DEVELOPMENT AND APPLICATION OF NITROGEN-14 SOLID-STATE NMR FOR THE ANALYSIS OF BIOLOGICAL MOLECULES

James Alexander Jarvis

THESIS SUBMITTED TO THE CENTRE FOR BIOLOGICAL SCIENCES GRADUATE SCHOOL IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY BY THE UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

CENTRE FOR BIOLOGICAL SCIENCES

UNIVERSITY OF SOUTHAMPTON

OCTOBER 2015

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

CENTRE FOR BIOLOGICAL SCIENCES

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEVELOPMENT AND APPLICATION OF NITROGEN-14 SOLID-STATE NMR FOR THE ANALYSIS OF BIOLOGICAL MOLECULES

JAMES ALEXANDER JARVIS

Solid-state NMR has progressed significantly in recent decades into a powerful tool for the analysis of structure and dynamics in complex biological molecules such as proteins. The vast majority of solid-state NMR experiments on proteins rely on the isotopic enrichment of the spin-½ isotopes, ^{13}C and ^{15}N. However, there is a >99% naturally abundant, NMR active, isotope of nitrogen, ^{14}N, which has a nuclear spin of 1. Its large quadrupole coupling generally precludes its use in biological NMR experiments, as it is challenging to manipulate and detect directly, though the quadrupolar coupling is highly sensitive to the local electrostatic environment.

Recently, a class of solid-state NMR experiments have been introduced where ^{14}N is studied in the indirect dimension of a 2D experiment using a spin-½ nucleus such as ^{13}C or ^{1}H for detection. These experiments greatly benefit from the increased sensitivity and resolution afforded by detecting ^{14}N indirectly, but currently still lack the necessary sensitivity to be generally applicable to complex biomolecules such as proteins. These experiments typically rely on delays for free evolution in order to generate and reconvert coherence between spy nuclei and ^{14}N. In this thesis, a novel variation on this class of experiment is presented whereby coherence is generated by long, moderate amplitude RF fields applied on the ^{14}N channel. The experiment is demonstrated detected on both ^{13}C and ^{1}H on small organic molecules; demonstrating increased ^{14}N transfer efficiencies over previous HMQC-style experiments. It is shown that the technique is robust with respect to ^{14}N RF amplitude and pulse widths, meaning a variety of sites with different ^{14}N environments can be detected simultaneously with minimal optimisation. Furthermore, it is shown that high quality correlation spectra can be acquired with ^{14}N lineshapes that have well-defined features. This allows for a rigorous analysis of ^{14}N lineshapes using numerical simulations to define ^{14}N quadrupolar coupling parameters at a variety of sites. Finally, using the indirect detection method described herein, we demonstrate indirect detection of ^{14}N in a 56 residue microcrystalline protein, GB3. To the best of our knowledge, this is the first solid-state NMR measurement of ^{14}N made in a full-length protein. An analysis of the magnitude and distribution of quadrupolar couplings in the protein has been performed, demonstrating that the 14N quadrupolar coupling is highly sensitive to the local electrostatic environment, with large changes in quadrupolar interaction reflecting subtle differences in hydrogen bonding in different secondary structures.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DECLARATION OF AUTHORSHIP

I, James Alexander Jarvis, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

DEVELOPMENT AND APPLICATION OF NITROGEN-14 SOLID-STATE NMR FOR THE ANALYSIS OF BIOLOGICAL MOLECULES

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

Jarvis, J. A., Haies, I. M., Williamson, P. T., Carravetta, M. An Efficient NMR Method for the Characterisation of ^{14}N Sites Through ^{13}C Indirect Detection. *Phys Chem Chem Phys.* **2013**, *15*, 7613-7620.

Signed:.....

Date: .....

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisors, Phil Williamson and Marina Carravetta, for their constant support and encouragement over the last 4 years. I have been allowed a vast degree of freedom in my research and have been lucky enough to have had the opportunity to travel numerous places with their support to build a range of scientific contacts and experience research in a variety of laboratories. Both have been unbelievably generous with their time, patient and extremely enlightening. I can't think of two better people to receive a practical education in solid-state NMR with.

