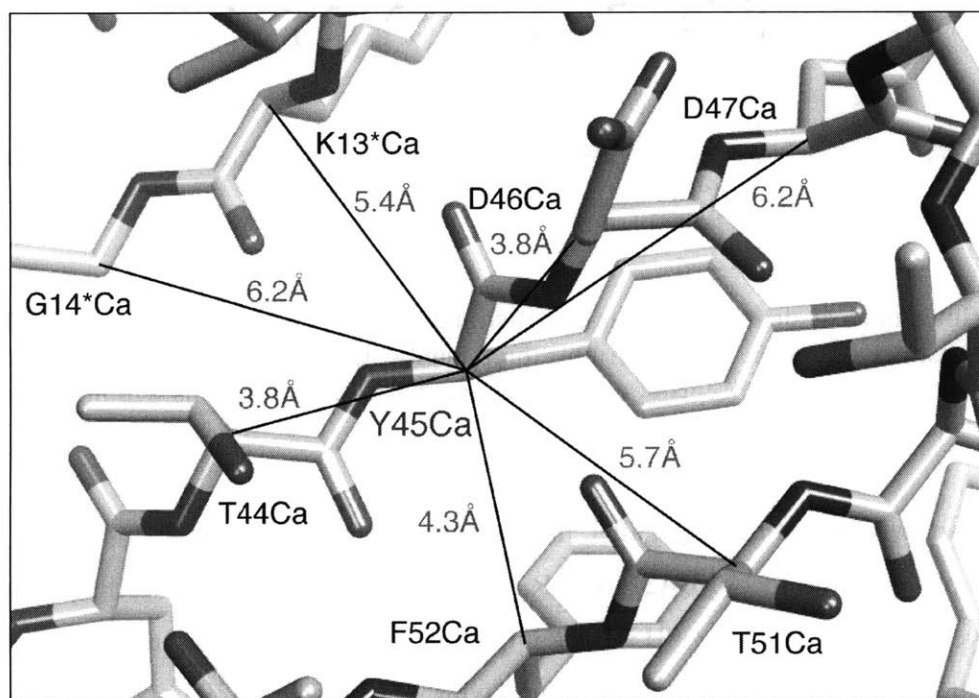
Figure 2. Inter-nuclear distances in the crystal structure of GB1 corresponding to cross-peaks observed in BASE RFDR spectra between Y45Cα and other ^{13}Cα sites: within its own strand (T44, D46, and D47), to an intra-molecular strand (T51 and F52), and to an inter-molecular strand (K13 and G14).

The pattern of BASE RFDR cross-peaks in the antiparallel β-strands of 2-GB1 is also expected for parallel β-strands since the inter-nuclear ^{13}Cα-^{13}Cα distances involved are similar in both cases. As Figure 3a shows, given three distinct parallel strands, multiple cross-peaks would be expected (as long as the nuclei are labeled) for such an arrangement. However, in the specific case of parallel, in-register formation of similar strands, which is common in amyloid fibrils, would lead to degeneracy of cross-peaks between inter-strand and intra-strand contacts.
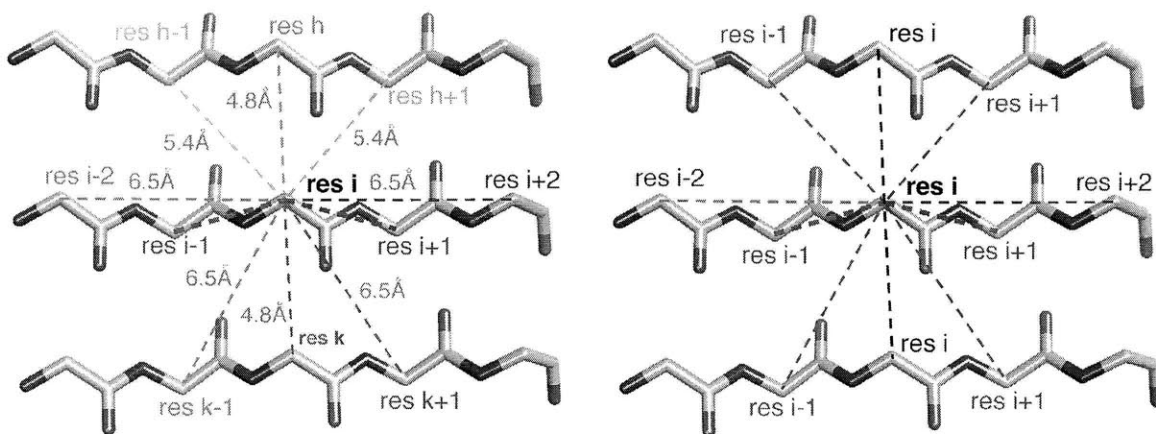
Figure 3. Inter-nuclear distances in parallel β-strands and resolvable $^{13}C\alpha$-$^{13}C\alpha$ correlations in (left) three distinct strands *h*, *i*, *k*, and (right) three identical, in-register strands *i*, *i*, *i*. Inter-strand correlations in the parallel, in-register case are degenerate with correlations within the strand.

8.2.2 Indirect determination of parallel, in-register tertiary structure

We proceeded to record BASE RFDR spectra of PI3-SH3 amyloid fibrils produced with [2-^{13}C] glycerol labeling (2-PI3-SH3). An important consideration was to attempt to obtain spectra with similar sensitivity as that obtained with 2-GB1. The 2-GB1 spectrum from which Figure 1 was extracted was recorded in 8 hours. Since ^{13}C cross-polarization spectra of microcrystalline GB1 present approximately four times more signal-to-noise per ^{13}C site than those of PI3-SH3 fibrils (due to higher sample density and the smaller size of GB1, 56 versus 86 residues), we needed to record BASE RFDR spectra of 2-PI3-SH3 for ~5.3 days. Long-term acquisition of robust two-dimensional experiments such as BASE RFDR or PDSD is feasible with some degree of adjustment between consecutive runs, which allows running the experiments for several days. Figure 4 illustrates a section of a BASE RFDR spectrum of 2-PI3-SH3 acquired with a 24 ms mixing period, and other parameters similar to those described for the experiment on 2-GB1. Several sequential $^{13}C\alpha$-$^{13}C\alpha$ cross-peaks are indicated for different regions of the PI3-SH3 sequence, including M3-S4-A5, R11-A12-L13-Y14, and F44-S45-D46. Interestingly, in addition to sequential cross-peaks between adjacent residues (i to i+1), there are also cross-peaks between $^{13}C\alpha$ nuclei separated by two residues (i to i+2), which correspond to

inter-nuclear distances of up to ~6.5 Å. However, no cross-peaks to $^{13}C\alpha$ nuclei distant in sequence (i to i±4 or longer) are observed, which is consistent with the highly degenerate structure of a parallel, in-register conformation.
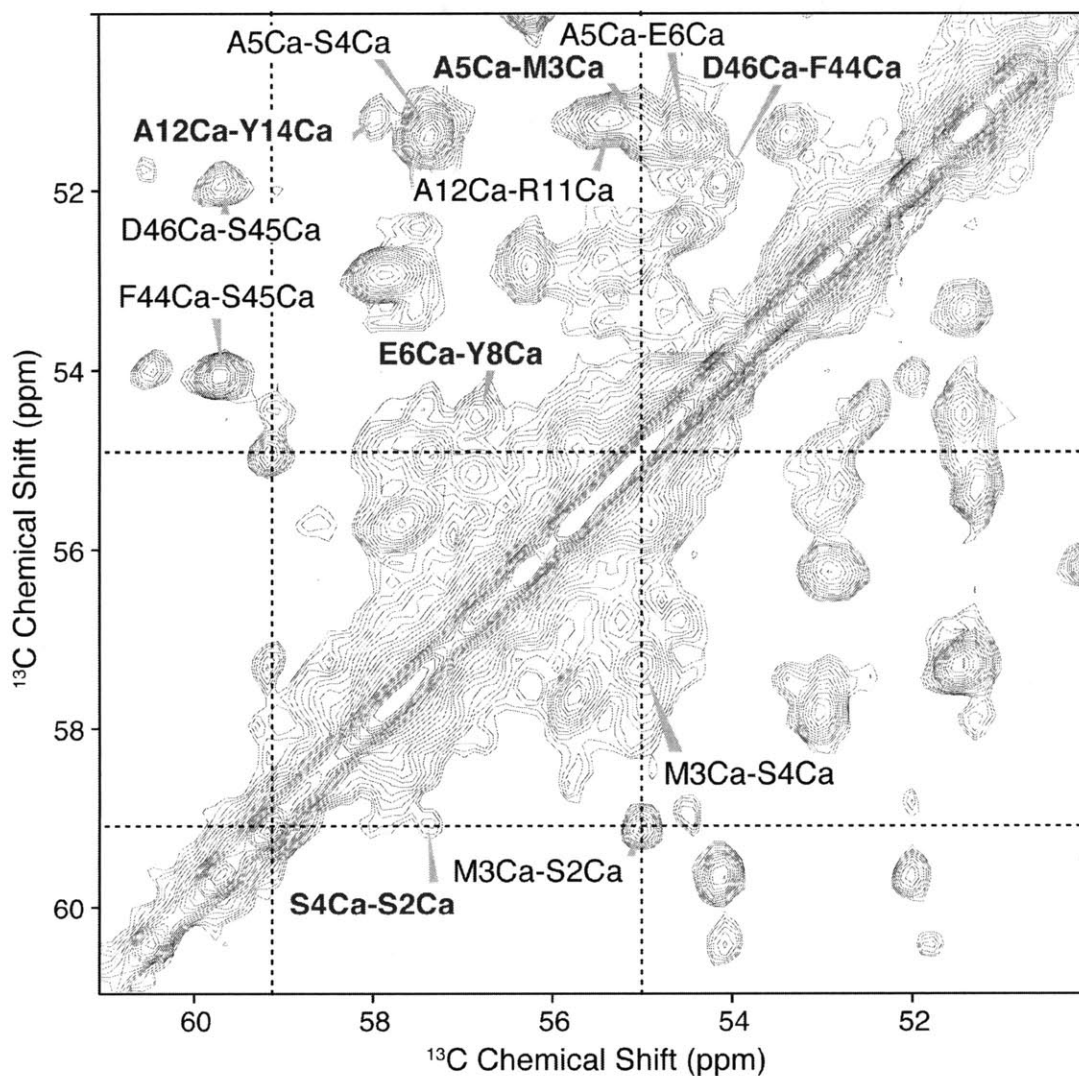
Figure 4. Section of a BASE RFDR spectrum of 2-PI3-SH3 recorded with a mixing time of 24 ms and other experimental parameters identical as those described for 2-GB1 in the caption of Figure 1. The experiment was averaged for ~5.3 days in order to match the signal-to-noise ratio obtained for 2-GB1, for comparison purposes. Blue bold labels indicate cross-peaks between $^{13}C\alpha$ nuclei separated by two residues, with an inter-nuclear distance corresponding to ~6.5 Å.
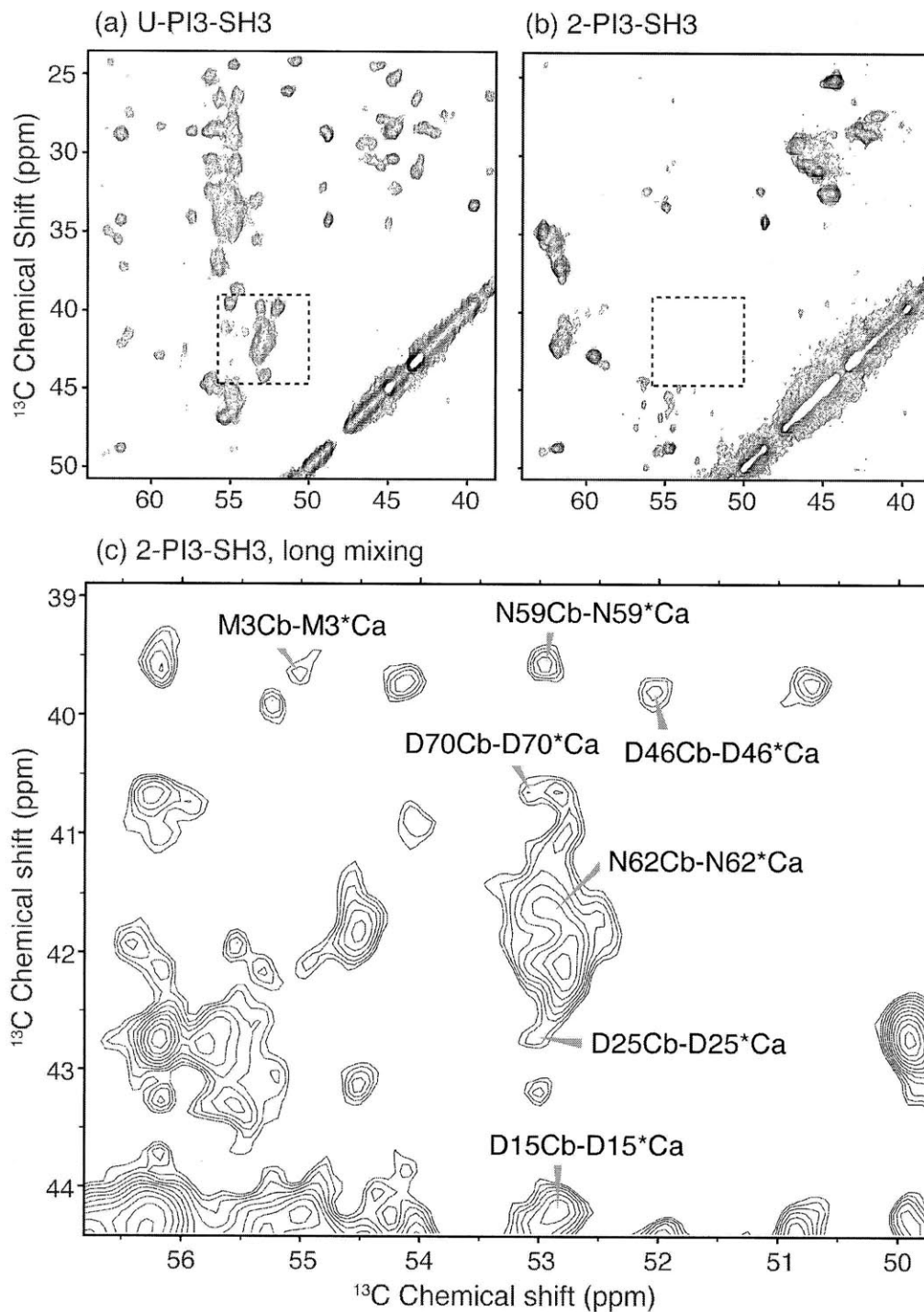
Figure 5. PDSD correlation spectra of (a) U-PI3-SH3 acquired with a short mixing time of 20 ms and 2-PI3-SH3 spectra acquired with (b) 20 ms and (c) 500 ms. The boxes in (a) and (b) are the same region as that shown in (c), in which asterisks mark correlations we conclude must occur between neighboring molecules in a parallel, in-register β-sheet.

8.2.3 Direct determination of parallel, in-register tertiary structure

Examination of long-range ^{13}C-^{13}C correlation spectra of 2-PI3-SH3 recorded with BASE RFDR, PDSD, or DARR, indicated a number of apparent intra-residue cross-peaks between ^{13}C sites that are not expected to be labeled concurrently in the same molecule in samples produced with alternating labeling. Therefore, we recorded PDSD ^{13}C-^{13}C correlation spectra of 2-PI3-SH3 with short and long mixing times in order to investigate the origin of such cross-peaks. We also recorded similar spectra in a PI3-SH3 sample prepared with uniform ^{13}C labeling (U-PI3-SH3). The dashed boxes in Figures 5a and 5b, and all of Figure 5c, denote the position of these interesting cross-peaks in the spectra of U-PI3-SH3 recorded with a short mixing time (5a), 2-PI3-SH3 recorded with a short mixing time (5b), and 2-PI3-SH3 recorded with a long mixing time (5c). We observe that the intra-residue ^{13}Cα-^{13}Cβ cross-peaks are present at short mixing time in the spectrum of the uniformly labeled sample (as expected) but not in the spectrum of the alternatingly labeled sample. However, at long mixing times in the latter sample, these cross-peaks emerge. Figure 6 illustrates the likely explanation of their origin, namely, inter-strand ^{13}Cα-^{13}Cβ correlations in a parallel, in-register β-sheet arrangement between ^{13}Cα in one strand and ^{13}Cβ in the adjacent strand. The random fractional labeling found for some type of residues ensures that these type of cross-peaks may be observed between adjacent molecules forming the β-sheet at long mixing periods. However, it is essential to perform the control experiment, that is, short mixing spectra of 2-PI3-SH3, which, as seen in Figure 5b, must show no cross-peaks to ensure the ^{13}Cα-^{13}Cβ pairs of interest are not labeled in the same molecule.
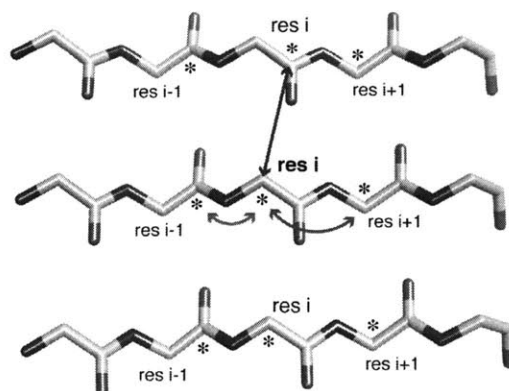
Figure 6. Inter-strand correlations (straight blue arrow) between fractionally labeled ^{13}Cα-^{13}C' sites, analogous to the ^{13}Cα-^{13}Cβ case of Figure 5. Curved arrows mark expected sequential cross-peaks.

8.2.4 Heteronuclear recoupling in a mixed $^{15}N/^{13}C$ sample

In order to corroborate our conclusions regarding a parallel, in-register tertiary structure in PI3-SH3 fibrils we prepared a fibril sample starting with a monomeric mixture of ^{15}N-labeled protein (with natural abundance ^{13}C) and ^{13}C-labeled protein (with [2-^{13}C] glycerol and natural abundance ^{15}N), with the objective of obtaining inter-molecular ^{15}N-^{13}C correlations in 2D spectra that can potentially be assigned to individual resonances. However, the experiments proved to be very insensitive, which prevented us from recording high-quality 2D spectra. Nevertheless, we were able to record a polarization build-up curve via transferred echo double resonance (TEDOR) in 1D fashion, shown in Figure 7. TEDOR ^{13}C signal builds up from the reintroduction of ^{15}N-^{13}C couplings. The maximum for this build-up, for the 4.2-4.4 Å internuclear distance found in β-sheets, should be reached with a mixing time of about 15 ms, wich is approximately consistent with our results. Other pulse sequences utilized for ^{15}N-^{13}C mixing provided similar or less sensitive results than the TEDOR data shown here. A second important observation, in addition to the maximum build-up time, is the transfer efficiency, which peaks at about 2.6 % in our data. According to numerical simulations, including similar relaxation parameters, the transfer efficiency for a 4.3 Å internuclear distance should be ~5%, which matches well with our results, considering the two-fold ^{15}N dilution factor in this sample.
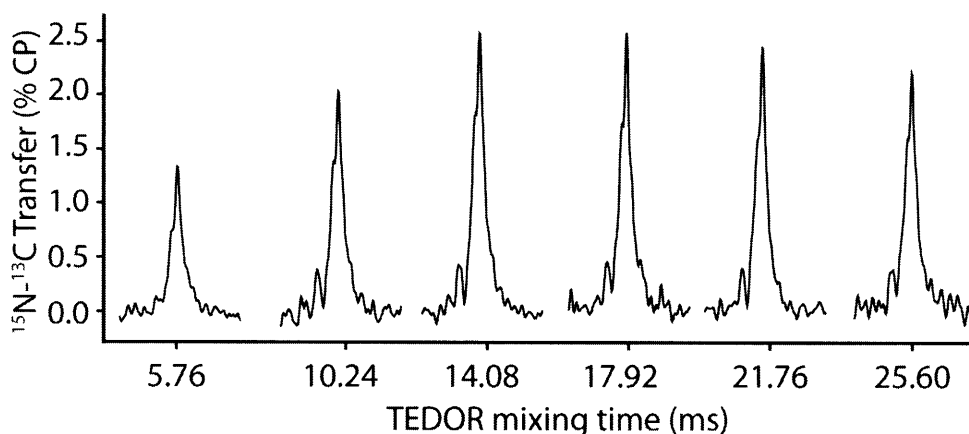
Figure 7. Carbonyl region of a TEDOR mixing build-up experiment recorded on a PI3-SH3 amyloid fibril sample obtained by mixing equal quantities of proteins labeled with either [^{15}N, ^{12}C] or [^{14}N, 2-^{13}C glycerol], prior to fibrillization. Each 1D spectrum was the average 16,000 scans and is scaled with respect to the full ^{13}C CP intensity.

The data shown in Figure 7 include all carbonyl resonances in the 86 residue-long polypeptide chain simultaneously, and while these "bulk" measurements do not provide the site-specific information that well-resolved multidimensional ^{15}N-^{13}C spectra could potentially yield, they are generally consistent with a ~4.3 Å distance between ^{13}C carbonyl sites in a given molecule and ^{15}N sites in a nearby molecule. This general observation is not sufficient to ascertain the specific type of tertiary structure, since for example similar inter-strand ^{15}N-^{13}C distances are found in parallel and anti-parallel β-sheets. However, the transfer efficiency would be considerably lower in a β-solenoid architecture, in which only a fraction (roughly a half) of the β-sheet interactions are inter-molecular. Thus this particular class of structure may be ruled out by the bulk measurements. Two-dimensional heteronuclear experiments on this kind of mixed sample will be necessary to corroborate the conclusions reached by the analysis of long-range ^{13}C-^{13}C correlation spectra in the strategies presented in the preceding sections.

8.3 Conclusions

We have described the identification of a parallel, in-register tertiary structure in PI3-SH3 amyloid fibrils via the examination of long-range ^{13}C-^{13}C MAS NMR correlation spectra of samples prepared with alternating labeling. We have demonstrated that BASE RFDR yields highly efficient polarization transfer between distant ^{13}C nuclei such as those in adjacent β-strands within and between molecules. This efficiency can be used to elucidate regions of high structural degeneracy in amyloid fibrils, which are consistent with a parallel, in-register inter-molecular organization. In a second, complementary approach, comparison of short-range and long-range ^{13}C-^{13}C correlation spectra results in the differentiation between intra-residue and inter-strand contacts due to mutually exclusive ^{13}C-^{12}C directly bonded pairs often found in alternating labeling. The major advantage of our strategies is that they rely on the analysis of experiments that are acquired routinely and efficiently in most structure determination efforts, and only require additional control experiments. The analysis described here should be applicable to fibrils formed by proteins of moderate and large sizes, depending on the spectral resolution attainable.

8.4 Materials and Methods

8.4.1 Protein samples

A sample of the $\beta 1$ domain of Immunoglobulin protein G (GB1, 56 residues) in microcrystalline form was used to demonstrate the long-range efficiency of BASE RFDR. This sample was prepared using 2-^{13}C glycerol and ^{13}C bicarbonate as the sole sources of carbon and uniform ^{15}N labeling (2-GB1). Further preparation details for GB1 are provided in chapter 6. Three different samples of PI3-SH3 amyloid fibrils were studied. The experimental approaches proposed in this chapter for the elucidation of inter-molecular organization in amyloid fibrils are based on alternating labeling with 2-^{13}C glycerol and ^{13}C bicarbonate as the sole sources of carbon, used for one of the PI3-SH3 samples (2-PI3-SH3), which was also uniformly ^{15}N labeled. A spectrum from a second sample, prepared with uniform ^{13}C and ^{15}N labeling (U-PI3-SH3), was recorded to compare to 2-PI3-SH3 spectra. The third PI3-SH3 fibril sample was obtained by mixing in solution, prior to fibrillization, equal quantities of two lyophilized protein preparations, one U-^{15}N labeled (with ^{13}C at natural abundance, 1 %) and the other produced with 2-^{13}C glycerol and ^{13}C bicarbonate as the sole sources of carbon with ^{15}N at natural abundance (~0.1 %). The fibrils were grown by incubation at pH 2.0 and 25 °C, in the presence of seeds of a well characterized form of PI3-SH3 fibrils (chapter 7), during a period of two weeks, resulting in the generation of PI3-SH3 amyloid-like fibrils that were subsequently spun down in a 60/40 d5-glycerol/water solvent (w/w), to help preserve hydration, maintain the sample integrity, and improve the spectral resolution. Approximately 18 mg of GB1 and 8 mg of PI3-SH3 were contained in each sample.

8.4.2 NMR spectroscopy

All experiments were performed in a spectrometer operating at 700 MHz ^{1}H Larmor frequency (courtesy of Dr. David J. Ruben, Francis Bitter Magnet Laboratory, Cambridge MA), corresponding to a 16.4 T magnetic field, using a triple resonance Varian/Chemagnetics (Palo Alto, CA) probe equipped with a 3.2 mm stator. Experiments were acquired with a 12.5 kHz spinning frequency and sample temperatures were maintained around 5 °C with a stream of cooled N_2 gas. Correlation experiments used ^{1}H-^{13}C cross-polarization preparation and high-power TPPM decoupling was applied on the ^{1}H channel during the chemical shift evolution and detection periods. 2D BASE RFDR experiments consisted of 320 real and 320 imaginary t1

points acquired in 56 μs increments with a 2.8 s recycle delay. The total acquisition times were 8.0 hours for 2-GB1 and 127.4 hours for 2-PI3-SH3, corresponding to 16 and 256 scans per t1 point, respectively. Similar acquisition parameters were used to record PDSD spectra, with 16 scans per t1 point for the U-PI3-SH3 spectrum and 192 scans per t1 point for the 2-PI3-SH3 spectra (with short and long mixing times). Data processing was performed with the NMRPipe suite and analyzed with the Sparky program.

References

(1)   Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol. 273*, 729-39.

(2)   Dobson, C. M. (2003) Protein folding and misfolding. *Nature 426*, 884-90.

(3)   Sunde, M., and Blake, C. C. (1998) From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Q. Rev. Biophys. 31*, 1-39.

(4)   Sacchettini, J. C., and Kelly, J. W. (2002) Therapeutic Strategies for Human Amyloid Diseases. *Nat. Rev. Drug Discovery 1*, 267-275.

(5)   Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem. 75*, 333-66.

(6)   Chiti, F., and Dobson, C. M. (2009) Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol. 5*, 15-22.

(7)   Jaroniec, C. P., MacPhee, C. E., Astrof, N. S., Dobson, C. M., and Griffin, R. G. (2002) Molecular conformation of a peptide fragment of transthyretin in an amyloid fibril. *Proc. Natl. Acad. Sci. U.S.A. 99*, 16748-53.

(8)   Jaroniec, C. P., MacPhee, C. E., Bajaj, V. S., McMahon, M. T., Dobson, C. M., and Griffin, R. G. (2004) High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A. 101*, 711-716.

(9) Heise, H., Hoyer, W., Becker, S., Andronesi, O. C., Riedel, D., and Baldus, M. (2005) Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A. 102*, 15871-6.

(10) Siemer, A. B., Ritter, C., Ernst, M., Riek, R., and Meier, B. H. (2005) High-resolution solid-state NMR spectroscopy of the prion protein HET-s in its amyloid conformation. *Angewandte Chemie (International ed in English) 44*, 2441-4.

(11) Siemer, A. B., Arnold, A. A., Ritter, C., Westfeld, T., Ernst, M., Riek, R., and Meier, B. H. (2006) Observation of highly flexible residues in amyloid fibrils of the HET-s prion. *Journal of the American Chemical Society 128*, 13224-8.

(12) Helmus, J. J., Surewicz, K., Nadaud, P. S., Surewicz, W. K., and Jaroniec, C. P. (2008) Molecular conformation and dynamics of the Y145Stop variant of human prion protein. *Proceedings of the National Academy of Sciences of the United States of America 105*, 6284-6289.

(13) Bayro, M. J., Maly, T., Birkett, N. R., Dobson, C. M., and Griffin, R. G. (2009) Long-Range Correlations between Aliphatic 13C Nuclei in Protein MAS NMR Spectroscopy. *Angew. Chem. Int. Ed. 48*, 5708-5710.

(14) Tycko, R. (2006) Methods in Enzymology. *Volume 413*, 103-122.

(15) Tycko, R. (2006) Molecular structure of amyloid fibrils: insights from solid-state NMR. *Q. Rev. Biophys. 39*, 1-55.

(16) Heise, H. (2008) Solid-state NMR spectroscopy of amyloid proteins. *ChemBioChem 9*, 179-189.

(17) Antzutkin, O. N., Balbach, J. J., Leapman, R. D., Rizzo, N. W., Reed, J., and Tycko, R. (2000) Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of beta-sheets in Alzheimer's beta-amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America 97*, 13045-50.

(18) Antzutkin, O. N., Leapman, R. D., Balbach, J. J., and Tycko, R. (2002) Supramolecular structural constraints on Alzheimer's beta-amyloid fibrils from electron microscopy and solid-state nuclear magnetic resonance. *Biochemistry 41*, 15436-50.