I would like to thank the numerous people who have helped me in the various labs I've done experiments in over the last 4 years. Here in Southampton, Stuart Findlow and Ole Johannessen have provided constant, expert support in the biochemical and NMR labs that has facilitated most all of the work in this thesis. Ilya Kuprov is acknowledged for his collaboration on simulating ^{14}N experiments, and has provided me with invaluable insight and instruction on NMR simulations. Raivo Stern and Ivo Heinmaa were extremely generous in welcoming me to their lab at the KBFI in Tallinn for low temperature MAS measurements. Dinu Iuga has provided plenty of helpful wisdom and assistance in performing experiments and packing tiny rotors at the UK 850MHz Solid-State NMR Facility at the University of Warwick, and funding for the Facility by the EPSRC, BBSRC and The University of Warwick is gratefully acknowledged. From the CRMN at the ENS de Lyon, I owe Moreno Lelli, Aaron Rossini, David Gajan, Lenaïc Leroux, Anne Lesage and Lyndon Emslely a huge debt of gratitude for assistance with all aspects of the DNP measurements and enlightening discussions. I'm sorry I broke so many rotors. Financial support for the DNP work from the TGIR-RMN-THC Fr3050, and Lyndon Emsley, is gratefully acknowledged.

I'd like to thank all members of the Williamson and Carravetta research groups that I have worked with, past and present, as well as others from the MagRes and structural biology sections, who have provided academic assistance and excellent company over the years I've been in Southampton.

Thanks goes to Phedra Marius and Garrick Taylor, I learnt a lot of from them both in the biochemistry lab. I worked very closely with Ibraheem Haies on the ^{14}N project for four years, and we visited a variety of beautiful places all over the world together to do experiments, or attend conferences. It was always a pleasure, and I wish him all the very best. Richard Bounds undoubtedly made the experience of the last 3 years of my PhD more enjoyable. I thank him for what amounts to many hours of variously enlightening, engaging and hilarious conversation in the pub, and in what is possibly the world's most brightly lit office that we shared for two years. Stuart Findlow was always great conversation. I had fun and learned a lot, not always about science. Andy Hutchin and Matt Rodriguez I thank for all the early morning bouldering sessions. I think they really helped keep me sane in the final months of my PhD.

I save my largest thanks for my family: I thank Emma for her unerring love and support throughout the ups and downs of the PhD. She has kept me on track in many ways during the PhD and I can't thank her enough. Finally, I thank my parents, for their continuous love, support, belief and encouragement that this work surely would not have been possible without.