(19) Balbach, J. J., Petkova, A. T., Oyler, N. A., Antzutkin, O. N., Gordon, D. J., Meredith, S. C., and Tycko, R. (2002) Supramolecular structure in full-length Alzheimer's beta-

amyloid fibrils: evidence for a parallel beta-sheet organization from solid-state nuclear magnetic resonance. *Biophysical journal 83*, 1205-16.

(20)   Petkova, A. T., Ishii, Y., Balbach, J. J., Antzutkin, O. N., Leapman, R. D., Delaglio, F., and Tycko, R. (2002) A structural model for Alzheimer's beta -amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci. U.S.A. 99*, 16742-7.

(21)   Ritter, C., Maddelein, M. L., Siemer, A. B., Lv√°hrs, T., Ernst, M., Meier, B. H., Saupe, S. J., and Riek, R. (2005) Correlation of structural elements and infectivity of the HET-s prion. *Nature 435*, 844-8.

(22)   Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A. B., Riek, R., and Meier, B. H. (2008) Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science 319*, 1523-6.

(23)   Benzinger, T. L., Gregory, D. M., Burkoth, T. S., Miller-Auer, H., Lynn, D. G., Botto, R. E., and Meredith, S. C. (1998) Propagating structure of Alzheimer's beta-amyloid(10-35) is parallel beta-sheet with residues in exact register. *Proceedings of the National Academy of Sciences of the United States of America 95*, 13407-12.

(24)   Petkova, A. T., Yau, W. M., and Tycko, R. (2006) Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils. *Biochemistry 45*, 498-512.

(25)   Kloepper, K. D., Zhou, D. H., Li, Y., Winter, K. A., George, J. M., and Rienstra, C. M. (2007) Temperature-dependent sensitivity enhancement of solid-state NMR spectra of alpha-synuclein fibrils. *Journal of biomolecular NMR 39*, 197-211.

(26)   Andronesi, O. C., von Bergen, M., Biernat, J., Seidel, K., Griesinger, C., Mandelkow, E., and Baldus, M. (2008) Characterization of Alzheimer's-like paired helical filaments from the core domain of tau protein using solid-state NMR spectroscopy. *J. Am. Chem. Soc. 130*, 5922-8.

(27)   Heise, H., Celej, M. S., Becker, S., Riedel, D., Pelah, A., Kumar, A., Jovin, T. M., and Baldus, M. (2008) Solid-state NMR reveals structural differences between fibrils of wild-type and disease-related A53T mutant alpha-synuclein. *J. Mol. Biol. 380*, 444-50.

(28)   Iwata, K., Fujiwara, T., Matsuki, Y., Akutsu, H., Takahashi, S., Naiki, H., and Goto, Y. (2006) 3D structure of amyloid protofilaments of beta-2-microglobulin fragment probed by solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A. 103*, 18119-18124.

(29)   Shewmaker, F., Ross, E. D., Tycko, R., and Wickner, R. B. (2008) Amyloids of shuffled prion domains that form prions have a parallel in-register beta-sheet structure. *Biochemistry 47*, 4000-7.

(30)   Wickner, R. B., Dyda, F., and Tycko, R. (2008) Amyloid of Rnq1p, the basis of the [PIN+] prion, has a parallel in-register beta-sheet structure. *Proceedings of the National Academy of Sciences of the United States of America 105*, 2403-8.

(31)   Gregory, D. M., Benzinger, T. L., Burkoth, T. S., Miller-Auer, H., Lynn, D. G., Meredith, S. C., and Botto, R. E. (1998) Dipolar recoupling NMR of biomolecular self-assemblies: determining inter- and intrastrand distances in fibrilized Alzheimer's beta-amyloid peptide. *Solid State Nuclear Magnetic Resonance 13*, 149-66.

(32)   Benzinger, T. L., Gregory, D. M., Burkoth, T. S., Miller-Auer, H., Lynn, D. G., Botto, R. E., and Meredith, S. C. (2000) Two-dimensional structure of beta-amyloid(10-35) fibrils. *Biochemistry 39*, 3491-9.

(33)   Petkova, A. T., Leapman, R. D., Guo, Z., Yau, W. M., Mattson, M. P., and Tycko, R. (2005) Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. *Science (New York, NY) 307*, 262-5.

(34)   Sciarretta, K. L., Gordon, D. J., Petkova, A. T., Tycko, R., and Meredith, S. C. (2005) Abeta40-Lactam(D23/K28) models a conformation highly favorable for nucleation of amyloid. *Biochemistry 44*, 6003-14.

(35)   Lim, K. H., Nguyen, T. N., Damo, S. M., Mazur, T., Ball, H. L., Prusiner, S. B., Pines, A., and Wemmer, D. E. (2006) Solid-state NMR structural studies of the fibril form of a mutant mouse prion peptide PrP89-143(P101L). *Solid state nuclear magnetic resonance 29*, 183-90.

(36)   Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D., and Dobson, C. M. (1998) Amyloid Fibril Formation by an SH3 Domain. *Proc. Natl. Acad. Sci. U.S.A. 95*, 4224-4228.

254

(37)   Zurdo, J., Guijarro, J. I., and Dobson, C. M. (2001) Preparation and characterization of purified amyloid fibrils. *J. Am. Chem. Soc. 123*, 8141-2.

(38)   Zurdo, J., Guijarro, J. I., Jimenez, J. L., Saibil, H. R., and Dobson, C. M. (2001) Dependence on solution conditions of aggregation and amyloid formation by an SH3 domain. *J. Mol. Biol. 311*, 325-40.

(39)   Ventura, S., Lacroix, E., and Serrano, L. (2002) Insights into the origin of the tendency of the PI3-SH3 domain to form amyloid fibrils. *J. Mol. Biol. 322*, 1147-58.

(40)   Polverino de Laureto, P., Taddei, N., Frare, E., Capanni, C., Costantini, S., Zurdo, J., Chiti, F., Dobson, C. M., and Fontana, A. (2003) Protein aggregation and amyloid fibril formation by an SH3 domain probed by limited proteolysis. *J. Mol. Biol. 334*, 129-41.

(41)   Ventura, S., Zurdo, J., Narayanan, S., Parreño, M., Mangues, R., Reif, B., Chiti, F., Giannoni, E., Dobson, C. M., Aviles, F. X., and Serrano, L. (2004) Short amino acid stretches can mediate amyloid formation in globular proteins: the Src homology 3 SH3 case. *Proc. Natl. Acad. Sci. U.S.A. 101*, 7258-63.

(42)   Carulla, N., Caddy, G. L., Hall, D. R., Zurdo, J., Gairi, M., Feliz, M., Giralt, E., Robinson, C. V., and Dobson, C. M. (2005) Molecular recycling within amyloid fibrils. *Nature 436*, 554-8.

(43)   Ahn, H. C., Le, Y. T., Nagchowdhuri, P. S., Derose, E. F., Putnam-Evans, C., London, R. E., Markley, J. L., and Lim, K. H. (2006) NMR characterizations of an amyloidogenic conformational ensemble of the PI3K SH3 domain. *Protein Sci. 15*, 2552-7.

(44)   Bader, R., Bamford, R., Zurdo, J., Luisi, B. F., and Dobson, C. M. (2006) Probing the mechanism of amyloidogenesis through a tandem repeat of the PI3-SH3 domain suggests a generic model for protein aggregation and fibril formation. *J. Mol. Biol. 356*, 189-208.

(45)   Orte, A., Birkett, N. R., Clarke, R. W., Devlin, G. L., Dobson, C. M., and Klenerman, D. (2008) Direct characterization of amyloidogenic oligomers by single-molecule fluorescence. *Proc. Natl. Acad. Sci. U.S.A. 105*, 14424-9.

(46)   Buell, A. K., Tartaglia, G. G., Birkett, N. R., Waudby, C. A., Vendruscolo, M., Salvatella, X., Welland, M. E., Dobson, C. M., and Knowles, T. P. (2009) Position-dependent electrostatic protection against protein aggregation. *Chembiochem 10*, 1309-12.

(47)   Carulla, N., Zhou, M., Arimon, M., Gairí, M., Giralt, E., Robinson, C. V., and Dobson, C. M. (2009) Experimental characterization of disordered and ordered aggregates populated during the process of amyloid fibril formation. *Proc. Natl. Acad. Sci. U.S.A. 106*, 7828-33.

(48)   LeMaster, D. M., and Kushlan, D. M. (1996) Dynamical mapping of E. coli thioredoxin via 13C NMR relaxation analysis. *J. Am. Chem. Soc. 118*, 9255-9264.

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

(50)   Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K., and Oschkinat, H. (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature 420*, 98-102.

(51)   Bayro, M. J., Huber, M., Ramachandran, R., Davenport, T. C., Meier, B. H., Ernst, M., and Griffin, R. G. (2009) Dipolar truncation in magic-angle spinning NMR recoupling experiments. *J. Chem. Phys. 130*, 114506.

Chapter 9. Solid-State NMR Analysis of Bacterial Organelles: Implications for a Model of Gas Vesicle Assembly

Adapted from "Solid-state NMR evidence for inequivalent GvpA subunits in gas vesicles" by Astrid C. Sivertsen, Marvin J. Bayro, Marina Belenky, Robert G. Griffin, and Judith Herzfeld. Journal of Molecular Biology 387, 1032-1039 (2009).

Summary

Gas vesicles are organelles that provide buoyancy to the aquatic microorganisms that harbor them. The gas vesicle shell consists almost exclusively of the hydrophobic 70-residue protein GvpA, arranged in an ordered array. Solid-state NMR spectra of intact, collapsed gas vesicles from the cyanobacterium *Anabaena flos-aquae* show duplication of certain GvpA resonances, indicating that specific sites experience at least two different local environments. Interpretation of these results in terms of an asymmetric dimer repeat unit can reconcile otherwise conflicting features of the primary, secondary, tertiary and quaternary structures of the gas vesicle protein. In particular, the asymmetric dimer can explain how the hydrogen bonds in the β-sheet portion of the molecule can be oriented optimally for strength while promoting stabilizing aromatic and electrostatic side-chain interactions among highly conserved residues and creating a large hydrophobic surface suitable for preventing water condensation inside the vesicle.

9.1 Introduction

Gas vesicles are buoyancy organelles that are found in a wide range of aquatic microorganisms. By assembling and disassembling these vesicles, organisms are able to regulate their depth in the water column according to their needs for light, air and nutrients. The gas content of the hollow vesicles reflects passive equilibrium with gas molecules dissolved in the aqueous phase. Given that permeable species include not only such small, non-polar molecules as H_2, N_2, O_2, Ar, and CO_2, but also the polar CO molecule and the large perfluorocyclo-butane, C_4F_8 molecule (6.3 Å diameter)[1], it is assumed that H_2O is also permeable and that the absence of condensed water inside the vesicles reflects a highly hydrophobic and highly concave inner surface without suitable nucleation sites[1,2]. This view is supported by accumulating evidence that proximal hydrophobic surfaces in bulk water produce bubbles between them[3].

Electron microscopy of intact vesicles[4,5] shows shapes ranging from acorn-like spindles to regular cylinders with conical end caps. In *Anabaena flos-aquae,* the cylinders dominate (see Figure 1) and are typically about 5000 Å long and 750 Å wide. Close examination shows that the vesicle is bipolar, with ribs forming a low-pitch helix[6] on each side of an apparent insertion seam located in the cylindrical region. A 45.7 Å rib-rib distance has been measured by X-ray diffraction[7] and by atomic force microscopy.[8]

The gas vesicle wall comprises a shell formed almost exclusively by repeats of the 70-residue gas vesicle protein A[9] (GvpA) with a small amount of the ~3-fold larger gas vesicle protein C (GvpC) adhering loosely to the outer surface[10]. However, little is known about the GvpA fold: high resolution electron microscopy is not possible because of multiple scattering, and solution NMR of intact vesicles is not possible because of their large size. Furthermore, since gas vesicles dissolve only under denaturing conditions and subsequent dialysis yields only amorphous precipitates, neither crystallographic nor solution NMR studies have been feasible. However, FTIR spectra (obtained in collaboration with Mark Braiman) provide indications of anti-parallel β-sheet, and X-ray diffraction[7] and AFM[8] agree that β-strands are tilted 36° from the cylindrical axis of the vesicle. The corresponding orientation of the inter-strand hydrogen bonds at 54° to the cylinder axis, is the ideal for mechanical stability in both the length and width directions[1].
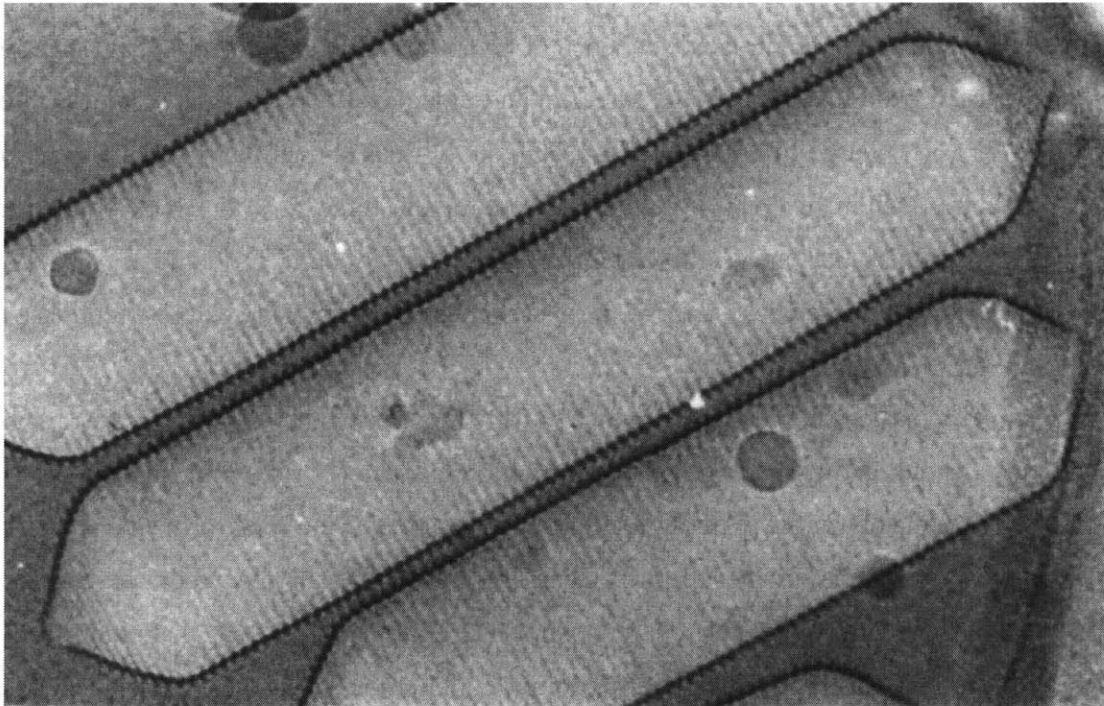
Figure 1. Electron micrograph (~150,000 x) of gas vesicles from *Anabaena flos-aquae* (courtesy of N. Grigorieff and A.E. Walsby).

```
(a)         1           11          21          31          41          51          61
GvpA        AVEKTNSSSS  LAEVIDRILD  KGIVIDAWVR  VSLVGIELLA  IEARIVIASV  ETYLKYAEAV  GLTQSAAVPA
Sec struc   CCCCCCCCCC  HHHHHHHHCC  CCCEEEEEEE  EEEECCEEEE  EEEEEEEEEH  HHHHHHHHHC  CCCCCCCCCC
Confidence  9741378771  5134433205  6617877899  9982641321  58888 76526 9999998642  8735576679


(b)         1           11          21          31          41          51          61
Anabaena    AVEKTNSSSS  LAEVIDRILD  KGIVIDAWVR  VSLVGIELLA  IEARIVIASV  ETYLKYAEAV  GLTQSAAVPA
Cyanobac    AVEK NS SS  L EV+DRILD  KG+V+D W R  +SLVGIEL+   +EAR V+ASV  ETYLKYAE +  GLT  A+ P+
All         ++      +   + +V+DR+L+  KG+ +D + +  ++LVGI+++   +E R V A +  +T+L Y++ +            +
```

Figure 2. (a) Amino acid sequence of *A. flos-aquae* GvpA and corresponding PSIPRED[12] secondary structure prediction (where H indicates α-helix, E β-sheet, and C coil) with associated estimates of confidence (where 0 = low, 9 = high). (b) BLAST[13] results for residue conservation among GvpA in cyanobacteria (middle row) and cyanobacteria, other bacteria, and archaea (bottom row). Letters indicate absolute conservation, + indicates conservative substitutions, and blank spaces no conservation.

Figure 2 shows the amino acid sequence of GvpA in *Anabaena flos-aquae*. There are six positively charged residues (three R and three K) and nine negatively charged residues (three D and six E), for a net charge of –3. A MALDI-TOF study[11] has shown that (1) there is no post-translational modification of GvpA, (2) the only one of the three R-X and K-X bonds that is accessible to trypsin is the one in the N-terminus, (3) none of nine D-X and E-X bonds are accessible to endoproteinase GluC, and (4) the C-terminus is accessible to carboxypeptidase only as far as the S65-A66 bond.

Secondary structure prediction using PSIPRED[12] (and other algorithms, not shown) defines likely α-helix and β-sheet regions with high confidence (as shown in Figure 2a). A coil prediction approximately midway through a long β-sheet stretch (too long to fit in one rib of the vesicle) suggests the location of a β-turn. Charges in each half of the predicted β-sheet are conspicuously arranged as oppositely charged pairs (D26,R30 and E52,R54) that could form salt bridges with the same pairs in a neighboring anti-parallel strand. This arrangement would also allow the unique tryptophan (W28) in each subunit to interact with that of a neighboring subunit.

BLAST[13] results comparing the *Anabaena flos-aquae* GvpA sequence with those of other cyanobacteria and with aquatic micro-organisms generally (Figure 2b) show a remarkably conserved core in the sequence, coinciding with the predicted α-helix and β-sheet segments. The putative β-turn region is absolutely conserved and the flanking β-sheet segments show absolute conservation or conservative substitution of all the aromatic and charged residues, including the tryptophan and all of the charge pairs noted above. This conservation suggests the importance of aromatic and electrostatic interactions for the structure of gas vesicles.
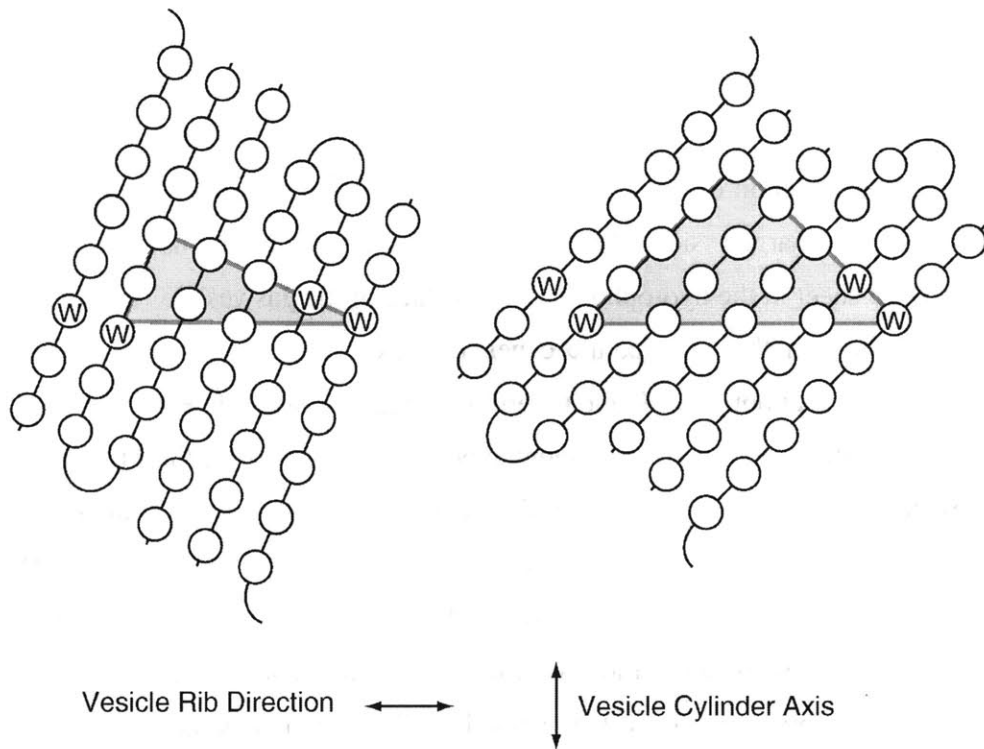
Figure 3. Dependence of the strand tilt relative to the vesicle cylinder axis (vertical) on the alignment of strands in the putative β-sheet (gray triangles). Whereas strands are tilted just 20° when the primary sequence is translated by just two residues over two monomers (left), they are tilted 36° when the primary sequence is translated by four residues over two monomers (right).
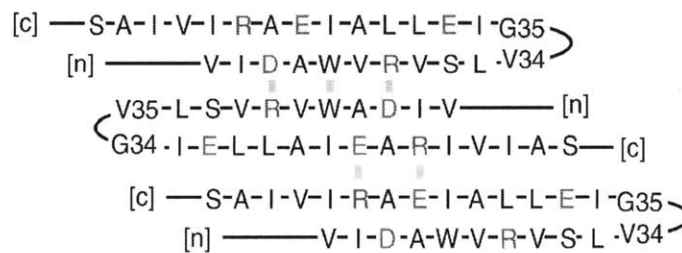
Figure 4. Model of the β-sheet portion of GvpA with even-numbered turns in identical subunits and in-register arrangement of the conserved aromatic and charged residues at subunit interfaces. Short gray bars indicate stabilizing side chain-side chain interactions, and [c] and [n] indicate C- and N-terminal ends of the polypeptide chain. The translation of four residues over two subunits gives the correct strand tilt. However, with an even-numbered turn, the two sets of salt bridges are on opposite sides of the sheet.

261

However, it is not so easy to satisfy these side chain interactions for the anti-parallel β-strands while also orienting the strands at the observed tilt of 36° from the vesicle axis, which requires a translation of four residues over two subunits (as shown in Figure 3). Figure 4 shows that this can be done with an even-numbered β-turn centered between V34 and G35, but even-numbered turns put charged side chains on both sides of the β-sheet, which precludes participation of the sheet in the hydrophobic inner surface of the gas vesicle. The alternative, that the hydrophobic faces of the amphipathic α-helices make up the inner surface of the vesicle, seems unlikely, given that both the N- and C-termini are accessible to digestion by proteases[11].

Figures 5-7 show options for odd-numbered β-turns. Centering the turn on G35 or I36 (Figure 5a,b) leads to a translation of two residues over two subunits in either direction, resulting in a tilt of only 20°. On the other hand, centering an odd-numbered turn on V34, (Figure 5c) leads to a translation of six residues over two subunits, resulting in a tilt of 48°. Evidently achieving the correct tilt with odd-numbered turns requires either shifting the register of subunits or relaxing the assumption of equivalent subunits. These two approaches are illustrated in Figures 6 and 7, respectively. Figure 6 shows the same turn as in Figure 5a with a two-residue shift in one subunit interface or the other: in Figure 6a, the DAWVR registration is preserved, while the EAR motifs are two residues out-of-register, whereas in Figure 6b, the DAWVR motifs are two residues out-of-register, while the EAR registration is preserved. In both cases, interactions between the out-of-register residues would require side chains to be stretched along a diagonal between the strands. In contrast, the model in Figure 7 preserves the close aromatic and electrostatic interactions while achieving the correct tilt by centering the odd-numbered β-turns in alternating subunits on different residues, specifically V34 and G35, to give the correct translation of four residues over two subunits.

(a)

[c] ——S-A-I-V-I-R-A-E-I-A-L-L-E-I36 ~
[n] ———V-I-D-A-W-V-R-V-S-L-V34~ G35

G35~V34-L-S-V-R-V-W-A-D-I-V———[n]
~I36-E-L-L-A-I-E-A-R-I-V-I-A-S——[c]

[c] ——S-A-I-V-I-R-A-E-I-A-L-L-E-I36 ~
[n] ———V-I-D-A-W-V-R-V-S-L-V34~ G35

(b)

[c]——S-A-I-V-I-R-A-E-I-A-L-L-E37~
[n]——I-D-A-W-V-R-V-S-L-V-G35~ I36

I36~G35-V-L-S-V-R-V-W-A-D-I———[n]
~E37-L-L-A-I-E-A-R-I-V-I-A-S——[c]

[c]——S-A-I-V-I-R-A-E-I-A-L-L-E37~
[n]——I-D-A-W-V-R-V-S-L-V-G35~ I36

(c)

[c]—S-A-I-V-I-R-A-E-I-A-L-L-E-I-G35~
[n]——I-V-I-D-A-W-V-R-V-S-L33~ V34

V34~L33-S-V-R-V-W-A-D-I-V-I——[n]
~G35-I-E-L-L-A-I-E-A-R-I-V-I-A-S——[c]

[c]—S-A-I-V-I-R-A-E-I-A-L-L-E-I-G35~
[n]——I-V-I-D-A-W-V-R-V-S-L33~ V34
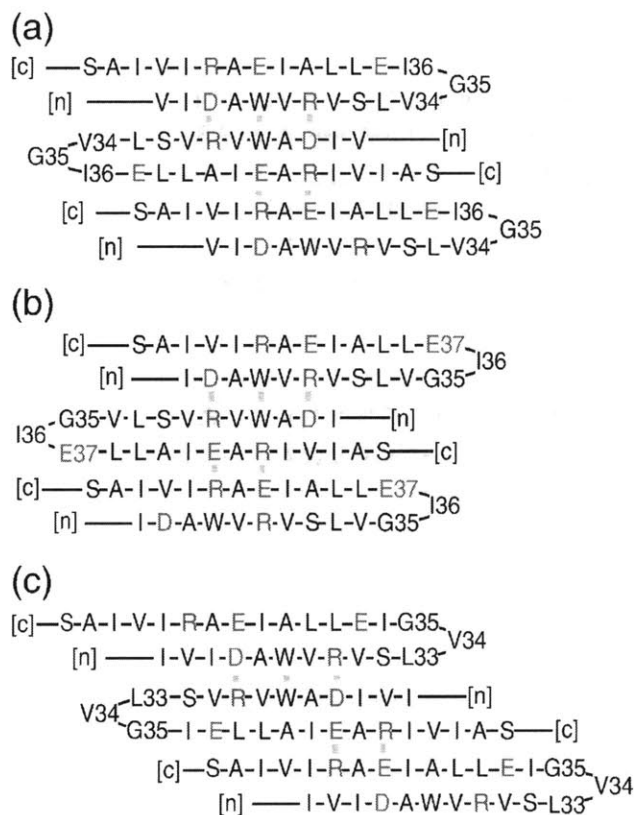
Figure 5. Models of the β-sheet portion of GvpA with all the salt bridges on one side of the sheet, identical subunits, and in-register arrangement of conserved aromatic and charged residues at the subunit interfaces. The notation is as in Figure 2. The tilts of the strands are wrong in all cases: (a) odd-numbered turns centered on G35 give a translation of two residues over two subunits; (b) odd-numbered turns centered on V34 also give a translation of two residues over two subunits, but in the other direction; and (c) odd-numbered turns centered on I36 give a translation of 6 residues over two subunits.

The model in Figure 7 uniquely invokes an asymmetric dimer as the fundamental building block of gas vesicles. In this case, NMR spectra would exhibit two different chemical shifts for at least some residues. Since solid-state NMR does not rely on fast molecular tumbling or large three-dimensional crystals, it is ideally suited technique for studying gas vesicles. Here we present the first solid-state NMR results for vesicles from *Anabaena flos-aquae* in their intact, collapsed state. Our data indicate that there are inequivalent GvpA subunits in the gas vesicle structure, lending plausibility to the β-sheet structure proposed in Figure 7.

263

(a)

```
[c]——S-A-I-V-I-R-A-E-I-A-L-L-E-I36
                                       `G35
   [n]——I-V-I-D-A-W-V-R-V-S-L-V34´

    G35‚V34-L-S-V-R-V-W-A-D-I-V-I——[n]
       `I36-E-L-L-A-I-E-A-R-I-V-I-A-S——[c]

      [c]——S-A-I-V-I-R-A-E-I-A-L-L-E-I36
                                         `G35
         [n]——I-V-I-D-A-W-V-R-V-S-L-V34´
```

(b)

```
[c]——S-A-I-V-I-R-A-E-I-A-L-L-E-I36
                                       `G35
   [n]——I-V-I-D-A-W-V-R-V-S-L-V34´

    G35‚V34-L-S-V-R-V-W-A-D-I-V-I——[n]
       `I36-E-L-L-A-I-E-A-R-I-V-I-A-S——[c]

      [c]——S-A-I-V-I-R-A-E-I-A-L-L-E-I36
                                         `G35
         [n]——I-V-I-D-A-W-V-R-V-S-L-V34´
```
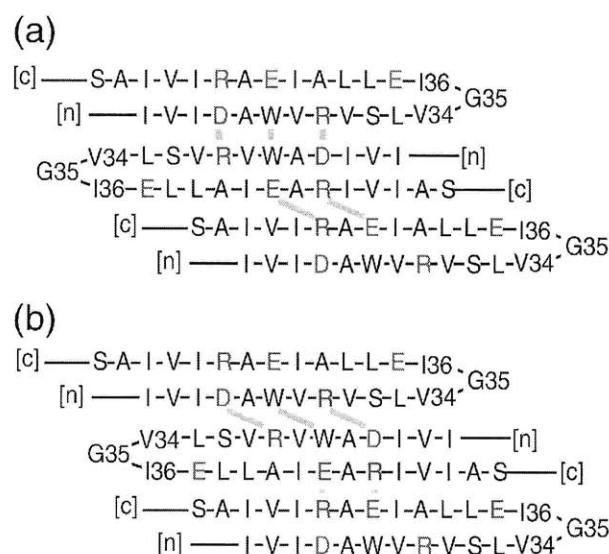
Figure 6. Models of the β-sheet portion of GvpA with all salt-bridges on one side of the sheet, identical subunits, and correctly tilted strands obtained by shifting subunits relative to each other. The notation is as in Figure 2. (a) The EAR segments are shifted. (b) The DAWVR segments are shifted.

```
[c]——S-A-I-V-I-R-A-E-I-A-L-L-E-I36
                                       `G35
   [n]——I-V-I-D-A-W-V-R-V-S-L-V34´

    V34‚L33-S-V-R-V-W-A-D-I-V-I——[n]
       `G35-I-E-L-L-A-I-E-A-R-I-V-I-A-S——[c]

      [c]——S-A-I-V-I-R-A-E-I-A-L-L-E-I36
                                         `G35
         [n]——I-V-I-D-A-W-V-R-V-S-L-V34´
```

Figure 7. Model of the β-sheet portion of GvpA with all salt-bridges on one side of the sheet, in-register alignment of conserved aromatic and charged residues at the subunit interfaces, and correctly tilted strands, obtained by allowing inequivalent subunits with odd-numbered turns centered alternately on G35 and V34. The notation is as in Figure 2.

264

9.2 Results

We have obtained NMR data with line widths (full width at half height) of 70-110 and 60-80 Hz in the ^{13}C and ^{15}N dimensions. The ^{13}C line widths are comparable to those in the microcrystalline proteins BPTI[14] and ubiquitin[15] (in both cases 100 Hz). Like amyloid fibrils, such as □-synuclein[16, 17] and HET-s(218-289)[18], where ^{13}C lines widths in the 30-100 Hz range are observed, gas vesicles can be regarded as natural 2-dimensional protein crystals with a high degree of short-range order, producing narrow NMR lines. Still narrower line widths, such as those observed for the microcrystalline proteins GB1[19] and Crh[20] (15-30 Hz line widths) are required in order to resolve one-bond ^{13}C-^{13}C J-couplings. With the line widths that we observe in gas vesicles, spectroscopy at high field is required to resolve the peaks. Therefore, the data in this paper were acquired at ^{1}H Larmor frequencies of 700 to 900 MHz. Partial assignments have been obtained, and will be reported in a later paper.

Figure 8 shows the 44-47 ppm ^{13}C/105-117 ppm ^{15}N region of a NCACX correlation spectrum, where only Gly CA-N cross-peaks are expected. There are 3 glycine residues in the sequence, but 6 peaks are observed in this region. It is possible to assign three of these peaks sequentially to G22, G35, and G61. The remaining peaks are labeled A, B, and C. While it is not yet possible to assign peaks A, B, and C sequentially, it has been determined that peak C has a valine neighbor and therefore must be due to either G35 or G61.

Figure 9 shows N-C correlations for the unique A-S pair in the GvpA sequence. The left two panels show the N-CA and N-CB correlation of S49 in a NCACX spectrum, and the two right panels show the correlations of the S49 N to the preceding A48 CA and CB sites in a NCOCX spectrum. Since there is only one A-S pair in the sequence, all peaks must be assigned to this pair. In the ^{15}N dimension, it is clear that all the correlation signals are found at two distinct chemical shifts for S49, 1.8 ppm apart. In the ^{13}C dimension, two distinct chemical shifts are also observed for S49-CB, 0.5 ppm apart. On the other hand, at S49-CA, A48-CA, and A48-CB different chemical shifts are not distinguishable within the experimental error.
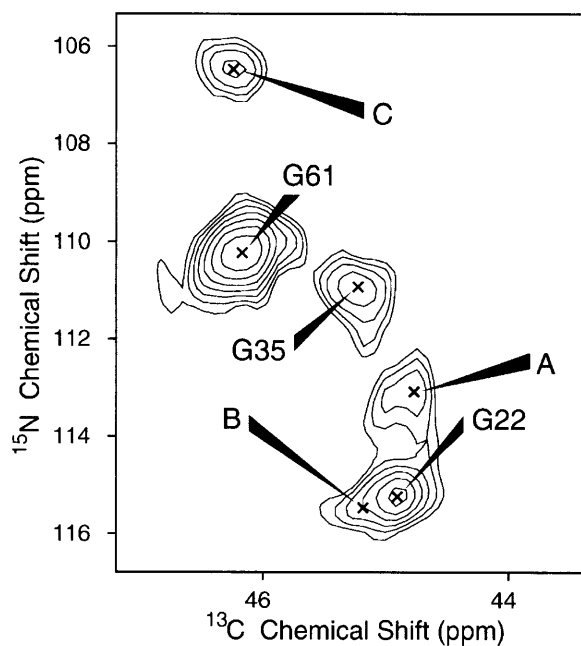
Figure 8. Glycine region of an NCACX correlation spectrum. The spectrum was acquired at 700 MHz. G22, G35, and G61 have been sequentially assigned, whereas the duplicated peaks A, B, and C have not. However, peak C must be due to either G35 or G61.
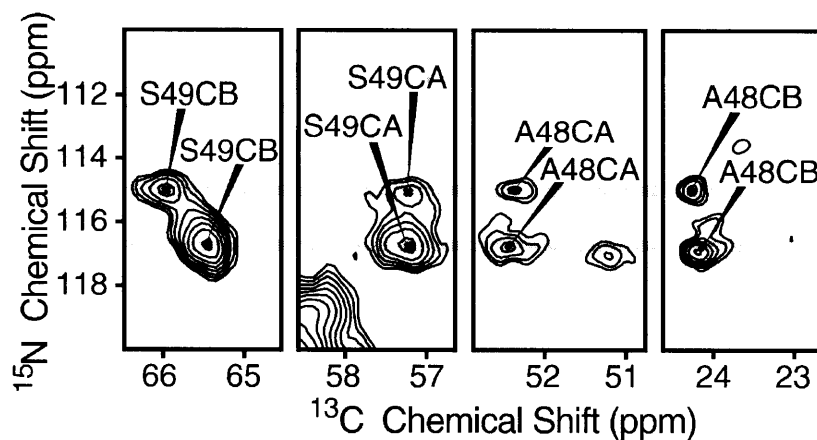
Figure 9. Correlations of S49 CA and CB and A48 CA and CB with two distinct S49 amide ^{15}N signals (grey lines). The two left panels are from a NCACX spectrum acquired at 700 MHz, while the two right panels are from a NCOCX spectrum acquired at 900 MHz.
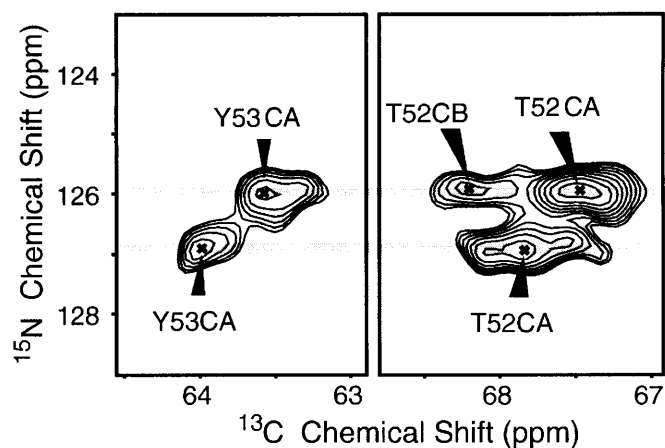
Figure 10. Correlations of Y53 CA and T52 CA and CB with two distinct Y53 amide ^{15}N signals (grey lines). The left panel is from a NCACX spectrum acquired at 700 MHz, while the right panel is from a NCOCX spectrum acquired at 900 MHz.
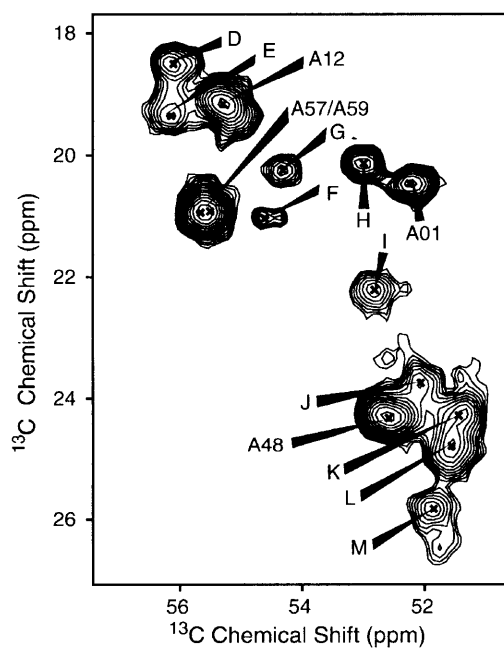
Figure 11. Alanine CA-CB region of an RFDR spectrum acquired at 750 MHz. At least 14 peaks are observed, although there are only 11 alanine residues in the sequence. The chemical shift dispersion clearly shows that both α-helix and β-sheet conformation is present in GvpA.

Figure 10 shows N-C correlations for the unique T-Y pair in the GvpA sequence. The left panel shows the N-CA correlation of Y53 in a NCACX spectrum, and the right panel shows the correlations of the Y53 N to the preceding T52 CA and CB sites in a NCOCX spectrum. Since there is only one T-Y pair in the sequence, all peaks must be assigned to this pair. In the ^{15}N dimension, it is clear that all the correlation signals are found at two distinct chemical shifts for Y53, 1.0 ppm apart. In the ^{13}C dimension, two distinct chemical shifts are also observed for Y53-CA, 0.4 ppm apart, and T52-CA, also 0.4 ppm apart.

Figure 11 shows the alanine CA-CB region of an RFDR spectrum. Although there are only 11 alanine residues in the sequence, at least 14 peaks are observed in this region, of which at least one represents two residues. From the distribution of the peaks, with secondary chemical shifts[21; 22] for both α-helix and β-sheet, it is clear that GvpA is a protein with mixed secondary structure. Although all three predicted α-helical alanine residues are assigned to one peak each, two peaks (D and E) still remain unassigned in the α-helix region. In addition, at least 10 peaks are observed for the 8 predicted coil and β-sheet residues, indicating duplication of peaks in various parts of the sequence. Finally, we note that the duplicated peaks in Figures 8-11 have comparable line widths in both the ^{13}C and ^{15}N dimensions. Hence, the extra peaks appear to represent two well ordered, but structurally different, protein fractions.

9.3 Discussion

Several sites in GvpA especially lend themselves to the identification of subunit inequivalence in gas vesicles. The glycine NCA correlations are found in a spectral region that does not contain other resonances, and there are only three glycine residues in this 70-residue protein. The unique AS and TY pairs are resolved in NCACX spectra. Of these residues, G22, G35, A48 and T52 are absolutely conserved across all organisms, while S49, Y53 and G61 are absolutely conserved among cyanobacteria, but not conserved more generally. All of these residues are located at or near predicted transitions in secondary structure (Figure 2).

As expected, the duplicated signals show larger variations in ^{15}N chemical shifts than in ^{13}C chemical shifts: the lone pair of nitrogen makes its shielding much more sensitive to changes in the local environment[23]. A sensitive reporter on the peptide backbone is clearly ideal for detecting secondary structure variations, although it may also reflect higher order variations.

The peaks in Figures 9 and 10 offer unambiguous evidence for structural variations in the GvpA subunits of *Anabaena flos-aquae*. Given the similar intensities within the pairs, it is likely that each peak in a duplicated pair comes from the same number of subunits, consistent with the asymmetric dimer model shown in Figure 7. The alternative of sample heterogeneity is implausible since x-ray diffraction shows just one characteristic distance in each of the three dimensions of the wall (i.e., the rib spacing, the wall-thickness, and the subunit repeat along a rib) and electron microscopy also shows just one rib spacing. Of course, electron microscopy also shows conical end-caps on otherwise cylindrical vesicles. However, Anabaena vesicles are long and the relatively small content of the end-caps is not expected to give signals of the strength of those seen in our duplicated peaks.

Figure 8 also shows peak duplication, with larger ^{15}N chemical shift variations, but it is more difficult to interpret. There are two options for assigning peak C because both G35 and G61 are preceded by a valine residue. However, given the strength of the peak already assigned to G61, it seems unlikely that any other peak belongs to G61 and reasonable to infer instead that the G61 signal is not split. With the alternative assignment of peak C to G35, the two G35 peaks are of comparable intensity and there is a difference in ^{15}N and ^{13}C chemical shifts of 4.5 and 1.0 ppm, respectively, indicating a significant difference in the two different subunit conformations at G35. In this scenario, the intensities of the remaining peaks A and B are most reasonably

assigned to G22. The chemical shift differences between peaks A, B, and G22 are small, and suggest correspondingly small conformation variations at this position.

The tentative glycine peak assignments obtained by the above consideration of peak intensities are in good agreement with the duplications expected for the asymmetric dimer model shown in Figure 7. G35 will be in significantly different local environments in alternating subunits depending on whether it is or is not at the center of the β-turn. In the above assignment, this is the glycine residue that displays the largest chemical shift change between the duplicated peaks. G22 is situated in a loop at the N-terminal end of the predicted β-sheet region and, like A48 at the C-terminal end, is expected to be less affected by the different turn positions in alternating subunits. G61, located on the C-terminal side of the C-terminal α-helix, would be affected still less. Thus, the asymmetric model shown in Figure 7 provides a clear rationale for the glycine peaks seen in Figure 8.

Taken together, the multiplicity of cross-peaks observed in well-resolved regions of heteronuclear 2D solid-state NMR spectra of intact gas vesicles support a model of the β-sheet portion of GvpA that achieves one completely hydrophobic face, complimentary charges and aromatic-aromatic interactions at subunit interfaces, and the stabilizing 36° strand tilt, all by a small folding variation in alternating GvpA subunits. That two conformations appear to contribute to function, places GvpA in a growing group of "metamorphic" proteins that adopt different folded states under native conditions[24].

9.4 Conclusions

Uniformly ^{13}C and ^{15}N labeled gas vesicles from *Anabaena flos-aquae* provide well-resolved solid-state NMR spectra at high magnetic fields. The sensitivity of the amide ^{15}N shifts provides a probe of subtle differences in protein structure. In gas vesicles we find that it distinguishes at least two different forms of the GvpA subunits. In particular, [15]N-^{13}C correlation spectra detect two different environments for the amides of S49 and Y53, and at least two different environments for the amides of G22 and G35. Thus the GvpA monomers are incorporated into gas vesicles in at least two different ways. The simplest explanation is that gas vesicles are formed by asymmetric dimers of GvpA. Such dimers can explain how the β-strands can be tilted at 36° relative to the vesicle axis while accommodating stabilizing interactions between highly conserved residues and forming a large hydrophobic surface suitable for inhibiting water condensation inside the vesicle. Future resonance assignments and structural constraints will test this interpretation and provide a larger context.

9.5 Methods

9.5.1 Sample Preparation

To label gas vesicles uniformly with ^{13}C and ^{15}N in *Anabaena flos-aquae* (Cambridge Collection of Algae and Protozoa (CCAP), Cambridge, UK, strain 1403/13f), the cells were grown under [13]CO_2 and [15]N_2. Floating cells were collected and lysed by osmotic shrinkage of the protoplasts using 0.7 M sucrose[25]. The released vesicles were isolated and washed by several rounds of centrifugally accelerated floatation at 100 x g in 5.0 mM NaCN, 10 mM potassium phosphate buffer at pH 8.0[26], followed by filtration on Millipore membrane filters with 0.65 and 0.45 □m pores[27]. No attempt was made to retain GvpC and any that might be retained would not disturb NMR experiments as a signal from less than 5 mol% of the sample is not detectable.

The isolated gas vesicles were collapsed by a sudden application of pressure to the plunger of a syringe holding a suspension of vesicles. Vesicle collapse was observed by clearing of the initial milky appearance. The collapsed vesicles were pelleted by 45 min of centrifugation at 158,420 x g, washed with 50 mM NaH_2PO_4, 1 mM NaN_3, pH 7.0, and then resuspended in the same buffer with 15% (w/w) D_8-glycerol (Cambridge Isotope Laboratories, Andover, MA). Glycerol prevents protein dehydration, and the deuteration prevents cross-polarization of the

natural abundance carbon. This suspension was centrifuged for 24 hours at 324,296 x g, and the resulting gel-like pellet was drained and packed into rotors using a tabletop centrifuge. 10.8 and 23.7 mg sample were packed into 2.5 and 3.2 mm rotors, respectively, and we estimate that less than half of this is protein. After closing the rotors, there was no dehydration over several months (as monitored by weight).

9.5.2 NMR Spectroscopy

CP/MAS (cross-polarization/magic-angle-spinning) NMR spectra were acquired on custom-designed spectrometers (courtesy of D. J. Ruben, Francis Bitter Magnet Laboratory, Massachusetts Institute of Technology) operating at 700 MHz and 750 MHz ^{1}H Larmor frequencies and a Bruker spectrometer (Billerica, MA) operating at 900 MHz ^{1}H Larmor frequency. The 700 MHz spectrometer was equipped with triple-resonance ^{1}H/^{13}C/^{15}N Varian-Chemagnetics (Palo Alto, CA) probe with a 3.2 mm stator, and the 750 and 900 MHz spectrometers were equipped with Bruker (Billerica, MA) probes with 2.5 mm stators. Samples were cooled with a stream of air during experiments, maintaining the exit gas temperature at –5 to 5°C. The MAS frequency was controlled to ±5 Hz using Bruker spinning speed controllers. All spectra were referenced to external DSS according to IUPAC convention[28], using the Ξ conversion factor[29].

The 2D NCOCX correlation spectrum was obtained at 900 MHz ^{1}H Larmor frequency, with 20 kHz MAS and 100 kHz TPPM ^{1}H decoupling[30], using ^{1}H-^{15}N CP, followed by specific DCP[31] for ^{15}N-^{13}CO polarization transfer, and PDSD[32] for subsequent ^{13}CO-^{13}CX transfer. 128 real and 128 imaginary points were acquired in the indirect dimension with a dwell time of 100 $\Box$s, and 2048 points were acquired in the direct dimension with a dwell time of 10 $\Box$s. 384 scans were acquired for each t$_1$ point.

The 2D NCACX correlation spectrum was obtained at 700 MHz ^{1}H Larmor frequency, with 12.5 kHz MAS and 83 kHz TPPM ^{1}H decoupling[30], using ^{1}H-^{15}N CP, followed by specific DCP[31] for ^{15}N-^{13}CA polarization transfer, and 2.56 ms RFDR[33; 34] with 20 kHz ^{13}C pulses for mixing from CA to other aliphatic ^{13}C nuclei. At this field, the typical 120-ppm difference between CO and CA resonances corresponds to 21 kHz and since the bandwidth of a 20 kHz pulse is less than 20 kHz, mixing occurs only between aliphatic carbons. This is advantageous in terms of sensitivity, as the magnetization is spread among fewer sites. 192 real and 192

imaginary points were acquired in the indirect dimension with a dwell time of 80 □s, and 1536 points were acquired in the direct dimension with a dwell time of 16 □s. 296 scans were acquired for each t_1 point.

The 2D RFDR[33; 34] spectrum was obtained at 750 MHz ^{1}H Larmor frequency, with 18.182 kHz MAS and 83 kHz XiX ^{1}H decoupling[35]. ^{13}C RFDR pulses of 40 kHz were used, and the ^{13}C-^{13}C mixing time was 3.52 ms. 512 real and 512 imaginary points were acquired in the indirect dimension with a dwell time of 28 μs, and 2048 points were acquired in the direct dimension with a dwell time of 12 μs. 112 scans were acquired for each t_1 point. The data were processed using NMRPipe[36], and examined using Sparky[37].

Acknowledgements

References

1.   Walsby, A. E. (1994). Gas vesicles. *Microbiol. Rev.* **58**, 94-144.

2.   Walsby, A. E. (1969). The permeability of blue-green algal gas-vacuole membranes to gas. *Proc. Royal Soc. London B.* **173**, 235-255.

3.   Ball, P. (2003). How to keep dry in water. *Nature* **423**, 25-26.

4.   Bowen, C. C. & Jensen, T. E. (1965). Blue-green algae: Fine structure of the gas vacuole. *Science* **147**, 1460-1462.

5.   Walsby, A. E. & Eichenberger, H. H. (1968). The fine structure of gas-vacuoles released from cells of the blue-green alga *Anabaena flos-aquae. Arch. Microbiol.* **60**, 76-83.

6.   Offner, S., Ziese, U., Wanner, G., Typke, D. & Pfeifer, F. (1998). Structural characteristics of halobacterial gas vesicles. *Microbiol.* **144**, 1331-1342.

7. Blaurock, A. E. & Walsby, A. E. (1976). Crystalline structure of the gas vesicle wall from *Anabaena flos-aquae. J. Mol. Biol.* **105**, 183-199.

8. McMaster, T. J., Miles, M. J. & Walsby, A. E. (1996). Direct observation of protein secondary structure in gas vesicles by atomic force microscopy. *Biophys. J.* **70**, 2432-2436.

9. Hayes, P. K., Walsby, A. E. & Walker, J. E. (1986). Complete amino acid sequence of cyanobacterial gas-vesicle protein indicates a 70-residue molecule that corresponds in size to the crystallographic unit cell. *Biochem. J.* **236**, 31-36.

10. Hayes, P. K., Buchholz, B. & Walsby, A. E. (1992). Gas vesicles are strengthened by the outer-surface protein, GvpC. *Arch. Microbiol.* **157**, 229-234.

11. Belenky, M., Meyers, R. & Herzfeld, J. (2004). Subunit structure of gas vesicles: A MALDI-TOF MS study. *Biophys. J.* **86**, 499-505.

12. McGuffin, L. J., Bryson, K. & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404-405.

13. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402.

14. McDermott, A., Polenova, T., Böckmann, A., Zilm, K. W., Paulsen, E. K., Martin, R. W. & Montelione, G. T. (2000). Partial NMR assignments for uniformly (^{13}C, ^{15}N)-enriched BPTI in the solid state. *J. Biomol. NMR* **16**, 209-219.

15. Igumenova, T. I., McDermott, A. E., Zilm, K. W., Martin, R. W., Paulson, E. K. & Wand, A. J. (2004). Assignments of carbon NMR resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* **126**, 6720-6727.

16. Heise, H., Hoyer, W., Becker, S., Andronesi, O. C., Riedel, D. & Baldus, M. (2005). Molecular-level secondary structure, polymorphism, and dynamics of full-length α-synuclein fibrils studied by solid-state NMR. *Proc. Natl. Acad. Sci. USA* **102**, 15871-15876.

17. Kloepper, K. D., Zhou, D. H., Li, Y., Winter, K. A., George, J. M. & Rienstra, C. M. (2007). Temperature-dependent sensitivity enhancement of solid-state NMR spectra of α-synuclein fibrils. *J. Biomol. NMR* **39**, 197-211.

18. Ritter, C., Maddelein, M.-L., Siemer, A. B., Lührs, T., Ernst, M., Meier, B. H., Saupe, S. J. & Riek, R. (2005). Correlation of structural elements and infectivity of the HET-s prion. *Nature* **435**, 844-848.

19. Franks, W. T., Zhou, D. H., Wylie, B. J., Money, B. G., Graesser, D. T., Frericks, H. L., Sahota, G. & Rienstra, C. M. (2005). Magic-angle spinning solid-state NMR spectroscopy of the β1 immunoglobulin binding domain of protein G (GB1): ^{15}N and ^{13}C chemical shift assignments and conformational analysis. *J. Am. Chem. Soc.* **127**, 12291-12305.

20. Böckmann, A., Lange, A., Galinier, A., Luca, S., Giraud, N., Juy, M., Heise, H., Montserret, R., Penin, F. & Baldus, M. (2003). Solid state NMR sequential resonance assignments and conformational analysis of the 2 × 10.4 kDa dimeric form of the *Bacillus subtilis* protein Crh. *J. Biomol. NMR* **27**, 323-339.

21. Wishart, D. S. & Sykes, B. D. (1994). Chemical shifts as a tool for structure determination. *Methods Enzymol.* **239**, 363-392.

22. Wishart, D. S. & Case, D. A. (2001). Use of chemical shifts in macromolecular structure determination. *Methods Enzymol.* **338**, 3-34.

23. Witanowski, M., Stefaniak, L. & Webb, G. A. (1993). Nitrogen NMR spectroscopy. *Ann. Rep. NMR Spectrosc.* **25**, 1-480.

24. Murzin, A. G. (2008). Metamorphic proteins. *Science* **320**, 1725-1726.

25. Walsby, A. E. (1972). Structure and function of gas vacuoles. *Bacteriol. Rev.* **36**, 1-32.

26. Dunton, P. G., Mawby, W. J., Shaw, V. A. & Walsby, A. E. (2006). Analysis of tryptic digests indicates regions of GvpC that bind to gas vesicles of *Anabaena flos-aquae*. *Microbiol.* **152**, 1661-1669.

27. Walsby, A. E. (1974). The isolation of gas vesicles from blue-green algae. *Methods Enzymol.* **31**, 678-686.

28. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L. & Sykes, B. D. (1995). ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**, 135-140.

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

30. Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V. & Griffin, R. G. (1995). Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* **103**, 6951-6958.

31. Schaefer, J., Stejskal, E. O., Garbow, J. R. & McKay, R. A. (1984). Quantitative determination of the concentrations of ^{13}C-^{15}N chemical bonds by double cross-polarization NMR. *J. Magn. Reson.* **59**, 150-156.

32. Suter, D. & Ernst, R. R. (1985). Spin diffusion in resolved solid-state NMR spectra. *Phys. Rev. B.* **32**, 5608-5627.

33. Bennett, A. E., Griffin, R. G., Ok, J. H. & Vega, S. (1992). Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys.* **96**, 8624.

34. Bennett, A. E., Rienstra, C. M., Griffiths, J. M., Zhen, W., Lansbury, P. T. & Griffin, R. G. (1998). Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys.* **108**, 9463.

35. Detken, A., Hardy, E. H., Ernst, M. & Meier, B. H. (2002). Simple and efficient decoupling in magic-angle spinning solid-state NMR: the XiX scheme. *Chem. Phys. Lett.* **356**, 298-304.

36. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293.

37. Goddard, T. D. & Kneller, D. G. SPARKY 3, University of California, San Francisco.

Chapter 10. Molecular Conformation of GvpA, the Building Block of Gas Vesicles, Characterized by MAS NMR Spectroscopy

Summary

Gas vesicles are gas-filled buoyancy organelles with walls that consist almost exclusively of gas vesicle protein A (GvpA). Intact, collapsed gas vesicles from the cyanobacterium *Anabaena flos-aquae* were studied by solid-state NMR spectroscopy, and most of the GvpA sequence was assigned. Chemical shift analysis indicates a coil-α-β-β-α-coil peptide backbone, consistent with secondary structure prediction algorithms, and complimentary information about mobility and solvent exposure yields a picture of the overall topology of the vesicle subunit that is consistent with its role in stabilizing an air-water interface.

10.1 Introduction

Gas vesicles are buoyancy organelles that are found in a wide range of microorganisms (50). The hollow protein shells (51) are assembled and disassembled *in vivo*, allowing the cells to change their overall density and thereby adjust their position in the water column in response to light intensity (50) and aeration (49). Of the 14 genes in the gas vesicle operon of *Halobacterium halobium*, 8 are necessary and sufficient to code gas vesicle expression, and 6 of these have homologues in *Anabaena flos-aquae* (37). The gas vesicle operon has most likely spread by lateral gene transfer (11), such that the ability to form gas vesicles occurs more widely even than photosynthesis (50).

Electron microscopy of intact vesicles (12, 52) shows shapes ranging from regular cylinders with conical end caps to irregular acorn-like bodies. Vesicles are typically 5000 Å long and 750 Å wide, but proportions vary considerably across species, and even within a cell. The vesicles are bipolar with an apparent insertion seam. To either side of the insertion seam, monomers of the ~ 7 kDa gas vesicle protein A, GvpA, are assembled into a low-pitch helix (38). GvpA is highly hydrophobic and negatively charged at neutral pH (24), with a highly conserved sequence that PsiPred (32) and other algorithms predict to contain both α-helix and β-strand secondary structure segments. Substantial anti-parallel β-sheet is indicated by FT-IR spectroscopy (Mark Braiman, personal communication) and X-ray diffraction (10). MALDI-TOF studies of GvpA in intact *Anabaena flos-aquae* vesicles have shown that only five residues can be cleaved from the C-terminus by using carboxypeptidase Y, none of the nine D-X and E-X bonds in GvpA, situated throughout the sequence from E03-K04 to E58-A59 are accessible to endoproteinase GluC, and only the N-terminal K04-T05 bond and not the other R-X and K-X bonds situated from R17-I18 to K55-Y56 are accessible to trypsin (6) except for a rarely observed cleavage of the R17-I18 bond (19).