LIST OF ABBREVIATIONS

AAG: *Alanyl-Alanyl-Glycine*

AGG: *Alanyl-Glycyl-Glycine*

BRAIN-CP: *BRoadband Adiabatic INversion Cross Polarisation*

CE: *Cross Effect*

CP: *Cross Polarisation*

CSA: *Chemical Shift Anisotropy*

CSI: *Chemical Shift Index*

CTP: *Coherence Transfer Pathway*

DANTE: *Delays Alternating with Nutation for Tailored Excitation*

DARR: *Dipolar Assisted Rotational Resonance*

DNP: *Dynamic Nuclear Polarisation*

DQ: *Double Quantum*

EFG: *Electric Field Gradient*

EPR: *Electron Paramagnetic Resonance*

FID: *Free Induction Decay*

FOQI: *First Order Quadrupolar Interaction*

FSLG: *Frequency Switched Lee-Goldberg*

GB3: *Protein G IgG Binding Domain 3*

hIAPP: *Human Islet Amyloid Precursor Protein*

HMQC: *Heteronuclear Multiple Quantum Coherence*

HSQC: *Heteronuclear Single Quantum Coherence*

IgG: *Immunoglobulin G*

INEPT: *Insensitive Nuclei Enhanced by Polarisation Transfer*

IPTG: *Isopropyl β-D-1 ThioGalactopyranoside*

LB: *Luria Broth*

MAS: *Magic Angle Spinning*

MIRROR: *MIxed Rotational and ROtary Resonance*

MQMAS: *Multiple Quantum Magic Angle Spinning*

NAV: *N-Acetyl-Valine*

NMR: *Nuclear Magnetic Resonance*

ORR: *Overtone Rotary Resonance*

PAGE: *Polyacrylamide Gel Electrophoresis*

PAS: *Principal Axis System*

PCS: *PseudoContact Shift*

PDB: *Protein Databank*

PDSD: *Proton Driven Spin Diffusion*

PMLG: *Phase Modulated Lee-Goldberg*

ppm: *parts per million*

PRE: *Paramagnetic Relaxation Enhancement*

QCPMG: *Quadrupolar Carr-Purcell-Meiboom-Gill*

RDS: *Residual Dipolar Splitting*

REAPDOR: *Rotational Echo Adiabatic Passage DOuble Resonance*

REDOR: *Rotational Echo DOuble Resonance*

RESPDOR: *Rotary Resonance Echo Saturation Pulse DOuble Resonance*

RF: *RadioFrequency*

RMSD: *Root Mean Squared Deviation*

SDS: *Sodium Dodecyl Sulphate*

SOQI: *Second Order Quadrupolar Interaction*

SPINAL: *Small Phase INcremental ALternation*

SQ: *Single Quantum*

ssNMR: *solid-state Nuclear Magnetic Resonance*

STMAS: *Satellite Transition Magic Angle Spinning*

TPPI: *Time Proportional Phase Incremented*

TPPM: *Two Pulse Phase Modulated*

TRAPDOR: *TRAnsfer of Populations in DOuble Resonance*

WURST: *Wideband Uniform Rate Smooth Truncation*

CHAPTER 1

MOTIVATION AND BACKGROUND

1.1 Introduction

Since its initial application to solids and liquids over 60 years ago, nuclear magnetic resonance (NMR) has become an indispensible tool for the analysis of structure, dynamics, and molecular interactions of a variety of inorganic, organic and biological samples. In the solution-state, the fast molecular tumbling of a molecule allows for the acquisition of high-resolution spectra, dominated by isotropic chemical shifts. This has led to the adoption of solution-state NMR as a fundamental spectroscopic technique in many branches of the chemical sciences. In structural biology, alongside X-ray crystallography, solution-state NMR has become a cornerstone technique for elucidating structures and dynamics in large biomolecules.

In solids, reduced molecular motion compared to liquids complicates acquisition of NMR spectra. This is because many of the interactions that a nucleus experiences in a magnetic field depend on the orientation of the nucleus with respect to the external magnetic field, or other nuclei in the sample. In this case, the interactions are said to anisotropic. The lack of motion of molecules in the solid-state prevents the averaging of these anisotropic interactions, leading to far broader signals than are observed in solution, and spectra with reduced resolution. The features of these broad lines, however, are rich in structural information. The difficulty associated with extracting information from the broad lineshapes associated with solid-state NMR (ssNMR) means that ssNMR has thus far failed to attain the ubiquity of its solution-state counterpart. However, a great number of theoretical, methodological and technological advancements over the last 70 years has seen the development of a plethora of techniques to increase resolution, and extract structural information from ssNMR signals to the point that it is now a well established technique for structural and dynamic studies of inorganic, organic and biological materials. Indeed, ssNMR can now be used to determine structures of large biomolecules in the solid-state that are

not amenable to solution-state NMR or X-ray crystallography, such as amyloid fibrils[6-8] and membrane proteins[9-11].

Most frequently, NMR is performed on spin-½ nuclei which are often easily studied by NMR since the size of their internal magnetic interactions is small compared to the external applied magnetic field. However, a large number of elements have a spin >½, meaning they interact with electric field gradients at the nucleus, as well as local magnetic fields. This interaction is known as the quadrupolar interaction. The quadrupolar interaction is an anisotropic interaction that is frequently large; often larger than the amplitude of RF pulses that can be applied with typical amplifiers and probes, and in some cases large enough to rival or even exceed the magnitude of the external magnetic field used for NMR experiments. Investigation of these nuclei with NMR can, therefore, represent a significant technical challenge, though the quadrupolar interaction is extremely sensitive to the local electrostatic environment, governed by the structure and dynamics, and characterisation of the quadrupolar coupling can provide useful structural and dynamic information.