Among the other gas vesicle proteins, GvpC is the best characterized. It is about three times larger than GvpA (23) and adheres to the outside of the shell formed by GvpA, strengthening the overall structure (22). In addition to GvpA and GvpC, the proteins GvpF, GvpG, GvpJ, GvpL, and GvpM have been found in *Halobacterium salinarum* gas vesicles in amounts only detectable by immunoblotting (i.e., well under 1% of overall gas vesicle protein). GvpJ and GvpM have partial sequence similarity with GvpA and are hypothesized to be involved in cap formation, while GvpF, GvpL, and GvpG are hypothesized to be involved in nucleation of

the gas vesicle assembly process at the tips of the caps (43). In addition, the *gvpo* and *gvpk* genes are required for synthesis of gas vesicles, although GvpO and GvpK proteins have not been detected in gas vesicles, indicating that these proteins may act as chaperones in the gas vesicle assembly process.

The low-pitch helix of GvpA subunits gives rise to regularly spaced ribs running nearly perpendicular to the vesicle long axis, and it has been shown by X-ray diffraction from intact, partially aligned vesicles (10) and by atomic force microscopy (33) that the distance between neighboring ribs is 45.7 Å, and that the ribs incorporate β-strands that are tilted at an angle of 36° from the vesicle long axis. This orients the H-bonds at a 54° angle relative to the vesicle axis, which is the ideal angle for mechanical stability (50) as the strength is the same in the length and width directions of the vesicle.

The hollow gas vesicles are filled with air of atmospheric composition, in equilibrium with gas molecules dissolved in the aqueous phase. Although the gas vesicle wall is permeable to H_2, N_2, O_2, Ar, CO, and CO_2, and even a molecule as large as perfluorocyclo-butane, C_4F_8 (6.3 Å diameter), water vapor does not condense inside. This is believed to be due to a highly hydrophobic and tightly curved inner face without suitable nucleation sites (48, 50). However, the details of the formation of the air-water interface are not known.

A recent solid-state NMR study (44) has shown duplication of the resonances for G22, G35, S49, T52, Y53, and several alanine residues of GvpA in intact, collapsed vesicles from *Anabaena flos-aquae*. Thus, GvpA seems to be a member of the growing group of metamorphic proteins (36). The duplicated resonances are interpreted in terms of asymmetric GvpA dimers that can explain how the strands of an anti-parallel β-sheet, can be tilted at an angle of 36° relative to the gas vesicle axis, while pairing oppositely charged side chains in salt bridges on one side of the β-sheet and leaving the other side fully hydrophobic.

In this paper, we describe magic-angle spinning NMR studies aimed at characterizing the secondary structure of GvpA in intact, collapsed gas vesicles. Resonance assignments have been obtained for the majority of the GvpA sequence, allowing us to elucidate previously uncharacterized features of gas vesicle structure at a site-specific level including backbone conformation, mobility, and solvent accessibility.

10.2 Materials and Methods

10.2.1 Sample Preparation

Anabaena flos-aquae cells were grown under a $^{13}CO_2$ and $^{15}N_2$ atmosphere, and the gas vesicles were isolated, stripped of GvpC, collapsed, and prepared for magic-angle spinning experiments as described previously (44). Approximately 10 and 24 mg of sample were packed into 2.5 and 3.2 mm rotors, respectively. A ^{2}H-exchanged sample was prepared by multiple rounds of centrifugally accelerated flotation of inflated vesicles in 99% ^2H_2O (Cambridge Isotope Laboratories, Andover, MA). Vesicles were then collapsed and approximately 28 mg of sample was packed in a 3.2 mm rotor. The solvent used included 15% $^2H_8$-glycerol (wt/wt) to prevent dehydration of the sample.

10.2.2 NMR Spectroscopy

Solid-state NMR experiments were carried out using custom-designed spectrometers (D. J. Ruben, Francis Bitter Magnet Laboratory, Massachusetts Institute of Technology) operating at 700 MHz and 750 MHz ^{1}H Larmor frequencies, and a Bruker spectrometer (Billerica, MA), operating at 900 MHz ^{1}H Larmor frequency. The 750 and 900 MHz spectrometers were equipped with triple-resonance $^1H/^{13}C/^{15}N$ Bruker probes with 2.5 mm stators, and the 700 MHz spectrometer was equipped with a triple-resonance $^1H/^{13}C/^{15}N$ Varian-Chemagnetics (Palo Alto, CA) probe with a 3.2 mm stator. The magic angle spinning (MAS) frequency was controlled to ± 5 Hz using Bruker MAS controllers, and the samples were cooled with a stream of cold N_2 gas during experiments, maintaining exit gas temperatures of -40 to 5 °C. Sample temperatures are estimated to be approximately 10 °C higher than the exit gas temperature, depending on the MAS frequency and the intensity of rf irradiation. All spectra were referenced to external DSS (34, 53).

One-dimensional spectra were acquired at 700 MHz ^{1}H Larmor frequency, 12.5 MHz MAS, and approximate sample temperatures of +3 and -30°C. The cross-polarization (39) (CP) and direct polarization spectra are the average of 64 scans each, while the INEPT (35) spectra are the average of 256 scans each. Homonuclear $^{13}C-^{13}C$ and heteronuclear $^{15}N-^{13}C-^{13}C$ correlation spectra were acquired using various two-dimensional (2D) dipolar recoupling experiments. A $^{13}C-^{13}C$ RFDR(5, 7, 9) spectrum was recorded at 750 MHz ^{1}H Larmor frequency and 18 kHz MAS. A 2D NCOCX spectrum was obtained at 900 MHz ^{1}H Larmor frequency and

20 kHz MAS, using specific CP (3, 41) for ^{15}N-^{13}CO polarization transfer and 40 ms of PDSD mixing (45, 46) for subsequent ^{13}CO-^{13}CX transfer. A NCACX spectrum was obtained at 900 MHz ^{1}H Larmor frequency and 20 kHz MAS, using specific CP for ^{15}N-^{13}CA polarization transfer followed by 6.4 ms of homonuclear mixing via RFDR. Both heteronuclear experiments were the recorded with 128 real and 128 imaginary points in the indirect dimension and 2048 points in the direct dimension, with dwell times of 100 µs and 12 µs respectively, and 83 kHz TPPM decoupling (8) during the chemical shift evolution and detection periods. The spectra are the average of 384 transients per t_1 point. A 2D INEPT-TOBSY spectrum was acquired at 750 MHz ^{1}H Larmor frequency and 8.333 kHz MAS. In this experiment, ^{1}H-^{13}C magnetization transfer via INEPT (35) was used to excite mobile sites, followed by 5.76 ms of TOBSY (1, 2) ^{13}C-^{13}C mixing using the P9$_3^1$ condition (21) with a 50 kHz ^{13}C rf field. This spectrum was acquired with 256 real and 256 imaginary points in the indirect dimension and 2048 points in the direct dimension, with dwell times of 60 µs and 12 µs respectively, and 50 kHz TPPM decoupling during the chemical shift evolution and detection periods. The spectra are the average of 224 scans for each t_1 point. All data sets were processed with the NMRPipe package (18) and analyzed using Sparky (20).

10.3 Results

10.3.1 Spectral Quality

Intact, collapsed gas vesicles are highly ordered molecular assemblies and present highly resolved magic-angle spinning NMR spectra with average line widths of 0.5 and 0.8 ppm for ^{13}C and ^{15}N resonances, respectively (equivalent to full width at half height of 80 and 70 Hz, respectively, at 700 MHz ^{1}H Larmor frequency). These line widths are comparable to those observed for such microcrystalline proteins as BTPI (31) and ubiquitin (26), and such amyloid fibrils as those formed by α-synuclein (25, 27), PI3-SH3(4), and HET-s(218-289) (40), but they do not allow one-bond ^{13}C-^{13}C J-couplings to be resolved. Fig. 1 shows cross-polarization (CP) and direct-polarization ^{13}C MAS spectra acquired at 3°C and –30°C. A certain degree of line broadening is observed at the lower temperature, indicating some freezing of dynamic disorder, in particular for side chain ^{13}C resonances. The CP signal enhancement observed is well below the ~2.5 factor observed for rigid peptides and proteins and reveals the highly mobile nature of GvpA. This dynamic nature also manifested itself in the low intensity of many of the cross peaks in dipolar correlation spectra, despite their narrow line widths, as can be observed in the ^{13}C-^{13}C spectrum of Fig. 2. However, as shown below, only a small number of residues exhibited solution-like mobility, and we therefore relied mainly on dipolar recoupling experiments for establishing intra-residue and sequential correlations.

10.3.2 Sequential Resonance Assignments

Narrow ^{15}N and ^{13}C line widths allowed us to obtain site-specific sequential resonance assignments via 2D dipolar correlation experiments. In particular, employing CC (Fig. 2), NCACX (Fig. 3), and NCOCX (not shown) 2D correlation spectra we were able to assign ^{15}N and ^{13}C resonances for 57 of the 70 residues, or 81% of the amino acid sequence, of GvpA. A list of assigned ^{15}N and ^{13}C chemical shifts is provided in the supplementary material (Table S1). As described and discussed earlier (44), certain resonances are duplicated, in particular those of several Gly and Ala residues along with S49, T52, Y53, and Y56, and sequential assignment of duplicated resonances was often impeded by the relatively low intensity of these signals.
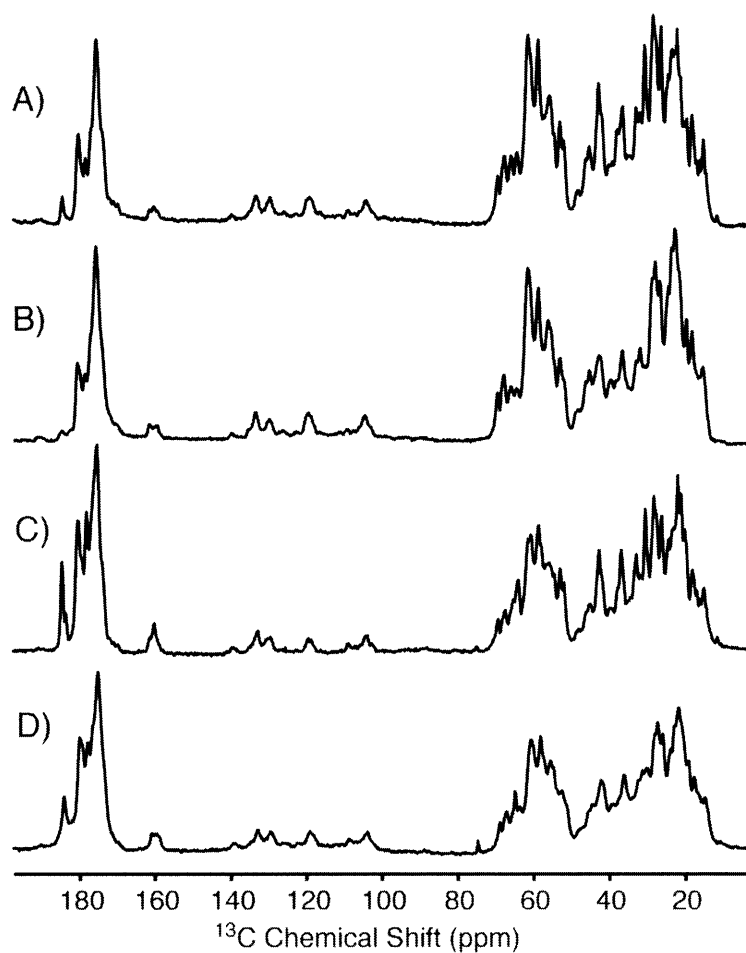
Figure 1. ^{13}C MAS spectra acquired at 700 MHz ^{1}H Larmor frequency and 12.5 kHz MAS: CP (*A* & *B*) and direct polarization (*C* & *D*) at 3°C (*A* & *C*) and –30°C (*B* & *D*). Each spectrum is an average of 64 scans, recorded with recycle delays of 3 s and 30 s in CP and DP experiments, respectively. A glycerol resonance is visible at ~75 ppm in (D).
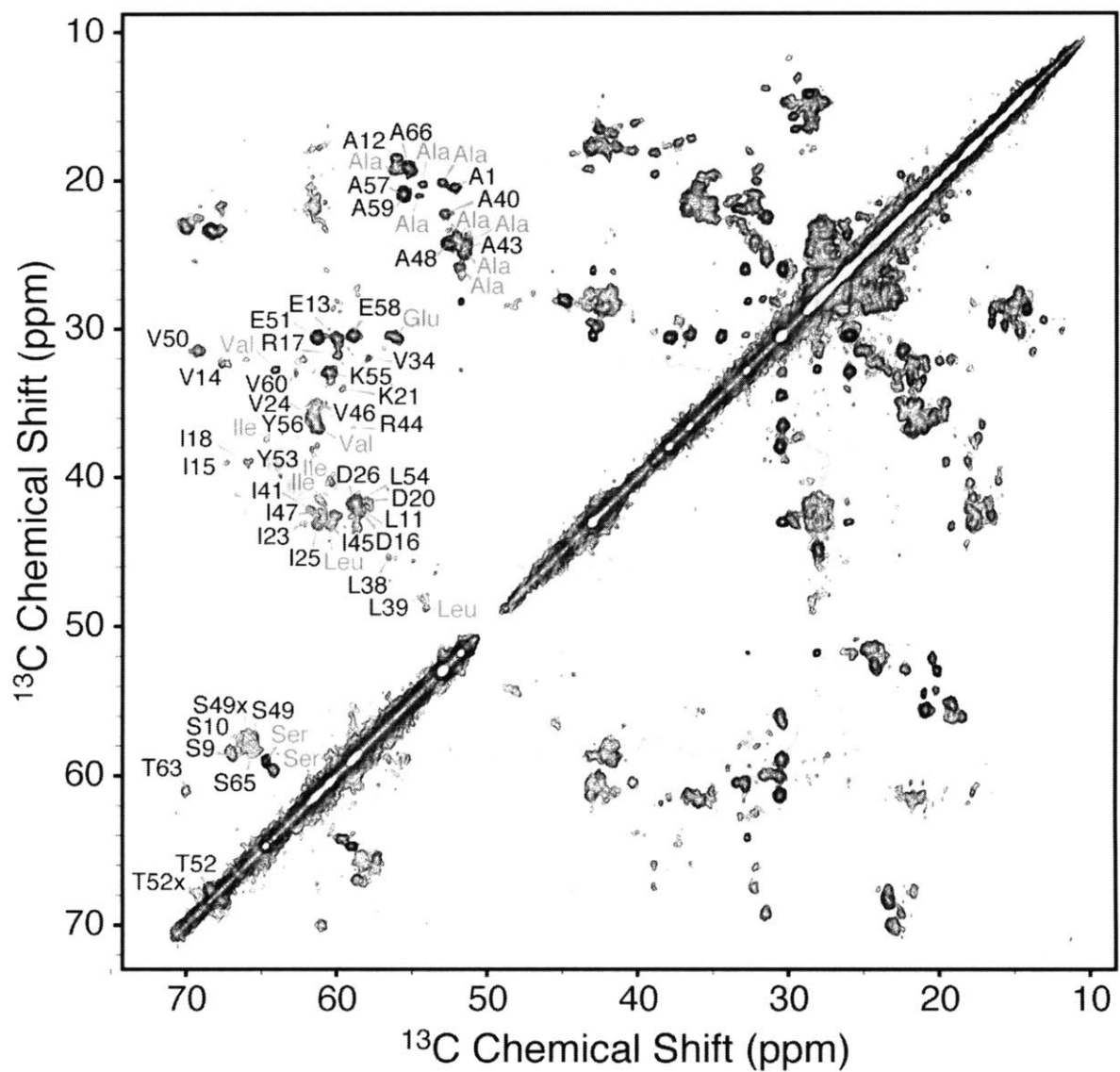
Figure 2. ^{13}C-^{13}C correlation spectrum acquired at 750 MHz ^1H Larmor frequency and 18 kHz MAS frequency with 3.52 ms of RFDR mixing. Cross-peaks corresponding to CA-CB correlations for several assigned residues are indicated.

The N-terminal residue, A1, can be identified due to its unique ^{15}N chemical shift, and sequential correlations can be established to V2 and E3. Residues K4-S8 have not been assigned because they present weak signals that severely limit inter-residue correlations. The segment S9-W28 has been assigned and gives rise to backbone and side-chain correlations that are generally of medium intensity. The following residues, V29-L33, again present only weak correlation signals and are not assigned. The segment V34-V66 has been assigned sequentially, and gives rise to backbone and side-chain signals of medium intensity for the segment V34-I47 and strong intensity for the segment A48-V66. The remaining four C-terminal residues have not been assigned, with the exception of P69, which presents weak signals that can be assigned unambiguously due to the characteristic downfield ^{15}N chemical shift of proline. Several side-chain resonances in the main assigned segments (S09-W28 and V34-V66) are not observed due to line broadening dynamics or remain unresolved in highly repetitive spin systems (e.g., 11 Ala, 6 Glu, 8 Ile, and 10 Val residues in the 70-residue protein).

While backbone ^{13}C and ^{15}N resonances tend to present significant degeneracy, it was possible to resolve assignment ambiguities in many instances by relying on side-chain correlations in NCACX and NCOCX spectra. Resonance intensities vary widely across the GvpA sequence, such that certain parts of the amino acid sequence, especially the C-terminal third, are predominant in the ^{13}C-^{13}C and ^{15}N-^{13}C correlation spectra. The segments S9-W28 and V34-I47 generally have weak or medium intensity correlations, but especially weak resonances arise from residues such as I15, L19, I36, E37, L38, L39, E42, and R44. The segment A48-V66 generally gives rise to correlations with high intensities, except for the duplicated residues Y53 and Y56, the duplicate form of T52, and residues V60, L62, and T63, which give rise to resonances of moderate intensity.

The side chains beyond CB of several residues in the GvpA sequence are observed in full (in the combined spectra, not necessarily in just one type of spectrum). These residues are L11, V14, I15, I18, I23, V24, I25, L38, L39, V46, I47, V50, T52, V60, L62, and T63. On the other hand, incomplete side-chains are observed for I36, V34, I41, I45, and L54. For almost all the acidic residues (Glu and Asp), carboxyl correlations are missing in most spectra. Only one Arg and one Lys side-chain are observed in full, among the three of each in the sequence. The Y53 side-chain readily shows up in full in RFDR and spin diffusion spectra with short mixing times, while longer mixing times are required for the entire Y56 side chain to be observed. This

dissimilarity is likely the result of different ring dynamics in the two Tyr residues. The ring system of W28 has not been observed in full; only a CB-CG correlation has been observed in ^{13}C-^{13}C spectra with long mixing times.

After assignment to the degree possible, two Ser, one Val, one Glu, and eight Ala CA-CB correlations in RFDR spectra remain unassigned. These Ser, Val and Glu correlations could come from unassigned parts of the protein, but only one Ala remains unassigned in the sequence. Therefore, the rest of the Ala CA-CB resonances are likely to be peak duplications. In the CA regions of NCACX and NCOCX spectra, approximately another 20 residues could be assigned to the yet unassigned peaks. This number is again higher than the actual number of unassigned residues, indicating duplications that are more extensive than the ones that have been assigned so far.

Figure 3. $^{13}C\alpha$ region of a NCACX spectrum acquired at 900 MHz ^{1}H Larmor frequency and 20.0 kHz MAS frequency. The experiment used ^{15}N-^{13}C cross-polarization followed by ^{13}C-^{13}C RFDR mixing. Labels correspond to N-CA correlations, unless otherwise noted.

286

10.3.3 Secondary Structure Analysis

Values of the CA–CB (28) secondary chemical shifts (54, 55) are shown in Fig. 4A, and display strong agreement with the PSIPRED secondary structure prediction shown in Fig. 4B. A TALOS (16) chemical shift analysis of the assigned residues gives the secondary structure pattern shown in Fig. 4C, which largely agrees with the secondary chemical shift result and with the PSIPRED prediction. TALOS results are only shown for one form for the residues that display peak duplications, as the TALOS results for the two forms of GvpA do not differ by more than 3-4°, which is within the standard deviation of the TALOS output.

Figure 4. Secondary structure analysis. (A) CA–CB secondary chemical shifts (gray indicates unassigned sequence); (B) PSIPRED prediction (α-helical, β-sheet, and coil regions); (C) TALOS results (notation as in B but with unassigned/non-calculated segments in gray).

The N-terminal α-helix is predicted by PSIPRED to extend from L11 to I18, and both secondary chemical shifts and the TALOS calculation indicate the beginning of the helix at L11. However, our chemical shift analysis indicates that the helix may continue until D20. It is worth noting that the T5-R17 fragment that has been observed in digestion studies (19) corresponds to cleavage at this transition between the N-terminal α-helix and the N-terminal β-sheet that follows. The latter is predicted by PSIPRED to extend from V24 to V34, and the secondary chemical shifts and TALOS results indicate that it begins at I23 and continues for a few residues. Resonance assignments were not obtained for the second half of this sequence, but there is a secondary chemical shift for V34 that has a β-sheet value. According to our chemical shift

287

analysis the C-terminal β-strand extends from I36 to S49, while the segment V50-A59 adopts an α-helical conformation, in close agreement with PSIPRED. The remainder of the sequence is predicted to adopt coil conformation, and neither secondary chemical shifts nor TALOS indicate any significant stretches of defined secondary structure for this part of the protein. TALOS results show non-standard angles for five residues: K21, G22, G35, V60, and G61. Because of their large conformational flexibility, it is not surprising that atypical ϕ and ψ values are observed for Gly residues, especially as they are found in intermediate regions. The unusual torsion angles of Gly neighbors K21 and V60 are also consistent with the Gly residues being found in turns between the secondary structure elements. The Ramachandran plot in Fig. 5 shows that the assigned residues reside in the allowed α-helix and β-sheet regions, except for the few mentioned above.
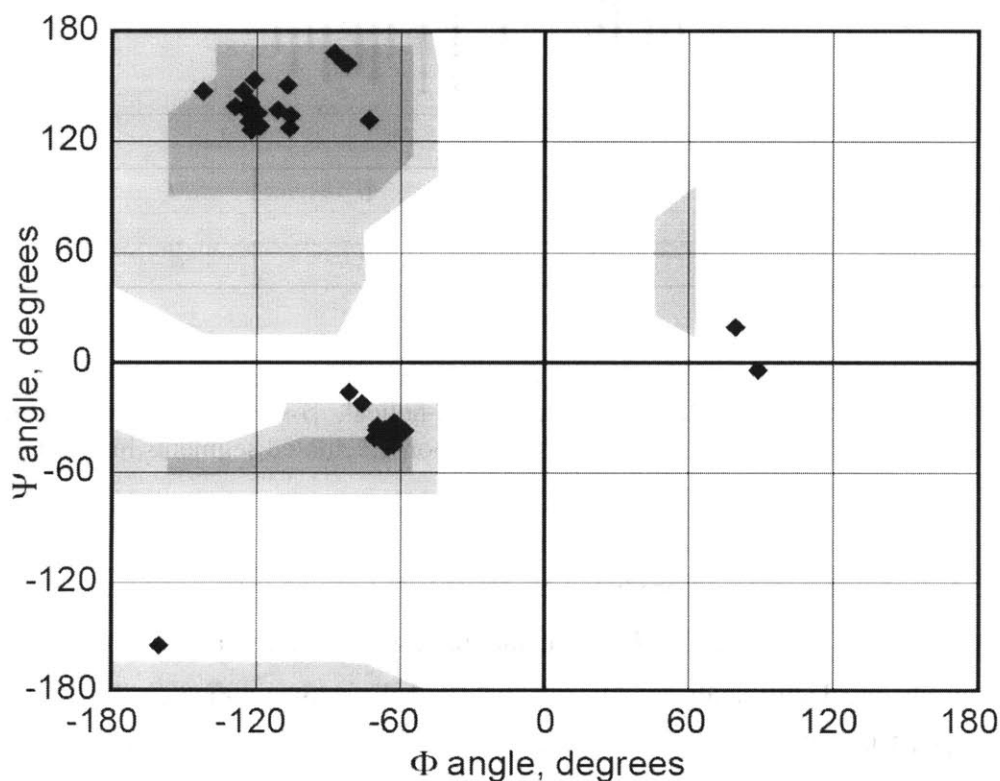
Figure 5. Ramachandran plot for GvpA using dihedral angles from TALOS. Gray areas indicate allowed regions for non-Glycine residues (dark gray indicates core areas).

10.3.4 Amphipathic α-helices

Fig. 6 shows nearest neighbor corrected (13, 54) secondary ^{15}N backbone chemical shift values for the α-helical segments of GvpA. As all residues in the figure are α-helical, the differences in ^{15}N chemical shifts are mainly governed by differences in NH-CO hydrogen bond length along the helix, where large secondary chemical shifts indicate short NH-CO hydrogen bonds (17, 55). The plots display variations of the secondary chemical shifts, with a period of roughly 3-4 residues, corresponding to an α-helical turn. A similar periodicity is found in amide proton resonances of curved amphipathic helices, where backbone hydrogen bonds are shorter on the hydrophobic side and longer on the hydrophilic side.(54, 56) Since amide ^{15}N chemical shifts exhibit trends with respect to hydrogen-bond length that are similar to those of amide ^{1}H shifts, we may speculate that the periodicity observed in the ^{15}N shifts for the N-teminal helix (left side of Fig. 6) might be the result of curvature in the helix, while the trend in the C-terminal helix is less defined. These observations are consistent with the amphipathic residue patterns seen in helical wheel renderings (see Fig. S1 in the supplementary material).
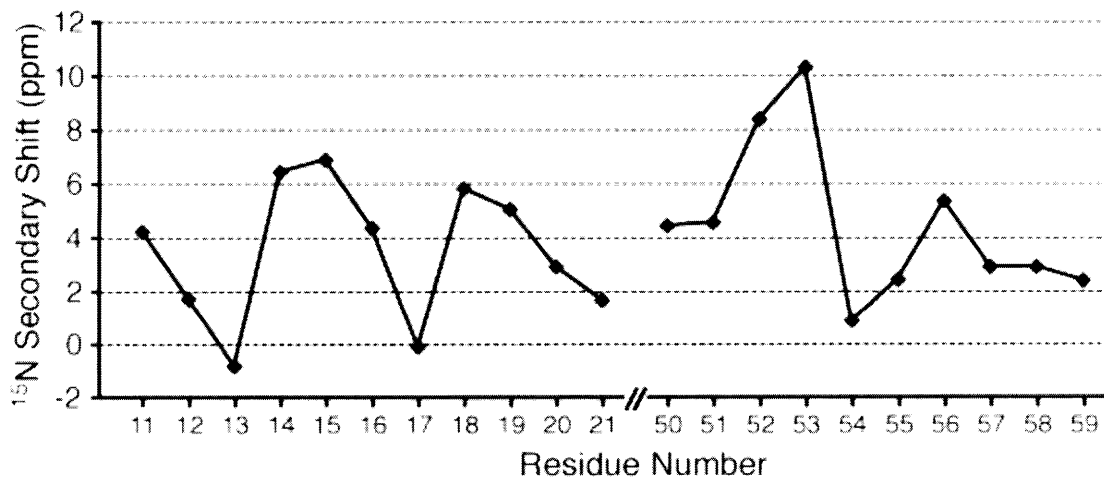
Figure 6. Nearest-neighbor corrected ^{15}N secondary shifts, showing periodic patterns characteristic of curved helices.

289

10.3.5 Molecular Mobility

Fig. 7 shows ^{1}H-^{13}C INEPT spectra acquired at 3°C (A) and −30°C (B). In the solid-state, the INEPT sequence excites those ^{13}C sites with sufficient flexibility to average out the dipolar interactions to the point that they do not interfere with the much weaker J-couplings. Typically, only mobile side-chains and highly flexible backbone regions give rise to INEPT signals. As Fig. 6 shows, a few signals are observed at 3°C, in most cases corresponding to single sites. About eight CA sites are observed, indicating that at this temperature most of the backbone is fairly rigid. The number of side-chains corresponds roughly to the number of backbone sites. At − 30°C, there is much less mobility, and no CA sites are observed, whereas some side chain sites are still observed, mostly methyl groups. The spectra in Figs. 1 and 7 together indicate that GvpA in intact collapsed vesicles is overall moderately dynamic but presents only a few sites with solution-like mobility.
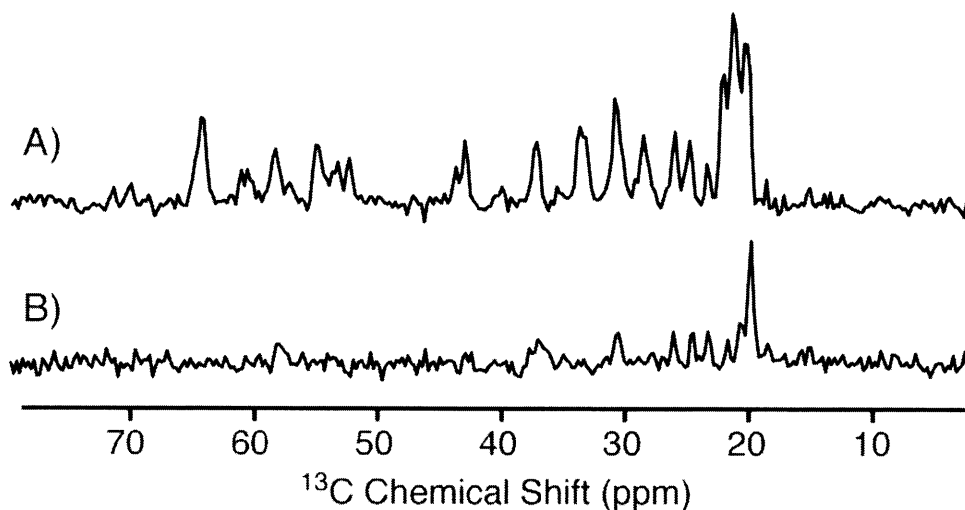
Figure 7. ^{1}H-^{13}C INEPT spectra acquired at 700 MHz ^{1}H Larmor frequency and 12.5 kHz MAS at (A) 3°C and (B) −30°C. Each spectrum is the average of 256 scans.
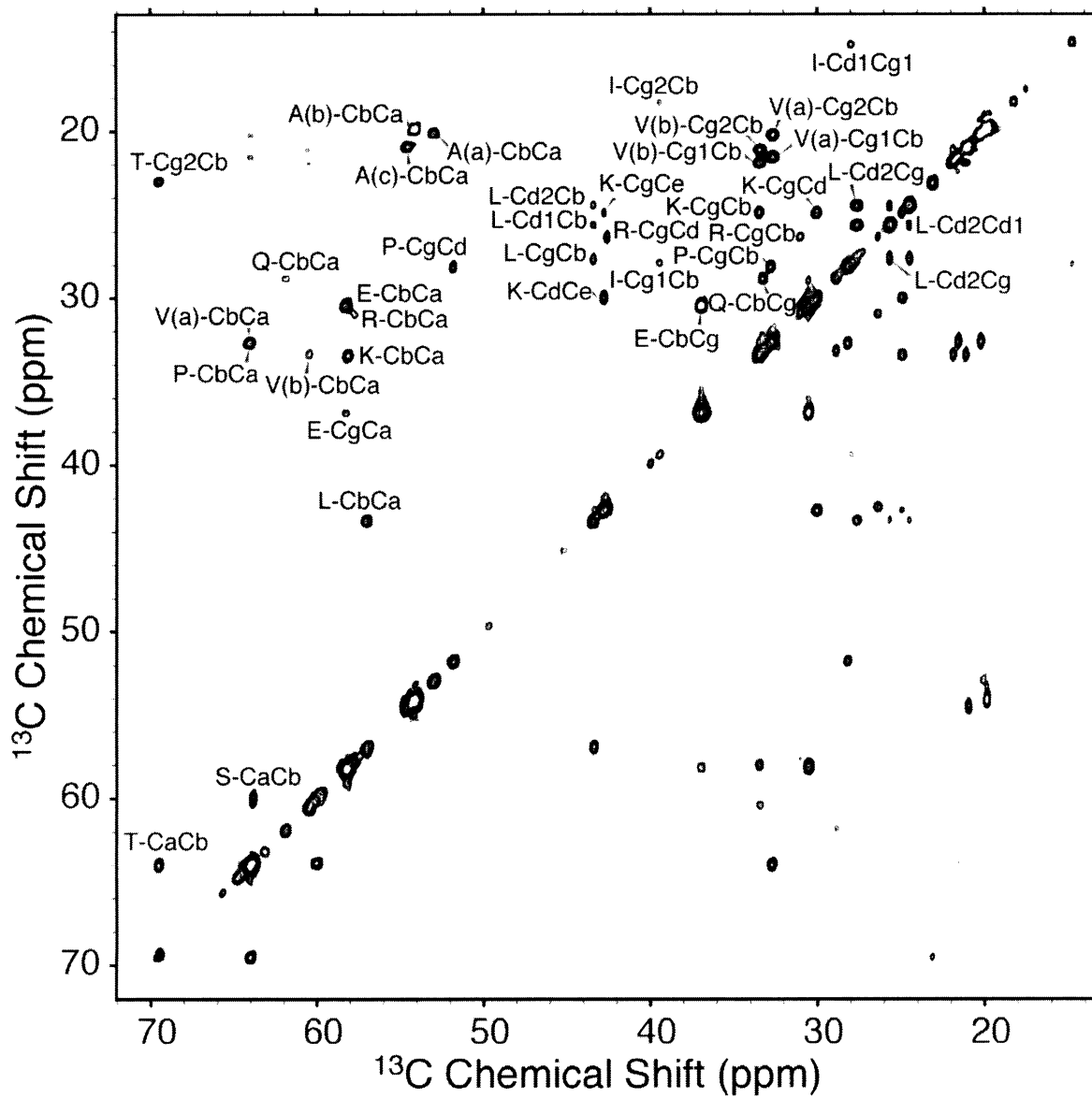
Figure 8. Magic-angle spinning ^{13}C-^{13}C INEPT-TOBSY spectrum recorded at 750 MHz ^1H Larmor frequency and 8.333 kHz MAS frequency.

To assign the highly mobile residues, a 2D ^{13}C-^{13}C INEPT-TOBSY spectrum was acquired (Fig. 8). In this experiment, after excitation of mobile ^{13}C sites, TOBSY mixing is applied to achieve homonuclear polarization transfer mediated by ^{13}C-^{13}C J-couplings. From the nearly complete ^{13}C spin systems in these data, we were able to identify three Ala, two Val and one Pro set of resonances that are also observed, at least weakly, in dipolar correlation spectra. In addition, resonance sets were detected for Ser, Thr, Leu, Ile, Glu, Gln, Lys, and Arg. None of the observed resonances could be assigned sequentially, but the Thr resonances are likely to correspond to T5, since T52 and T63 are already assigned in dipolar correlation spectra, while the Pro resonances are those of the sole P69.

10.3.6 Solvent accessibility

In order to gain information about the GvpA subunit topology that forms the gas vesicle wall, we acquired water-edited correlation spectra and studied a ^{2}H$_2$O-exchanged sample. Water-edited spectra can be used to identify solvent-accessible residues by delaying ^{1}H magnetization transfer to ^{13}C until the intrinsic ^{1}H polarization in the protein is relaxed by relatively strong dipolar interactions. However, the spectra we recorded in this fashion show that considerable ^{1}H polarization remains in the protein after the filter period, possibly due to motion of certain segments of the polypeptide chain. Therefore, water-edited 1D and 2D ^{13}C-^{13}C correlation experiments yielded ambiguous information.

On the other hand, a sample in ^{2}H$_2$O allowed the identification of exchangeable amide sites by analyzing ^{15}N-^{13}C DCP spectra recorded with different contact times for the initial ^{1}H-^{15}N CP step. Short ^{1}H-^{15}N CP periods are sufficient to transfer magnetization only to ^{15}N sites with directly bonded protons, while longer CP times allow polarization to build up at neighboring sites. Fig. 9 shows NCA spectra of vesicles collapsed in ^{2}H$_2$O recorded with ^{1}H-^{15}N CP contact times of 0.5 ms (A) and 8.0 ms (B). Most NCA cross-peaks appear at both CP times, but certain peaks that are not observed at 0.5 ms CP time are observed with an 8 ms contact time, including: A1, V2, T52, V60, T63, and several unassigned NCA peaks, some of which appear to correspond to residues with doubled resonances. Similarly, while Lys Nz-Ce correlations are observed at 8 ms CP time, they are not seen at 0.5 ms CP time (not shown), and thus we may conclude that all three Lys side-chains are solvent accessible. On the other hand, since the majority of NCA correlations are observed at 0.5 ms (Fig. 9B), most amide sites do not exchange

with ^2H_2O, either due to participation in secondary structure hydrogen bonding, or because they are buried within the solvent inaccessible protein core.
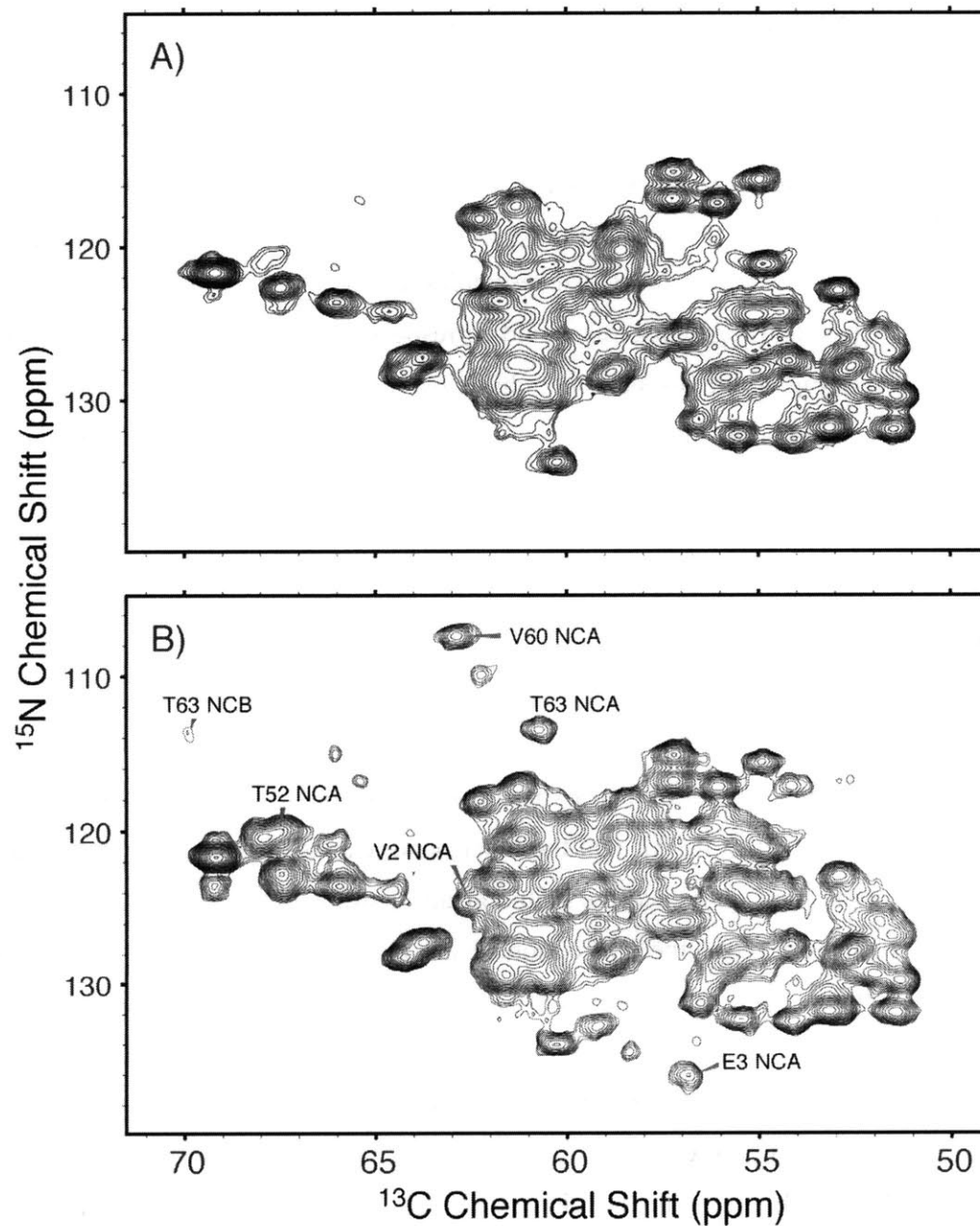
Figure 9. NCA spectra of vesicles collapsed in ^2H_2O with (*A*) 0.5, and (*B*) 8.0 ms ^1H-^{15}N CP time. The spectra were acquired at 700 MHz ^1H Larmor frequency and processed with 30 Hz Lorentzian to Gaussian line broadening in each dimension.

10.4 Discussion

10.4.1 GvpA subunit structure

Solid-state NMR assignments for 81% of the GvpA sequence support the coil-α-β-β-α-coil secondary structure that is predicted for this protein. This new information makes it necessary to reject the so-called "paperclip" model of gas vesicle subunit structure (50). In the "paperclip" model, GvpA is assumed to be in an all β-sheet conformation, and a correct tilt of the β-strands is obtained by placing two pairs of anti-parallel β-strands, of alternating 14- and 16-residue lengths, in an arrangement that resembles a paperclip. If a turn of 3 residues were assumed between β-strand segments, this would correspond to strand-turn regions of 33 residues. However, as seen in Fig. 4, the β-sheet region has a maximum length of only 27 residues, as it is flanked by α-helices on both sides. It is also clear that the termini beyond the helices are in coil conformation and solvent accessible, as indicated by ^2H_2O exchange at certain sites near the termini.

10.4.2 Air-water interface formation

Gas vesicles function by forming an interface between the aqueous phase and the air inside the vesicle. In contrast, ordinary proteins function fully immersed in condensed aqueous or lipid phases, and tend to unfold on exposure to an air-liquid interface (29, 47), with the hydrophobic core exposed to the air. On the other hand, extended β-sheets with an alternating polar-nonpolar amino acid sequence are stable at an air-water interface, as their amphipathic nature allows them to expose the hydrophobic side chains on one face to the air, and the hydrophilic side chains on the other face to the water. Indeed, amyloid fibrils have been found to self-assemble in air-water interfaces, with their hydrophobic side chains exposed to the air (42). Interfacial amyloids have even been found to occur naturally: the soil bacteria *Streptomyces coelicolor* secretes chaplin proteins (15) and filamentous fungi secrete hydrophobins (30), both proteins that self-assemble into amyloid-like fibrils at the air-water interface, thereby lowering the surface tension of water and presumably facilitating the growth of spore-forming hyphae into the air.

Gas vesicles have two segments that our NMR data indicate to be of β-sheet secondary structure, as predicted. Both segments have alternating polar-nonpolar residue patterns that are

highly conserved, although they are otherwise known to be disfavored by nature because of their propensity for amyloid formation (14). It is likely that this portion of GvpA is largely responsible for the air-water interface, as the presence of an inward facing hydrophobic surface would explain not only the ability of gas vesicles to prevent internal condensation of water, but also the strong 10 Å reflection, typical of stacked β-sheets, that is observed in X-ray diffraction studies of collapsed, but not of inflated gas vesicles (10).

Perhaps most intriguing about the functional amyloids is how cells prevent these molecules from accumulating in dysfunctional deposits. In particular, how do aquatic organisms control the assembly-disassembly of gas vesicles and the recycling of GvpA subunits? Chaperones (possibly the GvpO and GvpK proteins that are essential for gas vesicle expression, but are not detected in gas vesicles), are likely to be involved in the process, and the presence of the amphipathic α-helices in GvpA may help prevent catastrophic aggregation and precipitation. A similar issue arises in the secretion of chaplins and hydrophobins.

10.5 Conclusions

We have obtained well-resolved solid-state NMR spectra of intact, collapsed gas vesicles from *Anabaena flos-aquae*, allowing assignment of 81% of the GvpA sequence, followed by secondary chemical shift analysis, indicating a coil-α-β-β-α-coil secondary structure. Nearest-neighbor corrected ^{15}N secondary shifts provide evidence for curvature in the α-helices and ^2H_2O exchange experiments identify solvent exposed residues. These data provide insights into the topology of GvpA subunits in the gas vesicle shell and how it might stabilize the air-water interface. However, distance constraints are needed in order to obtain a structural model.

References

1. Baldus M, Iuliucci RJ, Meier BH. 1997. Probing through-bond connectivities and through-space distances in solids by magic-angle-spinning nuclear magnetic resonance. *J. Am. Chem. Soc* 119: 1121

2. Baldus M, Meier BH. 1996. Total correlation spectroscopy in the solid state. The use of scalar couplings to determine the through-bond connectivity. *J. Magn. Reson. Ser. A* 121: 65

3. Baldus M, Petkova AT, Herzfeld J, Griffin RG. 1998. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Molecular Physics* 95: 1197

4. Bayro MJ, Maly T, Birkett NR, Dobson CM, Griffin RG. 2009. Long-range correlations between aliphatic 13C nuclei in protein MAS NMR spectroscopy. *Angew. Chem. Int. Ed.* 48: 5708

5.   Bayro MJ, Ramachandran R, Caporini MA, Eddy MT, Griffin RG. 2008. Radio frequency-driven recoupling at high magic-angle spinning frequencies: Homonuclear recoupling sans heteronuclear decoupling. *J. Chem. Phys.* 128: 052321

6.   Belenky M, Meyers R, Herzfeld J. 2004. Subunit structure of gas vesicles: A MALDI-TOF MS study. *Biophys. J.* 86: 499

7.   Bennett AE, Griffin RG, Ok JH, Vega S. 1992. Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys.* 96: 8624

8.   Bennett AE, Rienstra CM, Auger M, Lakshmi KV, Griffin RG. 1995. Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* 103: 6951

9.   Bennett AE, Rienstra CM, Griffiths JM, Zhen W, Lansbury PT, Griffin RG. 1998. Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys.* 108: 9463

10.   Blaurock AE, Walsby AE. 1976. Crystalline structure of the gas vesicle wall from *Anabaena flos-aquae*. *J. Mol. Biol.* 105: 183

11.   Boucher Y, Douady CJ, Papke RT, Walsh DA, Boudreau MER, et al. 2003. Lateral gene transfer and the origins of prokaryotic groups. *Annu. Rev. Genet.* 37: 283

12.   Bowen CC, Jensen TE. 1965. Blue-green algae: fine structure if the gas vacuole. *Science* 147: 1460

13.   Braun D, Wider G, Wüthrich K. 1994. Sequence-corrected ^{15}N "random coil" chemical shifts. *J. Am. Chem. Soc* 116: 8466

14.   Broome BM, Hecht MH. 2000. Nature disfavors sequences of alternating polar and non-polar amino acids: implications for amyloidogenesis. *J. Mol. Biol.* 296: 961

15.   Claessen D, Rink R, de Jong W, Siebring J, de Vreugd P, et al. 2003. A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in *Streptomyces coelicolor* by forming amyloid-like fibrils. *Genes Dev.* 17: 1714

16.   Cornilescu G, Delaglio F, Bax A. 1999. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* 13: 289

17. de Dios AC, Pearson JG, Oldfield E. 1993. Secondary and tertiary structural effects on protein NMR chemical shifts: an ab initio approach. *Science* 260: 1491

18. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6: 277

19. Dunton PG, Mawby WJ, Shaw VA, Walsby AE. 2006. Analysis of tryptic digests indicates regions of GvpC that bind to gas vesicles of *Anabaena flos-aquae*. *Microbiol.* 152: 1661

20. Goddard TD, Kneller DG. SPARKY 3. University of California, San Francisco

21. Hardy EH, Verel R, Meier BH. 2001. Fast MAS total through-bond correlation spectroscopy. *J. Magn. Reson.* 148: 459

22. Hayes PK, Buchholz B, Walsby AE. 1992. Gas vesicles are strengthened by the outer-surface protein, GvpC. *Arch. Microbiol.* 157: 229

23. Hayes PK, Lazarus CM, Bees A, Walker JE, Walsby AE. 1988. The protein encoded by *gvpC* is a minor component of gas vesicles isolated from the cyanobacteria *Anabaena flos-aquae* and *Microcystis* sp. *Mol. Microbiol.* 2: 545

24. Hayes PK, Walsby AE, Walker JE. 1986. Complete amino acid sequence of cyanobacterial gas-vesicle protein indicates a 70-residue molecule that corresponds in size to the crystallographic unit cell. *Biochem. J.* 236: 31

25. Heise H, Hoyer W, Becker S, Andronesi OC, Riedel D, Baldus M. 2005. Molecular-level secondary structure, polymorphism, and dynamics of full-length α-synuclein fibrils studied by solid-state NMR. *Proc. Nat. Acad. Sci. USA.* 102: 15871

26. Igumenova TI, McDermott AE, Zilm KW, Martin RW, Paulson EK, Wand AJ. 2004. Assignments of carbon NMR resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* 126: 6720

27. Kloepper KD, Zhou DH, Li Y, Winter KA, George JM, Rienstra CM. 2007. Temperature-dependent sensitivity enhancement of solid-state NMR spectra of α-synuclein fibrils. *J. Biomol. NMR.* 39: 197

28. Luca S, Filippov DV, van Boom JH, Oschkinat H, de Groot HJM, Baldus M. 2001. Secondary chemical shifts in immobilized peptides and proteins: a quantitative basis for structure refinement under magic angle spinning. *J. Biomol. NMR* 20: 325

29. Maa Y-F, Hsu CC. 1997. Protein denaturation by combined effect of shear and air-liquid interface. *Biotechnol. Bioeng.* 54: 503

30. Mackay JP, Matthews JM, Winefield RD, Mackay LG, Haverkamp RG, Templeton MD. 2001. The hydrophobin EAS is largely unstructured in solution and functions by forming amyloid-like structures. *Structure* 9: 83

31. McDermott A, Polenova T, Böckmann A, Zilm KW, Paulsen EK, et al. 2000. Partial NMR assignments for uniformly (^{13}C, ^{15}N)-enriched BPTI in the solid state. *J. Biomol. NMR.* 16: 209

32. McGuffin LJ, Bryson K, Jones DT. 2000. The PSIPRED protein structure prediction server. *Bioinformatics* 16: 404

33. McMaster TJ, Miles MJ, Walsby AE. 1996. Direct observation of protein secondary structure in gas vesicles by atomic force microscopy. *Biophys. J.* 70: 2432

34. Morcombe CR, Zilm KW. 2003. Chemical shift referencing in MAS solid state NMR. *J. Magn. Reson.* 162: 479

35. Morris GA, Freeman R. 1979. Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc* 101: 760

36. Murzin AG. 2008. Metamorphic proteins. *Science* 320: 1725

37. Offner S, Hofacker A, Wanner G, Pfeifer F. 2000. Eight of fourteen gvp genes are sufficient for formation of gas vesicles in halophilic archaea. *Journal of Bacteriology* 182: 4328

38. Offner S, Ziese U, Wanner G, Typke D, Pfeifer F. 1998. Structural characteristics of halobacterial gas vesicles. *Microbiol.* 144: 1331

39. Pines A, Gibby MG, Waugh JS. 1973. Proton-enhanced NMR of dilute spins in solids. *J. Chem. Phys.* 59: 569

40. Ritter C, Maddelein M-L, Siemer AB, Lührs T, Ernst M, et al. 2005. Correlation of structural elements and infectivity of the HET-s prion. *Nature* 435: 844

41.   Schaefer J, Stejskal EO, Garbow JR, McKay RA. 1984. Quantitative determination of the concentrations of ^{13}C-^{15}N chemical bonds by double cross-polarization NMR. *J. Magn. Reson.* 59: 150

42.   Schladitz C, Vieira EP, Hermel H, Möhwald H. 1999. Amyloid-β-sheet formation at the air-water interface. *Biophys. J.* 77: 3305

43.   Shukla HD, DasSarma S. 2004. Complexity of gas vesicle biogenesis in *Halobacterium* sp. strain NRC-1: identification of five new proteins. *J. Bacteriol.* 186: 3182

44.   Sivertsen AC, Bayro MJ, Belenky M, Griffin RG, Herzfeld J. 2009. Solid-state NMR evidence for inequivalent GvpA subunits in gas vesicles. *Journal of Molecular Biology* 387: 1032

45.   Suter D, Ernst RR. 1985. Spin diffusion in resolved solid-state NMR spectra. *Phys. Rev. B.* 32: 5608

46.   Szeverenyi NM, Sullivan MJ, Maciel GE. 1982. Observation of spin exchange by two-dimensional fourier transform 13C cross polarization-magic-angle spinning. *J. Magn. Reson.* 47: 462

47.   Tronin A, Dubrovsky T, Dubrovskaya S, Radicchi G, Nicolini C. 1996. Role of protein unfolding in monolayer formation on air-water interface. *Langmuir* 12: 3272

48.   Walsby AE. 1969. The permeability of blue-green algal gas-vacuole membranes to gas. *Proc. Royal Soc. London B.* 173: 235

49.   Walsby AE. 1976. The buoyancy-providing role of gas vacuoles in an aerobic bacterium. *Arch. Microbiol.* 109: 135

50.   Walsby AE. 1994. Gas vesicles. *Microbiol. Rev.* 58: 94

51.   Walsby AE, Buckland B. 1969. Isolation and purification of intact gas vesicles from a blue-green alga. *Nature* 224: 716

52.   Walsby AE, Eichenberger HH. 1968. The fine structure of gas-vacuoles released from cells of the blue-green alga *Anabaena flos-aquae. Arch. Microbiol.* 60: 76

53.   Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, et al. 1995. ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR.* 6: 135

54.	Wishart DS, Case DA. 2001. Use of chemical shifts in macromolecular structure determination. *Methods Enzymol.* 338: 3

55.	Wishart DS, Sykes BD. 1994. Chemical shifts as a tool for structure determination. *Methods Enzymol.* 239: 363

56.	Zhou NE, Zhu BY, Sykes BD, Hodges RS. 1992. Relationship between amide proton chemical shifts and hydrogen bonding in amphipathic .alpha.-helical peptides. *J. Am. Chem. Soc.* 114: 4320

Appendix to Chapter 10

Table 1. MAS NMR ^{13}C and ^{15}N chemical shifts of GvpA (in ppm).

Res	CO	CA	CB	N	Side-chain
A01	173.7	52.1	20.5	43.4	
V02	177.0	62.7	32.2	125.4	CG1: 23.0, CG2: 22.0
E03	175.5	57.1	31.4	136.7	
S09	174.2	58.6	66.9	118.7	
S10	175.9	57.9	65.8	121.2	
L11	178.4	58.5	41.7	124.0	CG: 27.8, CD1: 27.5, CD2: 27.2
A12	180.1	56.1	18.6	123.9	
E13	180.0	59.9	30.5	120.2	
V14	178.3	67.5	32.3	123.7	CG1: 23.9, CG2: 21.7
I15	178.3	67.6	39.0	122.5	CD1: 14.0, CG1: 31.7, CG2: 18.3
D16	180.2	58.6	41.7	120.2	
R17	180.8	59.9	31.8	120.3	CG: 28.2, CD: 45.1, CZ: 160.5, NE: 88.4
I18	177.9	66.1	39.0	123.3	CG1: 30.5, CG2: 19.8, CD1: 15.1
L19	176.2	55.7	42.1	122.3	CG: 27.7
D20	180.3	57.7	41.5	121.7	
K21	180.0	59.5	34.0	122.4	CG: 26.1
G22	170.9	44.9		115.1	
I23	175.3	62.2	43.1	118.0	CD1: 14.9, CG1: 27.6, CG2: 18.0
V24	175.6	61.6	35.8	128.5	CG1: 22.5, CG2: 21.8
I25	173.8	61.2	43.3	127.9	CD1: 16.2, CG1: 28.4, 19.4
D26	180.4	58.9	41.5	128.4	
A27	175.6	52.5	24.2	128.2	
W28		56.0	32.2	121.6	CG: 111.0, CD1: 129.8
V34	180.0	57.8	32.0	125.6	CG1: 21.1, CG2: 21.0
G35	169.9	45.1		110.8	
I36	175.8	61.3	43.2	117.3	

E37	176.9	55.5	36.1	132.1	CG: 36.4
L38	175.9	56.7	45.5	131.4	CG: 28.2
L39	175.9	54.2	48.2	127.6	CG: 27.9, CD1: 27.3
A40	176.8	52.8	22.2	123.1	
I41	173.9	61.6	42.1	125.2	CG1: 28.3
E42	176.8	55.6	35.8	132.1	
A43	176.0	51.5	24.6	126.6	
R44	179.9	58.9	36.6	126.4	CG: 28.7, NE: 88.7
I45	175.8	60.8	42.8	128.4	CG1: 28.6, CD1: 16.0
V46	175.5	61.4	35.1	129.3	CG1: 21.3, CG2: 19.6
I47	175.5	60.9	42.3	127.6	CG1: 27.8, CG2: 17.5, CD1: 14.8
A48	175.6	52.6	24.3	128.1	
A48x	175.6	52.4	24.3		
S49	175.5	57.3	65.4	116.7	
S49x	175.4	57.1	65.9	115.0	
V50	178.4	69.3	31.6	121.5	CG1: 25.3, CG2: 22.5
E51	180.1	61.3	30.6	121.1	CG: 37.9, CD: 184.3
T52	176.8	67.6	68.3	120.5	CG2: 23.4
T52x	175.8	68.0	69.4	120.9	
Y53	178.5	63.8	39.6	127.1	CG: 130.4, CD: 133.4, CE: 119.0, CZ: 160.2
Y53x	178.3	64.1	39.4	127.9	
L54	178.6	58.6	41.8	120.4	CG: 27.3
K55	180.4	60.5	32.9	122.0	CG: 26.1, CD: 30.5, CE: 43.0, NZ: 36.0
Y56	175.9	61.8	38.2	123.2	CG: 129.4, CE: 119.6
Y56x		61.3	37.9		
A57	180.2	55.6	20.9	124.3	
A57x				124.1	
E58	180.1	58.7	30.4	123.9	
A59	180.5	55.4	20.9	124.2	
V60	176.9	62.9	32.8	107.7	CG1: 20.9, CG2: 18.0
G61	175.5	46.2		110.3	

L62	175.4	55.0	48.1	121.4	CG: 27.8, CD1: 24.2, CD2: 23.7
T63	174.2	60.9	70.0	113.6	CG2: 23.2
Q64	180.4	61.2	30.5	120.9	
S65	175.7	58.6	65.9	121.3	
A66	180.4	55.3	19.2	124.3	
P69	175.5	57.0	31.4	136.7	

Figure S1. Helical wheels for the two α-helices in GvpA. Grey indicates Non-polar/polar non-charged residues.
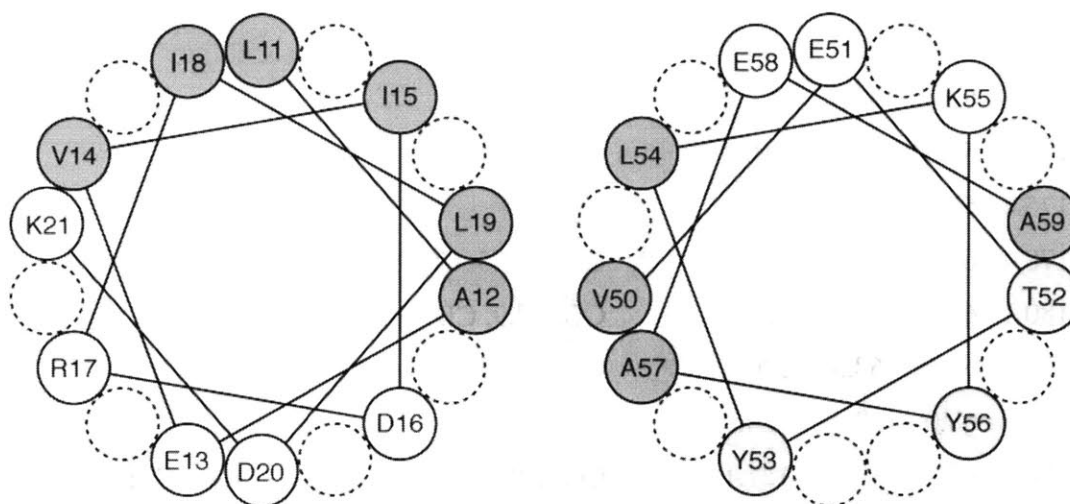
Part III

Manifestations of Protein Dynamics in MAS NMR Correlation Spectroscopy

Chapter 11. Low-temperature enhancement of long-range polarization transfer efficiency in solid-state NMR spectroscopy of amyloid fibrils

Adapted from a manuscript in preparation by Marvin J. Bayro, Neil R. Birkett, Christopher M. Dobson, and Robert G. Griffin

Summary

We have examined manifestations of protein backbone dynamics in long-range solid-state NMR correlation experiments of amyloid fibrils formed by the 86-residue SH3 domain of PI3 kinase (PI3-SH3). At temperatures below the sample freezing point, only few variations are evident in magic-angle spinning NMR spectra of these cryoprotected amyloid fibrils. However, we observe an overall increase in the homonuclear polarization transfer efficiency attainable between distant ^{13}C nuclei with band-selective radio frequency-driven recoupling (BASE RFDR), but not with spin difussion techniques. This enhancement in recoupling efficacy can be interpreted as a consequence of restricted molecular mobility at low temperatures, and it appears to be sensitive to the mechanism of polarization transfer under effect.