One such quadrupolar nucleus is ^{14}N. Despite the extremely high natural abundance (99.63%) of ^{14}N, and the prevalence of nitrogen in natural materials, ^{14}N has failed to become a routinely investigated nucleus by ssNMR. This is primarily due to it being a spin-1 nucleus with a frequently relatively large (1MHz – ~7MHz) quadrupolar interaction. Solid-state NMR studies of nitrogen are more frequently done on the ^{15}N isotope, a spin-½ nucleus, and potential information on structure and dynamics available from the quadrupolar coupling of ^{14}N is lost. Due to the low natural abundance of ^{15}N (0.37%), NMR studies on ^{15}N frequently require samples that have been isotopically enriched. This can be difficult, expensive, and in some instances impractical or impossible, which provides another motive for investigating the naturally abundant ^{14}N nucleus.

This thesis is concerned with developing methods for investigating ^{14}N by ssNMR, and characterising ^{14}N quadrupolar couplings in organic and biological samples, with a particular focus on applying the techniques to full length recombinantly expressed proteins, where ^{14}N quadrupolar couplings have not previously been measured.

In this Chapter, the nuclear spin interactions relevant for interpreting ssNMR spectra of ^{14}N are introduced and explained. A review of existing techniques to study ^{14}N by ssNMR is presented, with a particular focus on application or potential applicability to complex biological samples. Finally, model protein systems useful for method development of ^{14}N NMR methods are introduced and considered.

1.2 Solid-State NMR Spectroscopy

In this section, a brief theoretical description of some of the foundations of ssNMR is provided. The specific focus is on providing terms for the interactions that give rise to the spectral features observed in the ssNMR experiments described later in this thesis, in particular those of ^{14}N. It is therefore not exhaustive, and interested readers are referred to a number of other in depth texts on magnetic resonance for a more detailed, systematic description of the theory[12-15].

1.2.1 The Nuclear Spin Hamiltonian

The nuclear spin Hamiltonian operator contains all the information required to express the interactions of nuclei with spin≠0 in a magnetic field. It can be expressed as the sum of internal and external interactions:

$$\hat{H} = \hat{H}_{Ext} + \hat{H}_{Int} \tag{1.1}$$

where $\hat{H}_{Ext}$ describes the response of spins to the static magnetic field, B_0, and radiofrequency fields, B_1, and $\hat{H}_{Int}$ comprises the sum of the internal interactions of the spins:

$$\hat{H}_{Int} = \hat{H}_{CS} + \hat{H}_{D} + \hat{H}_{J} + \hat{H}_{Q} \tag{1.2}$$

where CS, D, J and Q respectively denote the chemical shift interaction, through space dipolar coupling, the electron mediated *J*-coupling, and the quadrupolar coupling. The quadrupolar coupling only exists for spins with $I > ½$. In the following, we consider the effects of each of the terms in Equations (1.1) and (1.2) in turn.

1.2.2 External Interactions

Initially the interactions of spins with external magnetic fields are considered.

1.2.2.1 The Zeeman Interaction

The Zeeman interaction is that of nuclei which have non-zero spin with a static, external magnetic field. When not in a magnetic field, the energy levels of spin states are degenerate. However, the application of a magnetic field splits the energy levels into 2*I*+1 levels, where *I* is the nuclear spin quantum number. The Zeeman Hamiltonian can be represented, in Cartesian notation, as:

$$\hat{H}_Z = \vec{\hat{I}} \cdot (-\gamma) \vec{\mathbf{1}} \cdot \vec{B}_0 \tag{1.3}$$

where $\vec{\hat{I}}$ is the nuclear spin operator, represented as a vector, $\vec{\hat{I}} = (\hat{I}_x, \hat{I}_y, \hat{I}_z)$. The gyromagnetic ratio of the spin, γ, describes the strength of the interaction of the spin with the magnetic field, $\vec{\mathbf{1}}$ is the 3x3 unit matrix, and $\vec{B}_0$ is the magnetic field, a vector quantity, which is customarily applied along the *z*-axis of the laboratory frame, $\vec{B}_0 = (0,0,B_{0z})$. One can show that Equation(1.3) is equal to:

$$\begin{aligned} \hat{H}_z &= -\gamma B_0 \hat{I}_z \\ &= \omega_0 \hat{I}_z \end{aligned} \tag{1.4}$$

where $\omega_0 = -\gamma B_0$ is the Larmor frequency in rad s^{-1}. This defines the spacing between energy levels in a magnetic field. This spacing is illustrated for an *I*=1 nucleus in Figure 1.