11.1 Introduction

The relationship between molecular motion and solid-state NMR spectral observables has long been a topic of great interest as the latter may provide unique insights regarding the rates and character of dynamic processes. In particular, dipole-dipole and quadrupolar couplings are highly sensitive anisotropic reporters of local motions in the solid state,[1-6] and a variety of experiments have been developed to extract related dynamical information. Conversely, molecular dynamics has also been shown to degrade magic-angle spinning (MAS) spectra due to interference with experimental timescales such as those associated with the radio-frequency (rf) fields applied during cross-polarization (CP) or heteronuclear decoupling, or with the sample rotation frequency.[1-3, 7-9] These interference processes typically manifest themselves as a broadening of resonance lines arising from dynamics at specific molecular sites but may also have a more general impact on the quality of the MAS NMR spectra attainable. In cases of extreme mobility, dipolar-mediated preparation (via CP) and mixing (via dipolar recoupling) may be limited severely and solution NMR schemes may be applied successfully in the solid state. However, in many biomolecular systems studied by solid-state NMR, such as amyloid fibrils and membrane proteins, the protein's fold or environment partially restricts its mobility, resulting in sensitive CP spectra that nevertheless may present lower quality or intensity than those of well-ordered, rigid solids.[4, 8-10]

Temperature has a profound effect on protein dynamics and thus on the shape and intensity of solid-state NMR lines.[6, 10, 11] Similar dynamic effects may influence the efficiency of MAS NMR polarization transfer experiments that rely on the reintroduction of dipolar interactions.[12] Molecular motion has the effect of averaging dipolar couplings into effective, reduced values. A consequence of such averaging is the decrease in dipolar recoupling efficiency for protein sites presenting significant conformational flexibility, as noted in various one-bond correlation spectra in previous studies.[8, 11, 13] This dynamic effect on dipolar recoupling is increasingly accentuated for longer internuclear distances, which are subject to multiple motional modes.[12]

In this chapter, we investigate the effect of sample temperature on the efficiency of polarization transfer between distant ^{13}C nuclei in MAS NMR correlation spectra of amyloid fibrils formed by PI3-SH3 (the SH3 domain of the p85α subunit of phosphatidylinositol 3-kinase), an 86-residue protein. We focus our analysis on two robust techniques for long-range

polarization transfer, proton-driven spin diffusion (PDSD)[14] and the recently described BASE RFDR (band-selective radio frequency-driven recoupling) scheme,[15] in a protein sample produced with [2-^{13}C] glycerol as the carbon source. This type of labeling[16] is optimal for generating sequential and long-range Cα-Cα correlations,[15] which one may anticipate to be sensitive to motions in the protein backbone. We note that the two different polarization transfer mechanisms lead to correlation spectra with different degrees of susceptibility to protein mobility, dependent on temperature. Notably, we observe considerable gains in recoupling efficiency for most cross-peaks in BASE RFDR spectra at temperatures below the sample freezing point.

11.2 Results and Discussion

11.2.1 Room-temperature dynamics and depolarization

In solid-state NMR correlation spectra, a large degree of variability in efficiency is observed between different mixing schemes utilized for polarization transfer.[11, 17] Polarization transfer efficiency is determined by inherent properties of the mixing pulse sequence (scaling factor, sensitivity to crystallite orientation, frequency offset dependence, etc.) as well as by specific experimental conditions (rf homogeneity, ^{1}H decoupling, heteronuclear interference,[17] MAS frequency, etc.) and by the nature of the spin system or molecule under study, which may present deleterious effects such as dipolar truncation[18] or interfering dynamic effects.[1] Indeed, in correlation spectra of proteins, the intensities of cross-peaks may vary significantly depending on the dynamic properties of different molecular sites with respect to the timescales of the experiments.[11, 13, 19] For example, backbone carbons, rigid side-chains, and mobile side-chains tend to have dissimilar degrees of mobility. These site-specific dynamics may have different effects in correlation spectra depending on the mixing scheme applied.

Sections of ^{13}C-^{13}C dipolar correlation spectra showing cross-peaks between ^{13}Cα and side-chain (^{13}Cγ, ^{13}Cδ, and ^{13}Cε) nuclei of Lysine residues are illustrated in Fig. 1, recorded at ambient temperature. In the extensive longitudinal mixing period required temperature. Lys ^{13}Cδ and ^{13}Cε side-chain resonances decay in the extensive longitudinal mixing period required for polarization transfer via PDSD (300 ms, in this case), and thus yield no cross-peaks. In contrast, the 18 ms mixing period of BASE RFDR is sufficient for efficient recoupling between ^{13}Cα and ^{13}Cδ/^{13}Cε nuclei so that these cross-peaks (corresponding to internuclear distances of 3.6 to 4.8

Å) appear with high sensitivity. The signal decay of these resonances during long PDSD mixing periods is likely the result of conformational exchange due to the flexibility of Lysine side-chains. It is interesting to note that out of the six Lys residues in the PI3-SH3 sequence, intense side-chain cross-peaks appear for only three of them, while the rest are not observed probably as a consequence of local dynamic or static disorder. A behavior similar to that exhibited by these Lys resonances is also observed for side-chains of other charged residues such as Glu, Asp, and Arg (evidenced, for example, in the difference between PI3-SH3 1D spectra obtained via ^{13}C Bloch decay and ^{1}H-^{13}C cross-polarization, shown in chapter 7), and also for ^{13}C signals from aromatic rings. While some of these sites may undergo more rapid motions than the rest of the protein (e.g., solvent-exposed side-chains), we may expect that even subtle differences in dynamics can lead to significant differences in cross-peak intensity. This is one of the reasons that different polarization transfer schemes tend to yield complementary information in protein solid-state NMR studies.[20]
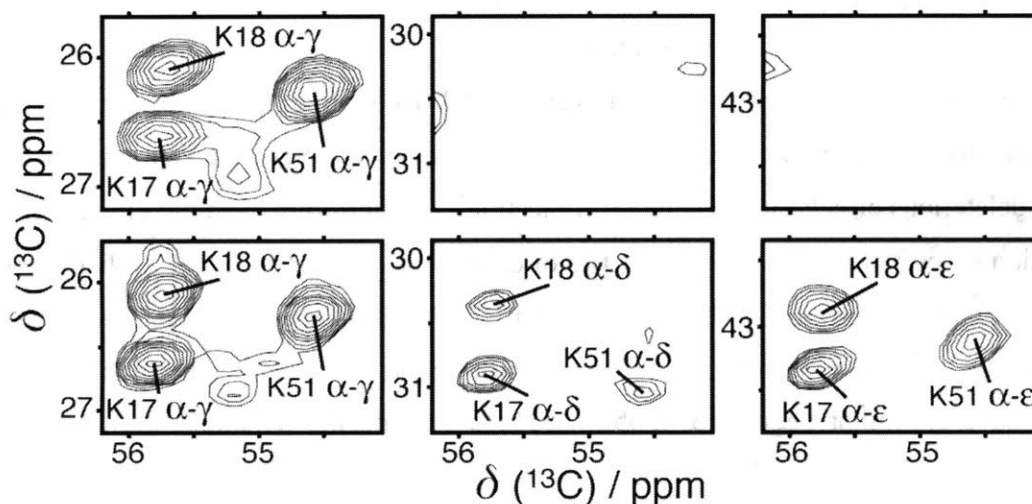
Figure 1. Sections of ^{13}C-^{13}C correlation spectra of PI3-SH3 fibrils corresponding to Lys side-chain resonances, recorded at ambient temperature with 300 ms PDSD mixing (top panels) and 18 ms BASE RFDR mixing (bottom panels).

11.2.2 Temperature effect on backbone and side-chain resonances

While certain side-chains present considerable dynamics in PI3-SH3 fibrils at room temperature, we investigated the extent of motion in the protein backbone by the application of variable-temperature experiments. As illustrated in Figure 2, when the sample temperature is decreased from 298 K to 278 K, a small number of sharp peaks in the aliphatic region of the ^{13}C spectrum begin to decrease in intensity. These peaks correspond mainly to Lys and Arg side-chains, which are charged and may be expected to be accessible to the solvent. On the other hand, as the sample is cooled to 228 K (roughly 10 degrees below the freezing point of the fibril/glycerol/water mixture), a slight broadening throughout the entire spectrum is evident. Nonetheless, a considerable degree of resolution is preserved. Contrary to the case of protein systems that exhibit extensive mobility, there is no appreciable intensity enhancement of one-dimensional (1D) ^{13}C CP spectra of PI3-SH3 fibrils at temperatures below the sample freezing point, *suggesting* that the majority of the protein is in a vastly rigid conformation both at low temperature and at room temperature.
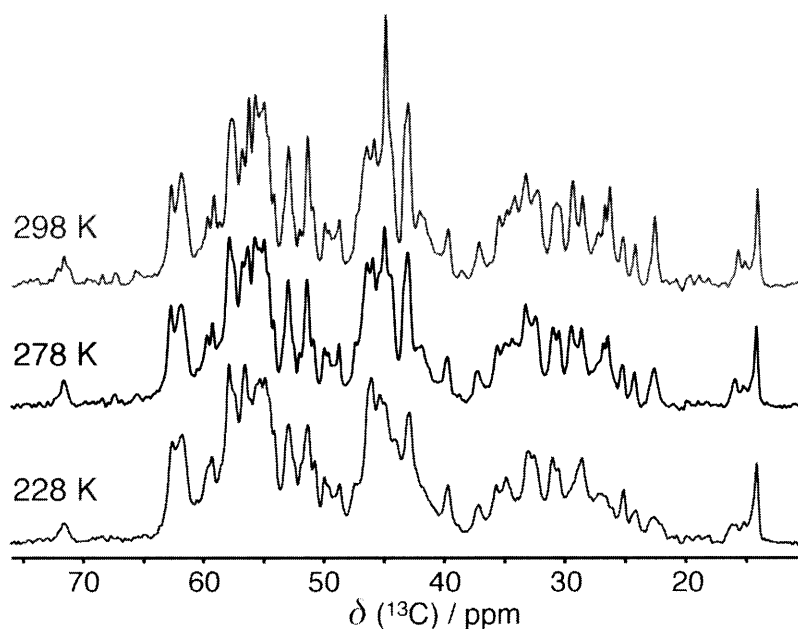

Figure 2. Aliphatic region of ^{13}C cross-polarization spectra of PI3-SH3 fibrils recorded at decreasing temperatures. The temperature-dependent spectral changes are fully reversible.

The limited temperature dependence of PI3-SH3 CP spectra is consistent with other characteristics of the fibrils. Indeed, PI3-SH3 subunits in mature amyloid fibrils adopt a compact form in which most of the protein is buried within the core of the fibril, as evidenced by resistance to limited proteolysis[21] and by spectroscopic signatures described in chapter 7. Specifically, the comparison of CP and direct polarization (Bloch decay) ^{13}C spectra at room temperature reveals few differences and an enhancement factor of about 2.5 arising from ^{1}H-^{13}C CP transfer, as shown in Figure 3 (top), which is the value typically observed in rigid crystalline peptides and proteins. In addition, J-coupling mediated ^{13}C excitation during MAS, which only succeeds for nuclei in highly mobile sites where the reduced dipolar interactions are too weak to interfere with the evolution of J-couplings, does not generate observable signals from PI3-SH3 fibrils, which verifies the absence of solution-like dynamics.

11.2.3 Observation of dynamic interference at low temperature

Despite the overall rigidity of PI3-SH3 molecules in the amyloid fibril state, and the absence of highly flexible regions, the remaining conformational fluctuations may be expected to have an effect on the observable spectral characteristics. At room temperature these fluctuations are rapid and of small amplitude. Such fast motions average homonuclear and heteronuclear dipolar couplings to lower effective magnitudes, which in the absence of other motions may lead to high-quality dipolar-mediated spectra. On the other hand, as the temperature is decreased, these conformational fluctuations slow down and may potentially interfere with the frequencies associated with MAS NMR experiments and lead to line broadening, as discussed above. For example, interference with radio-frequency irradiation would render ^{1}H decoupling and cross-polarization less effective, resulting in broad line shapes and decreased signal intensity. However, it is often difficult to identify such dynamic effects and distinguish them from static conformational heterogeneity often encountered at low temperatures. For example, the differences in 1D spectra of PI3-SH3 fibrils as a function of temperature shown in Figure 2 could arise from either static disorder or dynamic interference, which itself could be due to the r.f. fields applied during ^{1}H-^{13}C CP or ^{1}H decoupling or due to the spinning frequency.
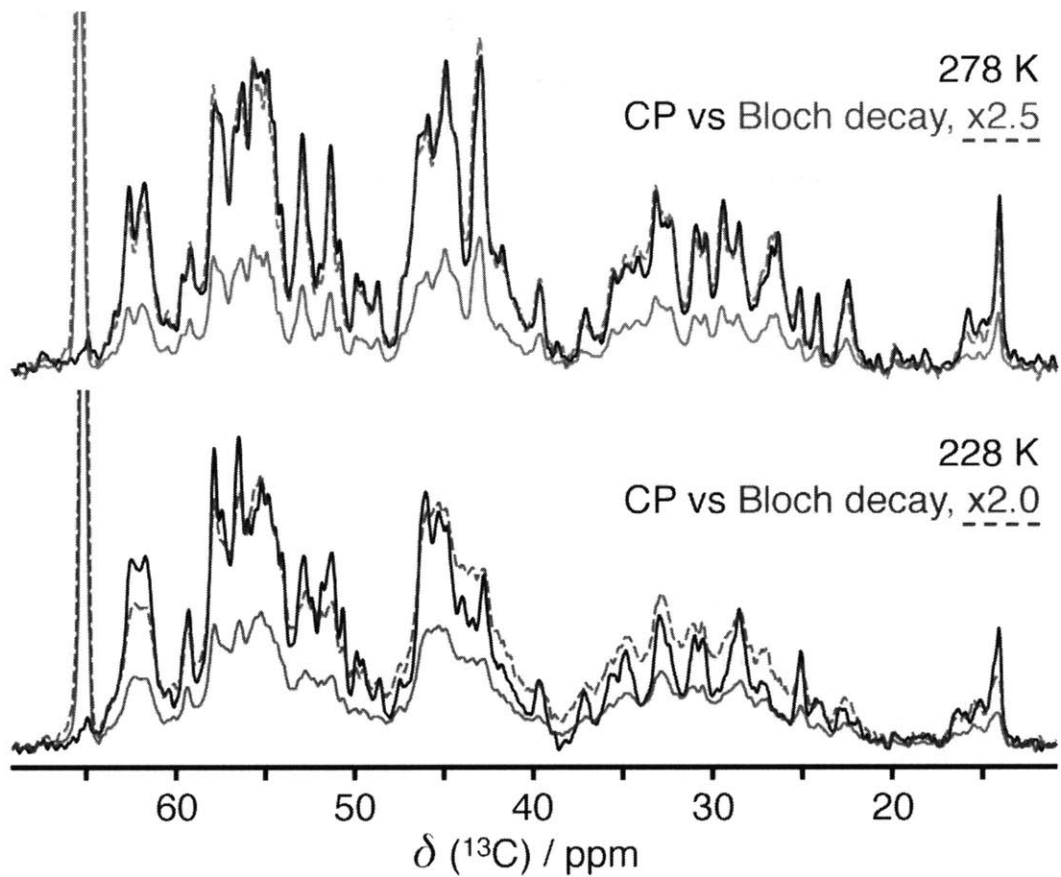
Figure 3. Dynamic interference at low temperature observed by comparison of cross-polarization (black trace) and Bloch decay (red or blue traces) spectra. An intensity enhancement of ~2.5 from CP is expected for proteins with low mobility and it observed at 278 K (top), but the enhancement factor is only ~2.0 at 228 K (bottom). Red and blue dashed traces are Bloch decay spectra, recorded at 278 K and 228 K, scaled by a factor of 2.5 and 2.0, respectively.

As illustrated in Figure 3, comparison of ^{13}C CP and Bloch decay spectra of PI3-SH3 fibrils recorded at different temperatures allows the identification of dynamic interference effects. In this figure, CP spectra (solid black trace) are overlaid with Bloch decay spectra (solid red and blue traces) and Bloch decay spectra scaled approximately by the observed CP enhancement factor (dashed red and blue traces), recorded at temperatures above and below the sample freezing point. The CP enhancement factor is 2.5 and very uniform throughout the spectrum recorded at 278 K, with a nearly identical overlay of CP and Bloch decay (scaled by 2.5). On the other hand, it is on average 2.0 at 228 K, with different spectral regions presenting

varying degrees of CP enhancement. Specifically, side-chain resonances (between 20 and 45 ppm) show lower CP/Bloch decay ratios than backbone resonances (between 45 and 65 ppm), consistent with a stronger dynamic interference effect in the side-chain groups.

In addition to the differences in intensity enhancement at low temperature, which can be attributed to dissimilar extents of interference with the r.f. fields applied during CP, there are also differences in line broadening for backbone and side-chain ^{13}C sites. Backbone resonances in the Bloch decay spectrum have an overall line shape that matches the CP spectrum quite closely, except for small variations in signal intensity. (See solid black and dashed blue traces in the bottom of Figure 3.) On the other hand, the side-chain region in the Bloch decay spectrum is much broader than the corresponding region in the CP spectrum. Indeed, side-chain resonances that cross-polarize well are also narrower than those that do not. We may then conclude that, in addition to interfering with CP, dynamic effects in PI3-SH3 fibrils at 228 K also influence the performance of ^{1}H decoupling irradiation. Therefore we have positively identified dynamic interference effects by comparing CP and Bloch decay (direct polarization) spectra and observing differences in line width between the two and variations in CP enhancement for different sites. Both of these observations would be absent in the case of a statically disordered sample without dynamic interference.

11.2.4 Temperature effect on backbone-backbone correlation spectra

The above observations indicate that PI3-SH3 fibrils exhibit a limited extent of dynamics, particularly in the peptide backbone, perhaps only restricted to small amplitude motions that are modulated slightly with temperature. Nevertheless, significant variations in long-range ^{13}C-^{13}C correlation spectra are visible at different temperatures, as depicted in Figure 4. The BASE RFDR experiment uses 16 ms of mixing, which is optimized for a 4 Å inter-nuclear distance, and multiple sequential Cα-Cα cross-peaks are readily observable at room temperature (Figure 4a). The low band width of the BASE RFDR pulses confine the recoupling process to aliphatic spins, thus reducing relayed polarization transfer (via carbonyl nuclei) and maximizing the sensitivity of aliphatic-aliphatic correlations.[15] Comparable, although slightly less sensitive, spectra can be obtained via PDSD at room temperature, with a 300 ms mixing time (Figure 4d). Increasing the mixing time up to 500 ms yields marginal increases in the sensitivity of sequential Cα-Cα cross-peaks (data not shown). At a sample temperature of 278 K, BASE RFDR (Fig. 3b) and PDSD

314

(Figure 4e) spectra vary only slightly with respect to those at 298 K, and no overall trend is observed besides the lower intensity of certain cross-peaks between backbone and side-chain nuclei. This intermediate temperature (278 K) is close to those that are typically used in protein solid-state NMR studies, in which caution must be taken to preserve the sample integrity and prevent dehydration due to heating. It is therefore interesting to note the potential advantage of moderately higher temperatures to observe certain side-chain cross-peaks with better sensitivity.

At a temperature of 228 K, below the sample freezing point, significant changes are finally observed. First, the moderate line broadening observed in 1D spectra is reproduced here in $C\alpha$-$C\alpha$ cross-peaks. On average, line widths increase from 80 Hz to 100 Hz between 278 K and 228 K spectra, respectively. This broadening is most evident in the PDSD spectrum of Figure 4f, which is considerably degraded compared to PDSD spectra at higher temperatures. Despite the lower apparent intensity at the same contour level, peak fitting reveals that the volumes of $C\alpha$-$C\alpha$ cross-peaks remain approximately constant for PDSD spectra at the various temperatures studied. The lower sensitivity in the PDSD spectrum is thus simply the outcome of line broadening at 228 K, while the transfer efficiency is unaltered.
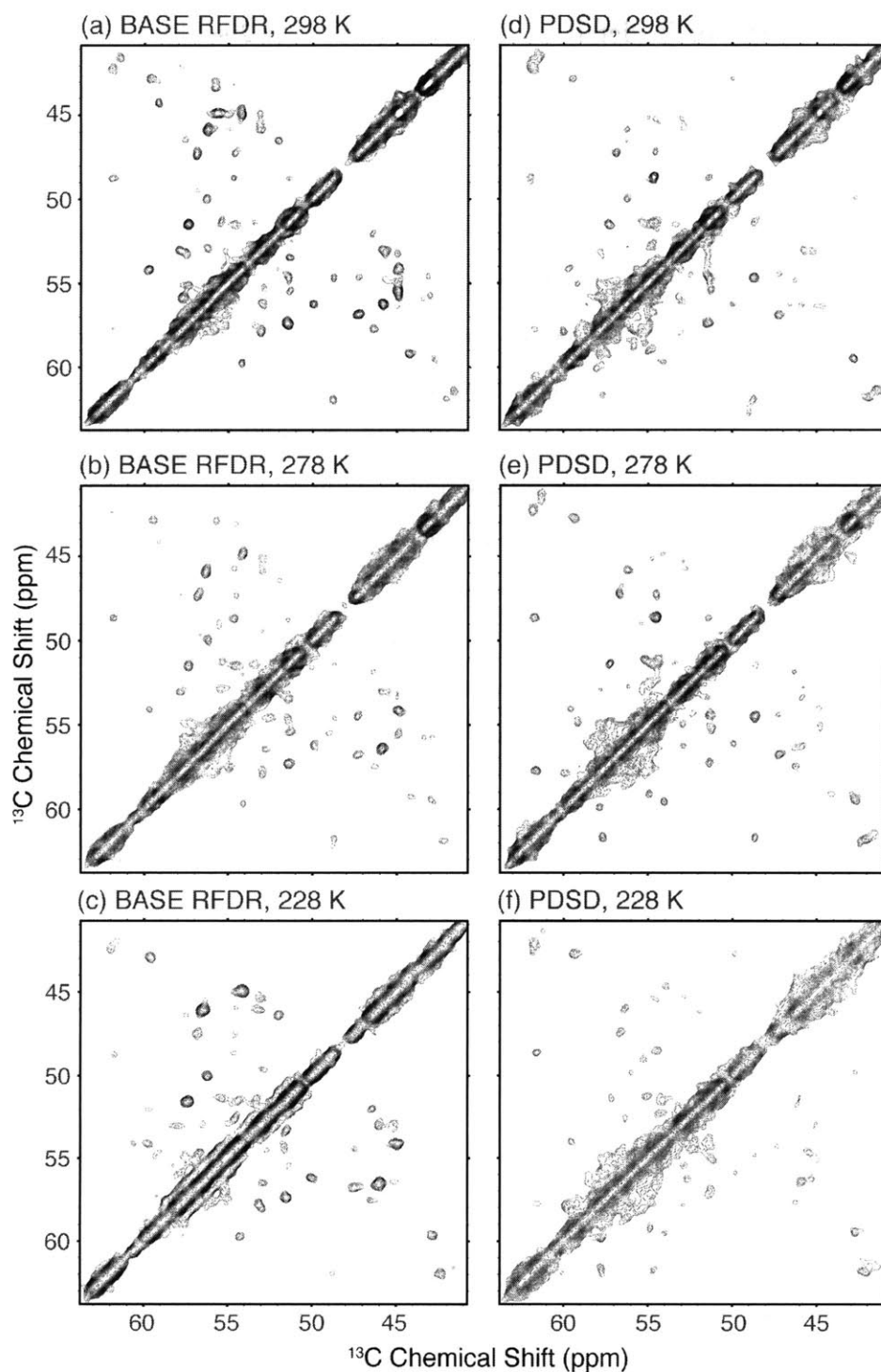
Figure 4. Long-range correlation spectra of PI3-SH3 fibrils recorded with 16 ms BASE RFDR (a-c) and 300 ms PDSD (d-f) at various temperatures. At low temperature, slight cross-peak broadening is accompanied by increased volumes in BASE RFDR but not PDSD experiments.

On the other hand, the BASE RFDR spectrum recorded at 228 K, shown in Figure 4c, presents strong cross-peaks of similar sensitivity as those at higher temperatures, despite the slightly larger line widths. This is the result of an enhancement in BASE RFDR polarization transfer efficiency at low temperature. The increase in recoupling efficiency is likely a consequence of the restriction of backbone dynamics at temperatures below the freezing point, such that the effective ^{13}C-^{13}C dipolar couplings are averaged to a lesser degree by such motions. Since this effect is negligible in PDSD spectra, we can speculate that this dynamic enhancement is sensitive to the mechanism of polarization transfer. In BASE RFDR, zero-quantum polarization transfer is mediated primarily by the direct reintroduction of homonuclear ^{13}C-^{13}C dipolar interactions, with a recoupled dipolar Hamiltonian similar to that encountered in (finite-pulse) RFDR at high spinning frequencies.[22] On the other hand, in PDSD experiments, polarization transfer occurs through the second-order interaction between homonuclear and heteronuclear couplings, yielding an effective Hamiltonian that depends weakly on direct ^{13}C-^{13}C couplings and more strongly on ^{1}H-^{13}C couplings and the strongly coupled bath of ^{1}H spins.[14, 23] It is feasible that the two types of dipolar interactions that drive ^{13}C-^{13}C polarization transfer in PDSD experiments (^{1}H-^{13}C and ^{1}H-^{1}H couplings) are not affected significantly by the change in amplitude of backbone dynamics expected between room temperature and temperatures below the sample freezing point.

As Figure 5 shows, the temperature dependence of BASE RFDR transfer efficiency is quite substantial. Cross-peak volumes increase by a factor of 50% for many of the observed sequential Cα-Cα correlations at 228 K, compared to those at 278 K. Considering the 25% increase in line width at 228 K, the actual improvement in sensitivity is approximately 35%. Nonetheless, such an enhancement is reasonably auspicious, in particular for cases of limited sensitivity. It is also possible that larger enhancement factors may be encountered for longer inter-nuclear distances than those examined here, as more dynamic modes are suppressed at low temperatures.
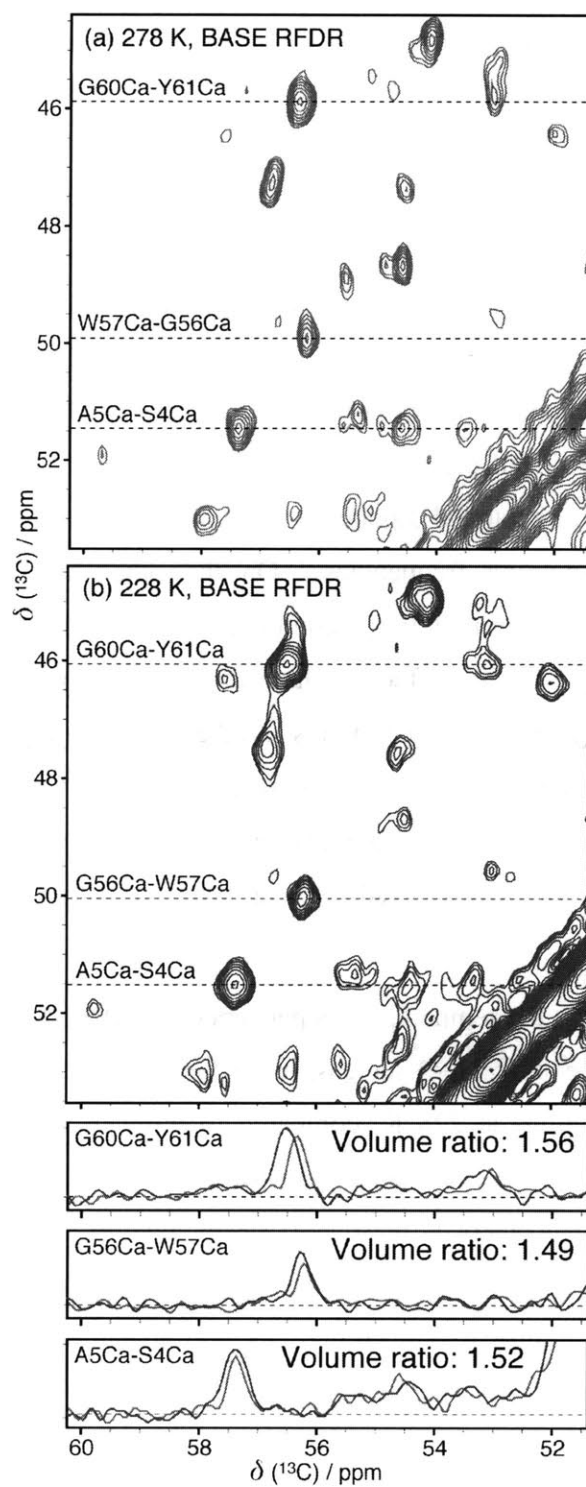
Figure 5. Temperature-dependent enhancement of BASE RFDR polarization transfer efficiency in sequential Cα-Cα correlations.

The temperature-dependent enhancement of BASE RFDR polarization transfer highlights some of the potential benefits of low-temperature studies of biomolecules by solid-state NMR. Recent developments in instrumentation for low-temperature MAS NMR[24, 25] has been driven in part by the need to increase the sensitivity of solid-state NMR experiments, either through the amplified spin polarization at cryogenic temperatures or via dynamic nuclear polarization (DNP) with stable radicals,[26, 27] an electron-nuclear process that yields large gains in NMR sensitivity at temperatures of about 100 K and below. Indeed, as variable-temperature and DNP instrumentation compatible with high-frequency spinning are becoming widespread, MAS NMR applications at low temperatures, including biomolecular studies, begin to present great promise. Similar studies to those described in this chapter are currently underway in our laboratory at cryogenic temperatures.

11.3 Conclusions

We have illustrated various examples of the influence of internal protein dynamics on solid-state MAS NMR spectra and in ^{13}C-^{13}C correlations between distant nuclei. We started with the observation that resonances from mobile side-chains often disappear after extended periods of longitudinal mixing, but can be observed with more efficient experiments with shorter mixing times. Then we investigated the effect of temperature on 1D spectra of PI3-SH3 fibrils. As may be expected, generally broader lines were observed at temperatures below the freezing point of the sample mixture. However, we were able to identify the source of broadening as a dynamic effect instead of static disorder by comparing cross-polarization and direct polarization spectra at different temperatures and observing differences in CP enhancement and line broadening in different regions of the spectra, namely backbone versus side-chain resonances. This approach may prove useful in the identification of line broadening at room temperature in the case of highly dynamic proteins. Finally, we demonstrated that restricting molecular motion by means of low temperatures can lead to gains in the polarization transfer efficiency between distant nuclei obtained with certain homonuclear mixing schemes (BASE RFDR in this case). Conversely, we may conclude that long-range correlation experiments require a judicious choice of pulse sequence that must take into account inherent molecular dynamics in the system under study.

11.4 Materials and Methods

11.4.1 Protein samples

A sample of PI3-SH3 in amyloid fibril form was prepared as described previously,[28] but with protein produced with [2-^{13}C]glycerol and NaH^{13}CO$_3$ as the sole sources of carbon and with uniform ^{15}N labeling. Fibrils were prepared by incubating 1.0 mM monomeric PI3-SH3 in aqueous solution at pH 2.0 for two to three weeks at 25 °C. The fibril morphology was monitored by transmission electron microscopy. After fibrillization, the fibrils were dispersed in a 60/40 (v/v) d$_5$-glycerol/water mixture, to serve as cryoprotectant, and concentrated again by ultracentrifugation. The total amount of protein packed in the MAS rotor was approximately 8 mg. Hydration levels remained constant during the experiments, as verified by rotor weight and ^{1}H water signal.

11.4.2 MAS NMR experiments

Solid-state NMR experiments were performed using a custom-designed spectrometer operating at 700 MHz ^{1}H Larmor frequency (courtesy of Dr. David J. Ruben, Francis Bitter Magnet Laboratory) equipped with a 3.2 mm Varian/Chemagnetics MAS probe (Palo Alto, CA). Correlation experiments were performed at 12.0 kHz spinning frequency and consisted of a 1.2 ms ^{1}H-^{13}C CP period followed by chemical shift evolution, mixing, and detection. TPPM decoupling[29] with an 83 kHz rf field was applied during evolution and detection periods. BASE RFDR was implemented via ^{13}C π pulses with a 12 kHz rf field applied in a 32-step XY sequence,[15] with concurrent 80 kHz ^{1}H continuous-wave decoupling. Each of the 384 total indirect points was the average of 16 scans, with a 3 s recycle delay, for a total measurement time of 5.1 hrs per 2D experiment. Similar acquisition parameters were used to record PDSD spectra. One-dimensional spectra are the average of 128 scans.

11.4.3 Temperature control

Variable-temperature (VT) experiments were performed using a stream of cooled N$_2$. The actual sample temperature deviates from the reading given by the temperature sensor located near the probe head due to diffusive losses and also to heating from rotor friction and prolonged high-power rf irradiation. In order to estimate this deviation, we measured the melting points of samples in water and in glycerol/water by observing line broadening transitions in ^{1}H spectra. For the VT flow rate and MAS frequency we employed in our solid-state NMR experiments,

these melting points were reached with VT target temperatures roughly 10° C lower than their nominal values. Therefore, the sample temperatures reported here correspond to readings of target temperatures 10° C lower.

Acknowledgements

References

1.   Rothwell, W. P.; Waugh, J. S., Transverse relaxation of dipolar coupled spin systems under rf irradiation: Detecting motions in solids. *J. Chem. Phys.* **1981,** 74, (5), 2721-2732.

2.   Long, J. R.; Sun, B. Q.; Bowen, A.; Griffin, R. G., Molecular dynamics and magic angle spinning NMR. *J. Am. Chem. Soc.* **1994,** 116, (26), 11950-11956.

3.   Maus, D. C.; Copie, V.; Sun, B. Q.; Griffiths, J. M.; Griffin, R. G.; Luo, S. F.; Schrock, R. R.; Liu, A. H.; Seidel, S. W.; Davis, W. M.; Grohmann, A., A solid-state NMR study of tungsten methyl group dynamics in [W(eta(5)-C(5)Me(5))Me(4)][PF6]. *J. Am. Chem. Soc.* **1996,** 118, (24), 5665-5671.

4.   Perry, A.; Stypa, M. P.; Tenn, B. K.; Kumashiro, K. K., Solid-state C-13 NMR reveals effects of temperature and hydration on elastin. *Biophys. J.* **2002,** 82, (2), 1086-1095.

5.   Cady, S. D.; Goodman, C.; Tatko, C. D.; DeGrado, W. F.; Hong, M., Determining the orientation of uniaxially rotating membrane proteins using unoriented samples: A H-2, C-13, and N-15 solid-state NMR investigation of the dynamics and orientation of a transmembrane helical bundle. *J. Am. Chem. Soc.* **2007,** 129, (17), 5719-5729.

6.   Bajaj, V. S.; van der Wel, P. C. A.; Griffin, R. G., Observation of a low-temperature, dynamically driven structural transition in a polypeptide by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2009,** 131, (1), 118-28.

7.   Saito, H.; Tsuchida, T.; Ogawa, K.; Arakawa, T.; Yamaguchi, S.; Tuzi, S., Residue-specific millisecond to microsecond fluctuations in bacteriorhodopsin induced by disrupted or disorganized two-dimensional crystalline lattice, through modified lipid-helix and helix-helix interactions, as revealed by C-13 NMR. *Biochim. Biophys. Acta, Biomembr.* **2002,** 1565, (1), 97-106.

8.   Kloepper, K. D.; Zhou, D. H.; Li, Y.; Winter, K. A.; George, J. M.; Rienstra, C. M., Temperature-dependent sensitivity enhancement of solid-state NMR spectra of alpha-synuclein fibrils. *J. Biomol. NMR* **2007,** 39, (3), 197-211.

9.   Yang, J.; Tasayco, M. L.; Polenova, T., Dynamics of Reassembled Thioredoxin Studied by Magic Angle Spinning NMR: Snapshots from Different Time Scales. *J. Am. Chem. Soc.* **2009,** 131, (38), 13690-13702.

10. Bodner, M. L.; Gabrys, C. M.; Parkanzky, P. D.; Yang, J.; Duskin, C. A.; Weliky, D. P., Temperature dependence and resonance assignment of C-13 NMR spectra of selectively and uniformly labeled fusion peptides associated with membranes. *Magn. Reson. Chem.* **2004,** 42, (2), 187-194.

11. Franks, W. T.; Zhou, D. H.; Wylie, B. J.; Money, B. G.; Graesser, D. T.; Frericks, H. L.; Sahota, G.; Rienstra, C. M., Magic-angle spinning solid-state NMR spectroscopy of the beta1 immunoglobulin binding domain of protein G (GB1): 15N and 13C chemical shift assignments and conformational analysis. *J. Am. Chem. Soc.* **2005,** 127, (35), 12291-305.

12. Ishii, Y.; Terao, T.; Hayashi, S., Theory and simulation of vibrational effects on structural measurements by solid-state nuclear magnetic resonance. *J. Chem. Phys.* **1997,** 107, (8), 2760-2774.

13. Schneider, R.; Seidel, K.; Etzkorn, M.; Lange, A.; Becker, S.; Baldus, M., Probing molecular motion by double-quantum (13C,13C) solid-state NMR spectroscopy: application to ubiquitin. *J. Am. Chem. Soc.* **2010,** 132, (1), 223-33.

14. Meier, B. H., Polarization transfer and spin diffusion in solid-state NMR. In *Adv. Magn. Opt. Reson.*, Warren, W. S., Ed. 1999; Vol. 18, p 1.

15. Bayro, M. J.; Maly, T.; Birkett, N. R.; Dobson, C. M.; Griffin, R. G., Long-range correlations between aliphatic 13C nuclei in protein MAS NMR spectroscopy. *Angew. Chem. Int. Ed.* **2009,** 48, (31), 5708-10.

16. LeMaster, D.; Kushlan, D., Dynamical Mapping of E. coli Thioredoxin via 13C NMR Relaxation Analysis. *J. Am. Chem. Soc.* **1996,** 118, (39), 9255-9264.

17. Bayro, M. J.; Ramachandran, R.; Caporini, M. A.; Eddy, M. T.; Griffin, R. G., Radio frequency-driven recoupling at high magic-angle spinning frequencies: Homonuclear recoupling sans heteronuclear decoupling. *J. Chem. Phys.* **2008,** 128, (5), 052321.

18. Bayro, M. J.; Huber, M.; Ramachandran, R.; Davenport, T. C.; Meier, B. H.; Ernst, M.; Griffin, R. G., Dipolar truncation in magic-angle spinning NMR recoupling experiments. *J. Chem. Phys.* **2009,** 130, 114506.

19. Gardiennet, C.; Loquet, A.; Etzkorn, M.; Heise, H.; Baldus, M.; Böckmann, A., Structural constraints for the Crh protein from solid-state NMR experiments. *J. Biomol. NMR* **2008,** 40, (4), 239-50.

20. Seidel, K.; Etzkorn, M.; Heise, H.; Becker, S.; Baldus, M., High-resolution solid-state NMR studies on uniformly [13C,15N]-labeled ubiquitin. *ChemBioChem* **2005,** 6, (9), 1638-47.

21. Polverino de Laureto, P.; Taddei, N.; Frare, E.; Capanni, C.; Costantini, S.; Zurdo, J.; Chiti, F.; Dobson, C. M.; Fontana, A., Protein aggregation and amyloid fibril formation by an SH3 domain probed by limited proteolysis. *J. Mol. Biol.* **2003,** 334, (1), 129-41.

22. Ishii, Y., 13C-13C dipolar recoupling under very fast magic angle spinning in solid-state nuclear magnetic resonance: Applications to distance measurements, spectral assignments, and high-throughput secondary-structure determination. *J. Chem. Phys.* **2001,** 114, (19), 8473-8483.

23. Grommek, A.; Meier, B. H.; Ernst, M., Distance information from proton-driven spin diffusion under MAS. *Chem. Phys. Lett.* **2006,** 427, 404-409.

24. Thurber, K. R.; Tycko, R., Biomolecular solid state NMR with magic-angle spinning at 25K. *J. Magn. Reson.* **2008,** 195, (2), 179-86.

25. Barnes, A. B.; Mak-Jurkauskas, M. L.; Matsuki, Y.; Bajaj, V. S.; van der Wel, P. C.; Derocher, R.; Bryant, J.; Sirigiri, J. R.; Temkin, R. J.; Lugtenburg, J.; Herzfeld, J.; Griffin, R. G., Cryogenic sample exchange NMR probe for magic angle spinning dynamic nuclear polarization. *J. Magn. Reson.* **2009,** 198, (2), 261-70.

26. Maly, T.; Debelouchina, G. T.; Bajaj, V. S.; Hu, K. N.; Joo, C. G.; Mak-Jurkauskas, M. L.; Sirigiri, J. R.; van der Wel, P. C.; Herzfeld, J.; Temkin, R. J.; Griffin, R. G., Dynamic nuclear polarization at high magnetic fields. *J. Chem. Phys.* **2008,** 128, (5), 052211.

27. Barnes, A. B.; De Paëpe, G.; van der Wel, P. C. A.; Hu, K.; Joo, C.; Bajaj, V. S.; Mak-Jurkauskas, M. L.; Sirigiri, J. R.; Herzfeld, J.; Temkin, R. J.; Griffin, R. G., High-field dynamic nuclear polarization for solid and solution biological NMR. *Appl. Magn. Reson.* **2008,** 34, (3-4), 237-263.

28. Guijarro, J. I.; Sunde, M.; Jones, J. A.; Campbell, I. D.; Dobson, C. M., Amyloid Fibril Formation by an SH3 Domain. *Proc. Natl. Acad. Sci. USA* **1998,** 95, 4224-4228.

29. Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G., Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* **1995,** 103, (16), 6951-6958.

Chapter 12. Cryogenic MAS NMR Spectroscopy of Amyloid Fibrils: Peptide Dynamics, Dipolar Recoupling, and Dynamic Nuclear Polarization

Summary

We investigate the influence of temperature on the spectral resolution, molecular conformation, structural integrity and dynamics of amyloid fibrils by conducting MAS NMR experiments at cryogenic temperatures enhanced by dynamic nuclear polarization (DNP). DNP utilizes the inherently larger polarization of electrons to enhance the sensitivity of conventional solid-state NMR experiments at low temperature. Recent advances in instrumentation development and sample preparation have transformed this field and have opened up new opportunities for its application to biological systems. Here, we present DNP-enhanced ^{13}C-^{13}C and ^{15}N-^{13}C correlation spectra of GNNQQNY amyloid fibrils acquired at 9.4 T and 100 K, and demonstrate that DNP can be used to obtain resonance assignments and site-specific structural information very efficiently. In addition, we assess the low-temperature performance of two commonly used solid-state NMR correlation experiments, PDSD and TEDOR, and discuss their potential as tools for measurement of structurally relevant distances at low temperatures in combination with reliable and reproducible DNP enhancement.

12.1 Introduction

In recent years, dynamic nuclear polarization (DNP) has gained ground as a highly promising method to increase the sensitivity of MAS solid-state NMR experiments.[1, 2] The technique utilizes the inherently larger polarization of electrons that are introduced in the sample in the form of biradical, typically TOTAPOL.[3, 4] The biradical's EPR spectrum is irradiated with microwaves at cryogenic temperatures and the electron polarization is transferred to the ^1H nuclei via the cross effect.[5, 6, 7] This is a three-spin mechanism that involves a spin flip-flop process of the two dipolar coupled electrons in the biradical and the generation of ^1H nuclear polarization when the following matching condition is satisfied $\omega_{0S1} - \omega_{0S2} = \omega_{0I}$, where ω_{0S1}, ω_{0S2} and ω_{0I} are the Larmor frequencies of the two electrons and the ^1H spin respectively. The theoretical enhancement for ^1H nuclei is ~ 660, while experimental enhancements on the order of 250 at 5 T and 100 at 9.0 T have been achieved.[8, 9]

Recent advances in instrumentation such as stable low-temperature MAS probes with sample insert/eject systems,[8, 10] reliable high-field microwave sources (gyrotrons)[11, 12] and improved cryogenic technology[13] have transformed the field of DNP and have placed the routine collection of multidimensional correlation experiments on various biological systems within reach.[14, 15] Here, we present DNP-enhanced ^{13}C-^{13}C and ^{15}N-^{13}C correlation experiments on GNNQQNY nanocrystals and amyloid fibrils acquired at 9.4 T and 100 K and demonstrate that DNP can be used to obtain assignments and site-specific structural information very efficiently.

GNNQQNY nanocrystals have previously been used as a model system to demonstrate the transfer of DNP-enhanced ^1H polarization across a glassy matrix-crystal interface at 5 T.[16] This 7-residue fragment can form two types of crystals (monoclinic and orthorhombic) and their crystal structures reveal the presence of the typical amyloid "cross-beta" motif and a steric zipper that was suggested to be the basis of amyloid fibrils in general[17, 18]. The steric zipper forms a dry interface while a network of water molecules forms a 0.7 nm wide channel that comprises the wet interface. TOTAPOL is too large to reside inside the water channel, and it was shown that it does not interact with the bulk molecules inside the monoclinic crystals.[16] ^1H-^1H spin diffusion, however, successfully transfers the polarization through the crystal interface and leads to an enhancement of $\sim$120 at 5 T. Although this enhancement allowed the rapid acquisition of a two-dimensional ^{13}C-^{13}C

correlation spectrum, the low magnetic field limited the spectral resolution and precluded the assignments of the resonances in the crystals.

Under slightly different experimental conditions, GNNQQNY can also form amyloid fibrils. ^{13}C-^{13}C and ^{15}N-^{13}C MAS spectra of these fibrils reveal the presence of three distinct conformational subsets that differ from the crystals both in their chemical shifts and mobility of the side chains.[19] One of these forms, in particular, exhibits significant non-β sheet character, which raises interesting questions regarding the fibril architecture. Obtaining long-range distances that could potentially establish the intermolecular arrangement in the fibrils, however, presents a significant challenge in this relatively small system due to the very degenerate amino acid composition and the multiplicity exhibited in the solid-state NMR spectra.

Here, we use DNP-enhanced ^{13}C-^{13}C and ^{15}N-^{13}C MAS correlation spectra of GNNQQNY amyloid-like fibrils to obtain resonance assignments of this system at 100 K and compare spectral characteristics with those observed at room temperature. The low temperature behavior of peptide systems can be affected by a variety of factors including but not limited to changes in the internal dynamics of the peptide backbone and side chains, structural rearrangements due to interactions with the solvent, and freezing out of disordered or heterogeneous regions of the molecule. These factors can lead to significant deviations in the chemical shifts compared to room temperature and can have profound effects on the line widths and resolution observed in the DNP-enhanced low temperature spectra. We evaluate the extent that these temperature-dependent effects are manifested in this peptide fibril system. Furthermore, we compare the performance of two commonly used solid-state MAS NMR experiments, PDSD[20] and TEDOR,[21; 22] to room temperature data and discuss their potential as tools for measurement of structurally relevant distances at low temperature in combination with DNP signal enhancement.

12.2 Materials and Methods

12.2.1 Sample preparation

Monoclinic crystals of 100% [U-^{13}C,^{15}N GNNQ]QNY and amyloid fibrils of 100% GNN[U-^{13}C,^{15}N QQN]Y peptide (New England Peptide, Gardner, MA) were obtained by following the previously described protocol.[19; 23] The crystals were suspended in 70/23/7 % (w/w/w) glycerol-d$_8$/D$_2$O/H$_2$O matrix containing 35 mM TOTAPOL[4] and centrifuged directly into a 3.2 mm Bruker sapphire rotor using a tabletop centrifuge (16,000 x g). The fibrils were placed in a 70/23/7 % (w/w/w) glycerol-d$_8$/D$_2$O/H$_2$O matrix containing 10 mM TOTAPOL and pelleted at 265,000 x g. The pellet was subsequently transferred into the rotor using a tabletop centrifuge. The [U-^{13}C,^{15}N GNNQ]QNY monoclinic crystal sample will be referred to as "GNNQqny" in the following sections, while the GNN[U-^{13}C,^{15}N QQN]Y fibrils will be abbreviated as "gnnQQNy". The labeling of the residues in the peptide follows the nomenclature used in ref. 19 and the fibril forms a, b and c in this work correspond to published forms 1, 2 and 3, respectively.[19] ^{13}C chemical shifts were referenced to aqueous DSS[24], while ^{15}N chemical shifts were referenced indirectly to liquid NH$_3$.[25; 26]

12.2.2 DNP Experiments at 100 K

MAS NMR experiments enhanced by dynamic nuclear polarization experiments were performed on a Bruker 263 GHz Solids DNP Spectrometer, consisting of a 263 GHz continuous-wave gyrotron source, microwave transmission line, 3.2 mm low temperature (100 K) MAS probe, gas cooling supply, and 400 MHz AVANCE III wide-bore NMR system.[10] 2.5 μs ^{1}H, 4.5 μs ^{13}C and 5.0 μs ^{15}N π/2 pulses were used for the 1D CP experiments and the 2D PDSD[20] and TEDOR[22] correlation experiments described below. 100 kHz ^{1}H TPPM decoupling irradiatoins was applied during the acquisition, evolution and TEDOR mixing periods. The recycle delay was set to 6 s for the crystal sample, and 6.5 s for the fibril sample. Each TEDOR experiment was acquired with 32 scans per t_1 point, 96 t_1 points and a dwell time of 111 μs (with total acquisition times of ~5 hr for the crystals and ~5.5 hr for the fibrils). The PDSD experiments were acquired with 4 scans per t_1 point , 384 t_1 points, and dwell time of 56 μs (with a total acquisition time of ~3 hr each).

12.2.3 Experiments at 300 K

Room temperature MAS NMR experiments were performed on custom-designed spectrometers (courtesy of D. J. Ruben, Francis Bitter Magnet Laboratory, Massachusetts Institute of Technology, Cambridge, MA) operating at 750 and 700 MHz ^{1}H Larmor frequency. The 750 MHz spectrometer was equipped with a 3.2 mm triple-resonance ^{1}H/^{13}C/^{15}N Bruker E^{free} probe (Billerica, MA), and the 700 experiments were performed on a 3.2 mm triple resonance Varian-Chemagnetics probe (Palo Alto, CA). The sample temperature during the MAS experiments was estimated to be ~300 K. The recycle delay was set to 3 s for all experiments at this temperature. A 2D TEDOR experiment was collected with the fibril sample at 750 MHz, spinning frequency of $\omega_r/2\pi$ = 9 kHz, mixing time of t_{mix} = 1.8 ms, 128 scans per t_1 point, 96 t_1 points and a dwell time of 111 μs, for a total acquisition time of ~10 hrs. A PDSD experiment was collected with a longitudinal mixing time of t_{mix} = 10 ms, 28 scans per t_1 point, 668 t_1 points, and a dwell time of 24 μs, for a total acquisition time of ~16 hrs.
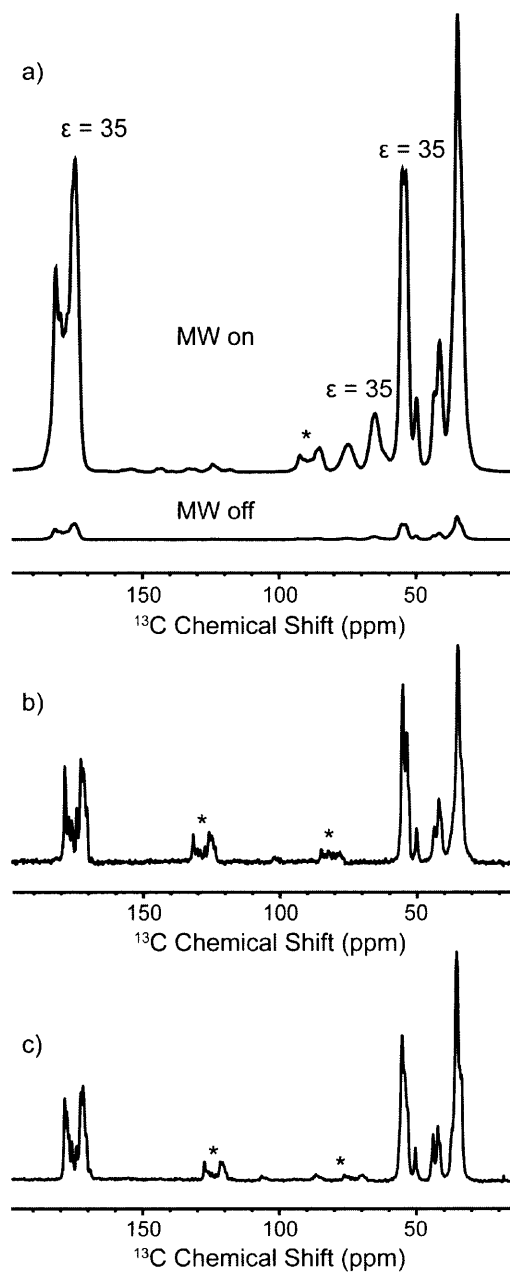
12.3 Results and Discussion

12.3.1 DNP enhancement

Solid-state MAS NMR spectra of gnnQQNy amyloid fibrils exhibited a DNP enhancement factor of 35, as shown in Figure 1a. Unlike the case of the crystals, the glycerol resonances in the ^{13}C CP spectra are enhanced by an identical factor as the fibril resonances. In addition, fibril and solvent signals present similar polarization build-up rates (5 s for the fibrils, and 5.5 s for glycerol). This behavior indicates that the biradical molecules, dispersed in the glycerol/water solvent matrix, are likely to be also in intimate contact with the fibril molecules. The small width of the fibrils (a few nanometers in their thinnest dimension)[19] allows close contact between the paramagnetic centers and a large fraction of the peptide molecules. Furthermore, enhanced ^{1}H polarization is spread via spin diffusion and may equilibrate rapidly throughout the fibril. Indeed, rapid polarization equilibration via spin diffusion is likely to be the reason why the DNP enhancement factor for the fibrils is considerably larger than for the much thicker crystals, despite a lower radical concentration in the fibrils (10 mM) than in the crystals (35 mM). The lower radical concentration, however, leads to the sightly lower enhancement observed for the glycerol resonances and their longer polarization buildup time.

12.3.2 Temperature effects on 1D spectra

Given the apparent close proximity between paramagnetic centers and fibril molecules, it is remarkable that high magnetic field spectra of gnnQQNy fibrils prepared with (Figure 1b) and without biradicals (Figure 1c) are equally well resolved, demonstrating that the resonance broadening effect of TOTAPOL biradicals at this concentration (10 mM) is marginal for these fibril samples. In addition, the room temperature spectrum in Fig. 1b was recorded on the same sample as that in Figure 1a after performing our DNP experiments at 100 K. Since both room temperature spectra (Figures 1b and 1c) present virtually no chemical shift discrepancies, it is evident that the fibrils experience little damage as a result of the low temperatures, likely due to adequate cryoprotection by the glycerol/water solvent. While it is clear that 400 MHz spectra recorded at 100 K with DNP present significantly larger line widths than those of 750 MHz spectra recorded at room temperature, chemical shift variations are very small, as described in detail below, which indicates that the fibrils preserve their integrity while at low temperatures.

Figure 1. (a) ^{13}C CP spectra of GNN[U-^{13}C,^{15}N QQN]Y fibrils obtained with DNP (top) and without DNP (bottom) at 400 MHz, $\omega_r/2\pi$ = 9 kHz, and 100 K, with a MW irradiation of 6.5 s and as the average of 16 scans. The enhancement factor fot the fibrils is 35, and it is the same for the glycerol resonances. (b) ^{13}C CP spectrum obtained with the same sample after the DNP experiments were performed, recorded at 750 MHz. (c) ^{13}C CP spectrum of fibril samples prepared without radical and glycerol, and never taken to low temperatures, recorded at 700 MHz. Experimental parameters for (b) and (c) are $\omega_r/2\pi$ = 9 kHz and 300 K, recorded as the average of 128 scans with recycle delay of 3 s. Asteriks denote side bands.