1.2.2.2 The Radiofrequency Interaction

In addition to the static magnetic field, pulses of radiofrequency (RF) irradiation are applied to spins during NMR experiments, along a direction perpendicular to

the B_0 field (aligned along the *z*-axis). This additional oscillating magnetic field, $\vec{B}_1$, results in mixing of Zeeman states if it is applied at a frequency, ω_{rf}, close to the Larmor frequency, ω_0, giving rise to detectable coherence states. The $\vec{B}_1$ field perpendicular to the $\vec{B}_0$ field can be described by:

$$\begin{aligned}\vec{B}_1(t) &= 2B_1\cos(\omega_{rf}t+\phi) \\ &= B_1(e^{i\omega_{rf}t+\phi}+e^{-i\omega_{rf}t+\phi})\end{aligned} \tag{1.5}$$

where $\pm\omega_{rf}$ are the angular frequencies of counter-rotating components, B_1, of the $\vec{B}_1$ RF field and ϕ is the phase of the RF pulse. Choosing the RF field to be close to the nuclear Larmor frequency, $\omega_{rf} \cong \omega_0$, allows the component with $-\omega_{rf}$ to be practically neglected, as only the $+\omega_{rf}$ component is close enough in frequency to the Larmor frequency to appreciably affect the magnetisation. The Hamiltonian for a pulse of RF irradiation at close to the Larmor frequency can therefore be expressed, in the same manner as the Zeeman interaction, as:

$$\begin{aligned}\hat{H}_{RF} &= \vec{\hat{I}}\cdot(-\gamma)\vec{\vec{1}}\cdot\vec{B}_1 \\ &= -\gamma B_1[\hat{I}_x\cos(\omega_{rf}t+\phi)+\hat{I}_y\sin(\omega_{rf}t+\phi)]\end{aligned} \tag{1.6}$$

In a frame of reference rotating at ω_{rf}, this can be simplified to:

$$\begin{aligned}\hat{H}_{rf}^{rot} &= -\gamma B_1[\vec{I}_x\cos(\phi)+\vec{I}_y\sin(\phi)] \\ &= \omega_1[\vec{I}_x\cos(\phi)+\vec{I}_y\sin(\phi)] \\ &= \omega_1\vec{I}_x \quad \text{if} \quad \phi = 0^\circ \\ &= \omega_1\vec{I}_y \quad \text{if} \quad \phi = 90^\circ\end{aligned} \tag{1.7}$$

where $\omega_1 = -\gamma B_1$, is the angular frequency of the RF pulse, known as the nutation frequency. The effect of the RF Hamiltonian is to rotate a magnetisation vector about an axis along which the pulse is applied. This can be exploited to

give rise to magnetisation in the transverse plane that can be detected by the receiver coil in an NMR spectrometer.

1.2.3 Internal Interactions

Here, we define the separate terms of Equation (1.2) that comprise the interactions of spins with internal fields. In the following, each Hamiltonian is expressed in the spherical tensor notation. This notation is the most convenient for expressing the effects of rotations on Hamiltonians, and interpreting the orientation dependence of interactions. In this notation, Hamiltonians describing internal spin interactions may be generally expressed by:

$$\hat{H}_\lambda = C_\lambda \sum_{l=0}^{2} \sum_{q=-l}^{l} (-1)^q A_{l,q} \hat{T}_{l,-q} \tag{1.8}$$

where l is the rank and q the order of the spherical tensors, $A_{l,q}$, and spherical tensor operators, $\hat{T}_{l,-q}$. Separate factors, $\hat{T}_{l,-q}$ and $A_{l,q}$, express spin operators, and their orientation with respect to the laboratory frame, respectively. In the Hamiltonians discussed here, spherical tensors up to rank $l = 2$ are required, and q may take values of $q = -l, -l+1, ..., l-1, l$. C_λ is a constant, invariant under rotations, that is specific for each of the interactions CS, D, J and Q.