12.3.3 Homonuclear correlation spectra

Magic-angle spinning homonuclear ^{13}C-^{13}C correlation spectra of gnnQQNy fibrils recorded with PDSD mixing present strong cross-peaks with multiple resolved ^{13}C sites. Figure 2 compares a 2D PDSD spectrum recorded at 750 MHz and 300 K without DNP (Figure 2a) to one recorded at 400 MHz and 100 K with DNP (Figure 2b). The line width differences observable in 1D spectra are more conspicuous in these 2D spectra. For example, the ^{13}C line widths for the resolved Asn12 cross-peaks at room temperature are 0.8-1 ppm, while they are 1.3-1.75 ppm at 100 K. On the other hand, the DNP-enhanced spectrum presents a number of resolved sites and allows the identification of various spin systems without referring to data at room temperature. For example, one of the fibril forms, here referred to as "form a", has been completely assigned with the low-temperature spectra presented here, while the three different forms of Asn12 are easy to identify.

Longitudinal mixing efficiency appears to improve at low temperature, as evidenced by increased cross-peak intensities. In addition, polarization transfer appears to occur at a faster rate at 100 K, with two-bond correlations appearing strongly after 5 ms of mixing while they are still relatively weak after 10 ms of mixing at room temperature. Such dynamic effects manifested as a low-temperature enhancement in mixing efficiency are advantageous from a practical perspective, but they might also be critical in the study of complex biomolecules that exhibit extensive dynamics. For example, form b of GNNQQNY fibrils appears to be more dynamic than forms a and c at room temperature, as evidenced by the lower CP signal intensity for residue N12 of form b (Fig. 5b and c), and ^{13}C and ^{15}N relaxation experiments.[27] The cross-peak corresponding to N12b CA-CO is at the noise level in the room temperature spectrum (Figure 2a) while other N12b cross-peaks do appear, indicating a degree of dynamic disorder. On the other hand N12b CA-CO is observed with sufficient intensity in the DNP-enhanced PDSD spectrum at 100 K (Figure 2b).
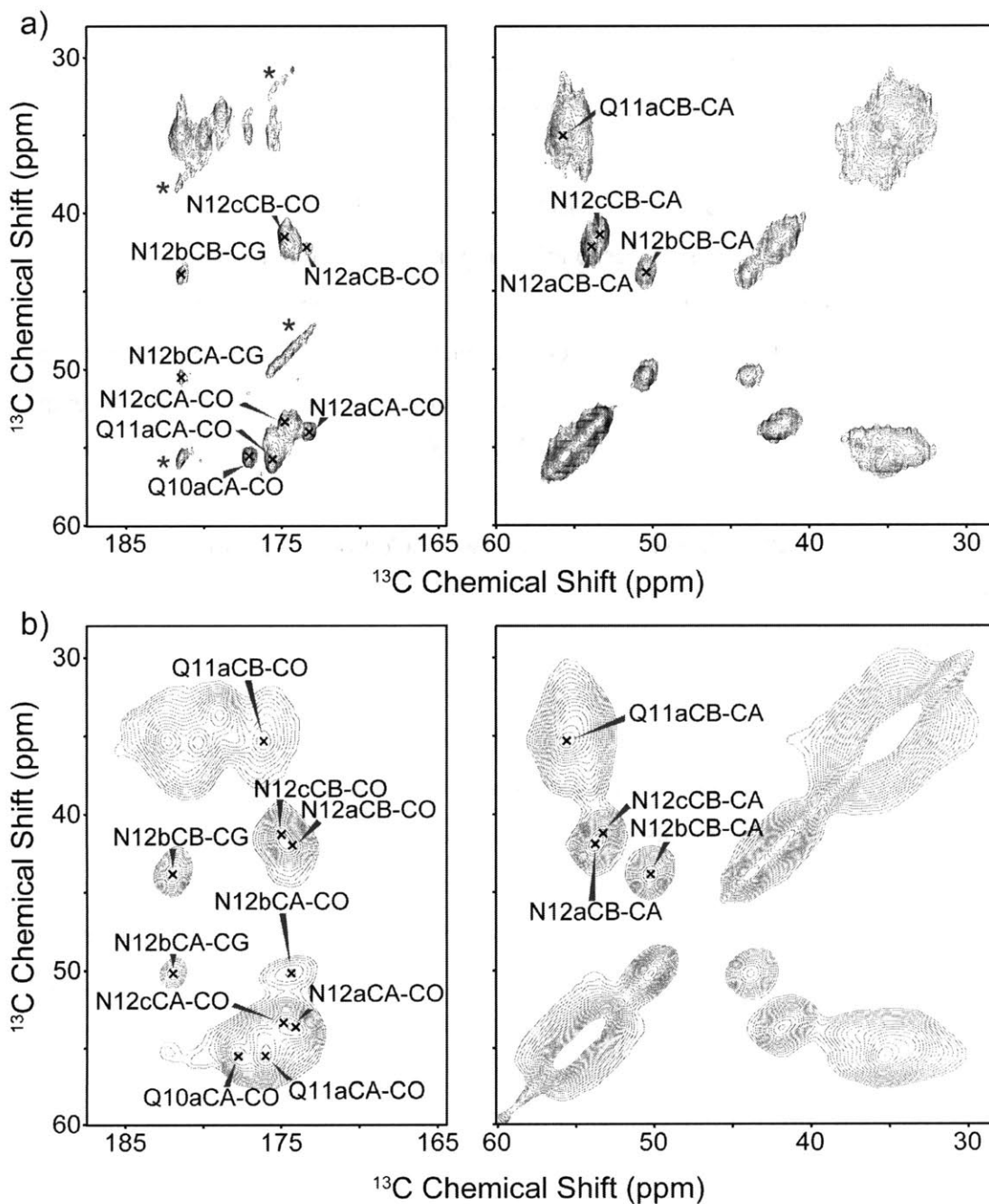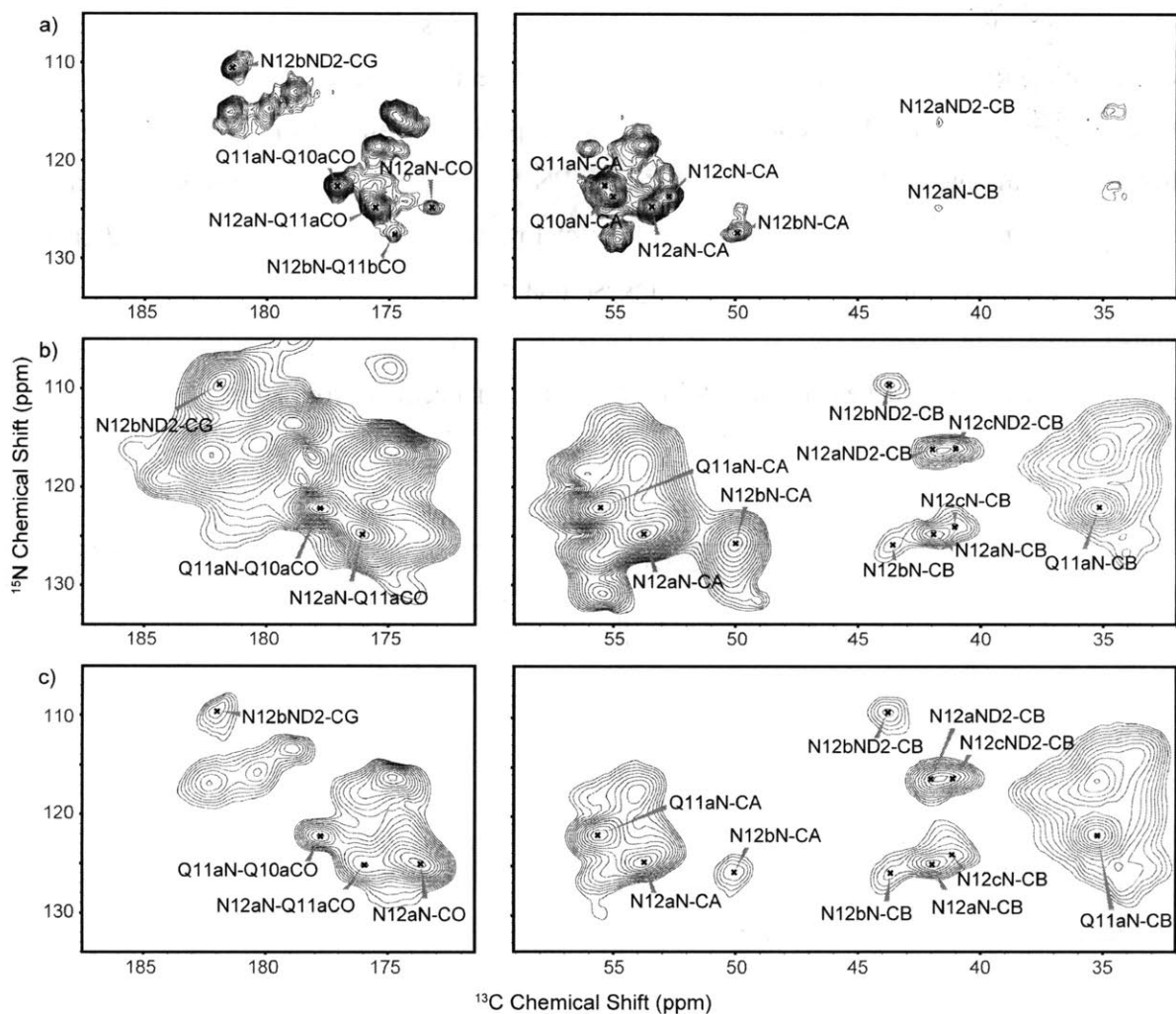
Figure 2. ^{13}C-^{13}C correlations of GNN[U-^{13}C,^{15}N QQN]Y fibrils collected at (a) 750 MHz without DNP, using PDSD τ_{mix} = 10 ms, and at $\omega_r/2\pi$ = 9 kHz, 300 K, and (b) 400 MHz with DNP, using PDSD τ_{mix} = 5 ms, and at $\omega_r/2\pi$ = 9 kHz, 100 K. The asterisks in (a) denote side bands.

12.3.4 Heteronuclear correlation spectra

The influence of temperature on the dynamics of the fibrils is also readily observed in heteronuclear correlation spectra. The TEDOR spectra presented in Figure 3 depict several well-resolved ^{15}N-^{13}C correlations at 750 MHz and room temperature (Fig. 7a) and at 400 MHz and 100 K with DNP (Figures 3b and 3c). TEDOR recorded with t_{mix} = 1.8 ms is optimal for one-bond heteronuclear correlations, corresponding in gnnQQNy to five backbone (three N_i-CA_i and two N_i-CO_{i-1}) and three side-chain contacts for each of the three forms of fibrils. At this short mixing time, the room temperature spectrum presents almost exclusively one-bond correlations, with two-bond contacts slightly above the noise level. On the other hand, the low-temperature, DNP-enhanced 1.8 ms TEDOR spectrum of Figure 3b, presents very intense two-bond ^{15}N-^{13}C correlations in addition to the expected one-bond contacts. The overall sensitivity of the DNP-enhanced TEDOR spectrum is excellent, given that in addition to the polarization enhancement factor, the recoupling efficiency itself at low temperature is comparable or superior to that at room temperature. Increasing the mixing time allows the build-up of two-bond correlations, shown at their approximate maximum in Figure 3c, which was recorded with t_{mix} = 3.6 ms.

We recorded several 2D DNP-enhanced TEDOR spectra of gnnQQNy fibrils with t_{mix} ranging from 0.9 ms to 7.2 ms, from which several build-up curves were extracted, plotted in Fig. 8a. Comparison of ^{13}C signals after CP and after TEDOR with a short mixing time in 1D allows us to evaluate the efficiency of one-bond recoupling. The corresponding cross-peaks in 2D spectra can then be compared to other cross-peaks to obtain an estimate of their recoupling efficiency, taking into account the original CP intensity of the nuclei involved. One-bond transfer efficiencies were ~20%, while efficiencies around 7% were obtained for two-bond transfers. These numbers are comparable to theoretical values, however, it is apparent in Figure 4 that the two-bond build-up curves decay somewhat quickly, after 4-5 ms, which is consistent with rapid transverse relaxation. Indeed, ^{13}C echo experiments at 100 K with high-power ^{1}H decoupling show T_1 times of ~5 ms for the fibrils. Since the TEDOR pulse sequence contains two echo segments in the ^{13}C channel during the ^{15}N-^{13}C recoupling period, significant polarization losses can be expected at increasing mixing times.

**Figure 3.** ^{15}N-^{13}C correlation experiments of GNN[U-^{13}C,^{15}N QQN]Y fibrils recorded with TEDOR mixing. (a) Spectrum recorded at 750 MHz, $\omega_r/2\pi$ = 9 kHz, 300 K and τ_{mix} = 1.8 ms. (b) and c) Spectra recorded with DNP at 400 MHz, $\omega_r/2\pi$ = 9 kHz, 100 K and τ_{mix} = 1.8 ms and 3.6 ms, respectively.

12.3.5 Dipolar recoupling and molecular dynamics

As a comparison to our DNP-enhanced, low-temperature TEDOR build-up curves, we recorded similar curves at room temperature, shown in Fig. 8b. Due to the low signal to noise ratio without DNP, we recorded the room temperature profiles in 1D fashion. A few interesting differences are readily observed. First, the decay of CA magnetization is much slower than at 100 K, consistent with reduced relaxation processes for the fibrils at room temperature. The decay of ^{13}C signals after polarization build-up in a TEDOR experiment is driven by both

relaxation and the strength of the effective ^{15}N-^{13}C couplings. The increased dynamics at room temperature, however, may also be responsible for a decreased recoupling efficiency for two-bond contacts (2.5 to 4%) compared to low-temperature data (7%). A third noteworthy difference is that while ^{13}C nuclei with the same ^{15}N-^{13}C coupling strengths (either one- or two-bond couplings) have similar TEDOR profiles at low temperature, they vary considerably at room temperature. These observations suggest that measuring internuclear distances at low temperatures with experiments such as TEDOR may have significant advantages over room temperature experiments besides the additional signal enhancement from DNP.
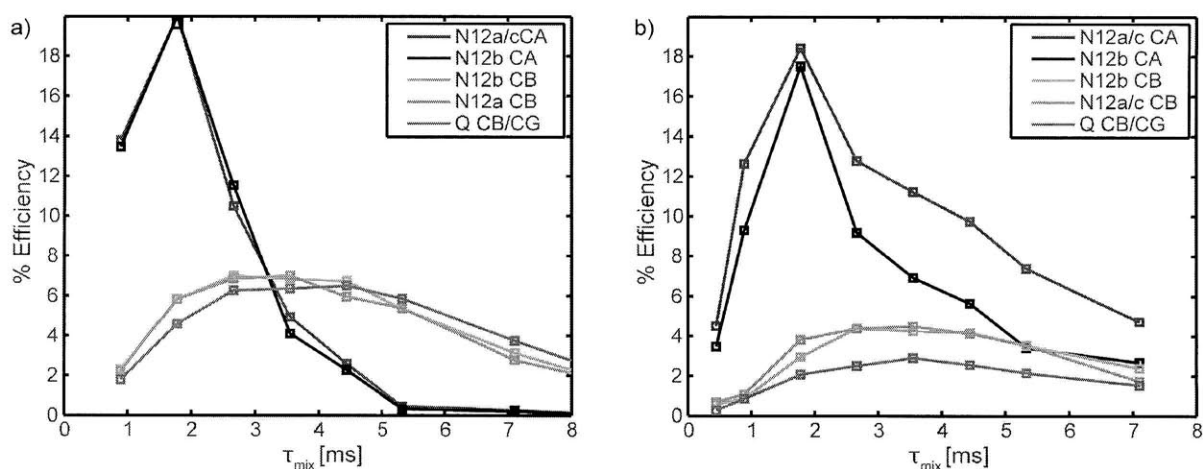
Figure 4. 1D TEDOR buildup curves of GNN[U-^{13}C,^{15}N QQN]Y fibrils recorded (a) with DNP at 400 MHz ^1H Larmor frequency and a temperature of 100 K, and (b) without DNP at 750 MHz ^1H Larmor frequency and a temperature of 300 K.

12.4 Conclusions

We have obtained high-quality DNP-enhanced solid-state MAS NMR spectra of GNNQQNY nanocrystals and amyloid fibrils, recorded at 100 K and 9.4 T. DNP leads to significant enhancements in the signal-to-noise ratio (20 for the nanocrystals and 35 for the fibril sample), allowing the acquisition of two dimensional correlation spectra with great time efficiency. The higher magnetic field utilized in this study has allowed us to record well-resolved low temperature PDSD and TEDOR spectra and to obtain assignments for all labeled residues in the crystal sample and for a large number of the labeled sites in the fibrils. Chemical shift and line width comparison to room temperature data indicates that the residues in close contact to the water molecules in the crystals undergo profound structural changes that can lead to inhomogeneous broadening. While little can be done in such cases where the sources of line broadening are inherent to the system, more work is necessary to understand the temperature response of these systems. In the fibril case, on the other hand, the chemical shifts change very little with temperature and while the line widths increase uniformly by 40-60 Hz, this process is reversible. Work with other fibril systems is currently in progress to assess the generality of this behavior.

The enhanced sensitivity due to DNP, combined with diminished dynamic processes at low temperature that could otherwise interfere with recoupling efficiency, lead to PDSD and TEDOR correlation spectra that are much richer in information content. This is especially the case for shorter mixing times where relaxation losses are less pronounced. Improvements in the sample preparation technique and cryoprotection could lead to samples with enhanced resolution and relaxation properties. Such developments will undoubtedly open new avenues for the application of DNP in biomolecular solid-state NMR investigations.

References

1. Maly, T., Debelouchina, G. T., Bajaj, V. S., Hu, K.-N., Joo, C.-G., Mak-Jurkauskas, M. L., Sirigiri, J. R., van der Wel, P. C. A., Herzfeld, J., Temkin, R. J. & Griffin, R. G. (2008). Dynamic nuclear polarization at high magnetic fields. *J. Chem. Phys.* **128**, 052211.

2. Barnes, A. B., De Paepe, G., van der Wel, P. C. A., Hu, K. N., Joo, C. G., Bajaj, V. S., Mak-Jurkauskas, M. L., Sirigiri, J. R., Herzfeld, J., Temkin, R. J. & Griffin, R. G. (2008). High-field dynamic nuclear polarization for solid and solution biological NMR. *Appl. Magn. Reson.* **34**, 237.

3. Hu, K. N., Song, C., Yu, H., Swager, T. M. & Griffin, R. G. (2008). High-frequency dynamic nuclear polarization using biradicals: A multifrequency EPR lineshape analysis. *J. Chem. Phys.* **128**, 052302.

4. Song, C., Hu, K. N., Joo, C. G., Swager, T. M. & Griffin, R. G. (2006). TOTAPOL: A Biradical Polarizing Agent for Dynamic Nuclear Polarization Experiments in Aqueous Media. *J. Am. Chem. Soc.* **128**, 11385.

5. Kessenikh, A. V., Manenkov, A. A. & Pyatnitskii, G. I. (1964). On Explanation of Experimental Data on Dynamic Polarization of Protons in Irradiated Polyethylenes. *Sov. Phys. - Sol. State* **6**, 641.

6. Hwang, C. F. & Hill, D. A. (1967). Phenomenological Model for the New Effect in Dynamic Polarization. *Phys. Rev. Lett.* **19**, 1011.

7. Wollan, D. S. (1976). Dynamic nuclear polarization with an inhomogeneously broadened ESR line. I. Theory. *Phys. Rev. B* **13**, 3671.

8. Barnes, A. B., Mak-Jurkauskas, M. L., Matsuki, Y., Bajaj, V. S., van der Wel, P. C. A., DeRocher, R., Bryant, J., Sirigiri, J. R., Temkin, R. J., Lugtenburg, J., Herzfeld, J. & Griffin, R. G. (2009). Cryogenic sample exchange NMR probe for magic angle spinning dynamic nuclear polarization. *J. Magn. Reson.* **198**, 261.

9. Matsuki, Y., Maly, T., Ouari, O., Karoui, H., Le Moigne, F., Rizzato, E., Lyubenova, S., Herzfeld, J., Prisner, T., Tordo, P. & Griffin, R. G. (2009). Dynamic Nuclear Polarization with a Rigid Biradical13. *Angew. Chem., Int. Ed.* **48**, 4996.

10. Rosay, M. (2010). *Phys. Chem. Chem. Phys.*, DOI:10.1039/c003685.

11. Bajaj, V. S., Farrar, C. T., Hornstein, M. K., Mastovsky, I., Vieregg, J., Bryant, J., Elena, B., Kreischer, K. E., Temkin, R. J. & Griffin, R. G. (2003). Dynamic nuclear polarization at 9T using a novel 250GHz gyrotron microwave source. *J. Magn. Reson.* **160**, 85.

12. Bajaj, V. S., Hornstein, M. K., Kreischer, K. E., Sirigiri, J. R., Woskov, P. P., Mak-Jurkauskas, M. L., Herzfeld, J., Temkin, R. J. & Griffin, R. G. (2007). 250 GHz CW gyrotron oscillator for dynamic nuclear polarization in biological solid state NMR. *J. Magn. Reson.* **189**, 251.

13. Allen, P. J., Creuzet, F., De Groot, H. J. M. & Griffin, R. G. (1991). Apparatus for low-temperature magic-angle spinning NMR. *J. Magn. Reson.* **92**, 614.

14. Bajaj, V. S., Mak-Jurkauskas, M. L., Belenky, M., Herzfeld, J. & Griffin, R. G. (2009). Functional and shunt states of bacteriorhodopsin resolved by 250 GHz dynamic nuclear polarization-enhanced solid-state NMR. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 9244.

15. Mak-Jurkauskas, M. L., Bajaj, V. S., Hornstein, M. K., Belenky, M., Griffin, R. G. & Herzfeld, J. (2008). Energy transformations early in the bacteriorhodopsin photocycle revealed by DNP-enhanced solid-state NMR. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 883.

16. van der Wel, P. C. A., Hu, K. N., Lewandowski, J. & Griffin, R. G. (2006). Dynamic nuclear polarization of amyloidogenic peptide nanocrystals: GNNQQNY, a core segment of the yeast prion protein Sup35p. *J. Am. Chem. Soc.* **128**, 10840.

17.   Nelson, R., Saway, M. R., Balbirnie, M., Madsen, A. O., Riekel, C., Grothe, R. & Eisenberg, D. (2005). Structure of the cross-b spine of amyloid-like fibrils. *Nature* **435**, 773.

18.   Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J. W., McFarlane, H. T., Madsen, A. O., Riekel, C. & Eisenberg, D. (2007). Atomic structures of amyloid cross-[bgr] spines reveal varied steric zippers. *Nature* **447**, 453.

19.   van der Wel, P. C. A., Lewandoswki, J. & Griffin, R. G. (2007). Solid state NMR study of amyloid nanocrystals and fibrils formed by the peptide GNNQQNY from yeast prion protein Sup35p. *J. Am. Chem. Soc.* **129**, 5117.

20.   Szeverenyi, N. M., Sullivan, M. J. & Maciel, G. E. (1982). Observation of spin exchange by two-dimensional Fourier-transform C-13 cross polarization-magic-angle spinning. *J. Magn. Reson.* **47**, 462.

21.   Hing, A. W., Vega, S. & Schaefer, J. (1992). Transferred-echo double resonance NMR. *J. Magn. Reson.* **96**, 205.

22.   Jaroniec, C. P., Filip, C. & Griffin, R. G. (2002). 3D TEDOR NMR Experiments for the Simultaneous Measurement of Multiple Carbon-Nitrogen Distances in Uniformly 13C,15N-Labeled Solids. *J. Am. Chem. Soc.* **124**, 10728.

23.   Balbirnie, M., Grothe, R. & Eisenberg, D. S. (2001). An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated beta-sheet structure for amyloid. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 2375.

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

25.   Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E. & Wuthrich, K. (1998). Recommendations for the presentation of NMR structures of proteins and Nucleic Acids (Reprinted from Pure and Applied Chemistry, vol 70, pgs 117-142, 1998). *J. Mol. Biol.* **280**, 933.

26. Harris, R. K., Becker, E. D., de Menezes, S. M. C., Goodfellow, R. & Granger, P. (2002). NMR nomenclature: Nuclear spin properties and conventions for chemical shifts - IUPAC recommendations 2001 (Reprinted from Pure Appl. Chem, vol 73, pg 1795-1818, 2001). *Solid State Nucl. Magn. Reson.* **22**, 458.

27. van der Wel, P. C. A., Lewandoswki, J. & Griffin, R. G. (2010). Structural characterization of GNNQQNY amyloid fibrils by Magic Angle Spinning NMR. *Biochemistry*, submitted.

Curriculum Vitae

Education
Massachusetts Institute of Technology, Cambridge, MA
PhD in Physical Chemistry, June 2010

Rutgers, The State University of New Jersey, Piscataway, NJ
B.A. in Chemistry and Physics, with honors, May 2003

Research Experience

01/04-06/10   Graduate Research Assistant
Group of Prof. Robert G. Griffin
Department of Chemistry and Francis Bitter Magnet Laboratory,
Massachusetts Institute of Technology
Thesis Topics: Solid-State NMR Methods Development and Applications to
Structural Studies of Protein Assemblies

09/00-07/03   Undergraduate Research Assistant
Group of Prof. Gaetano T. Montelione
Department of Molecular Biology and Biochemistry and Center for Advanced
Biotechnology and Medicine, Rutgers University
Undergraduate Thesis Topic: Development of Automated NMR Data Analysis
Software and Solution NMR Protein Structure Determination.

Publications

Bayro MJ, Maly T, Birkett NR, Dobson CM, Griffin RG. "High-resolution MAS NMR analysis of PI3-SH3 amyloid fibrils: Backbone conformation and implications for protofilament assembly." Submitted.

Sivertsen AC, Bayro MJ, Belenky M, Griffin RG, Herzfeld J. "Solid-state NMR characterization of gas vesicle structure." Biophys J. In press.

Debelouchina* GT, Bayro* MJ, van der Wel PCA, Caporini MA, Barnes AB, Rosay M, Maas WE, Griffin RG, "Dynamic nuclear polarization-enhanced solid-state NMR spectroscopy of GNNQQNY nanocrystals and amyloid fibrils" Phys Chem Chem Phys. In press.

Bayro MJ, Maly T, Birkett NR, Dobson CM, Griffin RG. "Long-range correlations between aliphatic 13C nuclei in protein MAS NMR spectroscopy." Angew Chem Int Ed (2009) 48, 5708-10.

Bayro MJ, Huber M, Ramachandran R, Davenport TC, Meier BH, Ernst M, Griffin RG. "Dipolar truncation in magic-angle spinning NMR recoupling experiments." J Chem Phys (2009) 130, 114506.

Sivertsen AC, Bayro MJ, Belenky M, Griffin RG, Herzfeld J. "Solid-state NMR evidence for inequivalent GvpA subunits in gas vesicles." J Mol Biol (2009) 387, 1032-9.

Lin J, Bayro MJ, Griffin RG, Khaneja N. "Dipolar recoupling in solid state NMR by phase alternating pulse sequences." J Magn Reson (2009) 197, 145-52.

Bayro MJ, Ramachandran R, Caporini MA, Eddy MT, Griffin RG. "Radio frequency-driven recoupling at high magic-angle spinning frequencies: homonuclear recoupling sans heteronuclear decoupling." J Chem Phys (2008) 128, 052321.

De Paëpe G, Bayro MJ, Lewandowski J, Griffin RG. "Broadband homonuclear correlation spectroscopy at high magnetic fields and MAS frequencies." J Am Chem Soc (2006) 128, 1776-7.

Baran MC, Moseley HN, Aramini JM, Bayro MJ, Monleon D, Locke JY, Montelione GT. "SPINS: a laboratory information management system for organizing and archiving intermediate and final results from NMR protein structure determinations." Proteins (2006) 62, 843-51.

Bayro MJ, Mukhopadhyay J, Swapna GV, Huang JY, Ma LC, Sineva E, Dawson PE, Montelione GT, Ebright RH. "Structure of antibacterial peptide microcin J25: a 21-residue lariat protoknot" J Am Chem Soc (2003) 125, 12382-3.

Manuscripts in preparation:

M.J. Bayro, G.T. Debelouchina, N.R. Birkett, C.M. Dobson, R.G. Griffin, "Solid-state NMR identification of a parallel, in-register architecture in PI3-SH3 amyloid fibrils."

M.J. Bayro, M.T. Eddy, G.T. Debelouchina, R.G. Griffin, "A protocol for rapid backbone resonance assignment of proteins in solid-state NMR spectroscopy."

M.J. Bayro, N.R. Birkett, C.M. Dobson, R.G. Griffin, "Low-temperature enhancement of polarization transfer between distant nuclei in protein MAS NMR spectroscopy."

M.J. Bayro, R.G. Griffin, "Efficient recoupling of weak homonuclear dipolar interactions in magic-angle spinning NMR spectroscopy: The BASE RFDR scheme."

M.J. Bayro, G.T. Debelouchina, R.G. Griffin, "Spinning frequency-dependent enhancement of long-range polarization transfer from aromatic nuclei in protein solid-state NMR spectroscopy."