We may simplify the general Hamiltonian of Equation (1.8) for Hamiltonians in the high field NMR case by writing the Hamiltonians in a rotating frame of reference, and making the secular approximation. Since the magnetic fields, B_0, typically used for high field NMR (~4.7T ≥ B_0 ≥ 23.5T) generally render the Zeeman interaction at least three orders of magnitude larger than the internal interactions, internal interactions may often be accurately described as perturbations to the Zeeman energy levels in terms of first-order perturbation theory. Only terms that commute with the Zeeman Hamiltonian contribute within first-order perturbation theory. In light of this, we may split internal Hamiltonians

into two terms:

$$\hat{H}_{Int} = \hat{H}'_{Int} + \hat{H}''_{Int} \tag{1.9}$$

where one part commutes with the Zeeman Hamiltonian, and one part does not:

$$\begin{aligned} [\hat{H}_Z, \hat{H}'_{Int}] &= 0 \\ [\hat{H}_Z, \hat{H}''_{Int}] &\neq 0 \end{aligned} \tag{1.10}$$

For a term to commute with the Zeeman Hamiltonian, it must be diagonal in the Zeeman eigenbasis, which is proportional to the $\hat{I}_z$ operator. Such terms will share this basis as their common eigenfunctions, and commute. Those diagonal elements in the commuting internal Hamiltonian, $\hat{H}'_{Int}$, are known as 'secular' terms which perturb the Zeeman energy levels to first order. The non-commuting parts of the Internal Hamiltonian, $\hat{H}''_{Int}$, are not considered to first-order. Only spherical tensor operators of order $q = 0$ have on-diagonal elements in the Zeeman eigenbasis, and commute with the Zeeman Hamiltonian, satisfying the relation:

$$[\hat{H}_Z, \hat{T}_{l,q}] = q\hat{T}_{l,q} \tag{1.11}$$

Hence, the Hamiltonian of Equation (1.8) may be simplified to:

$$\begin{aligned} H_\lambda &= C_\lambda \sum_{l=0}^{2} A_{l,0}\hat{T}_{l,0} \\ &= C_\lambda (A_{00}\hat{T}_{00} + A_{10}\hat{T}_{10} + A_{20}\hat{T}_{20}) \end{aligned} \tag{1.12}$$

in the high field limit. If we define a general second rank Cartesian tensor:

$$\overleftrightarrow{A} = \begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix} \tag{1.13}$$

we may express this tensor in terms of the irreducible components of a second rank spherical tensor:

$$\begin{aligned} A_{00} &= \frac{-1}{\sqrt{3}}(A_{xx} + A_{yy} + A_{zz}) \\ A_{10} &= -\frac{i}{2}(A_{xy} - A_{yx}) \\ A_{1,\pm1} &= -\frac{1}{2}(A_{zx} - A_{xz} \pm i(A_{zy} - A_{yz})) \\ A_{20} &= -\frac{1}{\sqrt{6}}(3A_{zz} - (A_{xx} + A_{yy} + A_{zz})) \\ A_{2,\pm1} &= \pm\frac{1}{2}(A_{xz} - A_{zx} \pm i(A_{yz} - A_{zy})) \\ A_{2,\pm2} &= \frac{1}{2}(A_{xx} - A_{yy} \pm i(A_{xy} - A_{yx})) \end{aligned} \tag{1.14}$$

Typically it will be more convenient to define a tensor in the principal axis system (PAS) of the interaction. This is a frame of reference in which the symmetric part of the Cartesian tensor of Equation (1.13) is diagonal. The orientation of this frame depends on the interaction and molecule in question. For a dipolar coupling, the *z*-axis of the PAS will be aligned along the vector that connects the two nuclei. For the chemical shift, the *z*-axis of the PAS will be defined by the local electronic environment, with the *z*-axis along the direction of the least shielding. In the PAS, the irreducible spherical tensor components of

Equation (1.14) may be expressed as:

$$\begin{aligned}
\rho_{00} &= \frac{-1}{\sqrt{3}} A_{iso} \\
\rho_{10} &= -i\sqrt{2} A_{xy} \\
\rho_{1,\pm 1} &= -A_{xz} \pm A_{yz} \\
\rho_{20} &= \sqrt{\frac{3}{2}} \delta \\
\rho_{2,\pm 1} &= 0 \\
\rho_{2,\pm 2} &= -\frac{1}{2} \delta \eta
\end{aligned} \tag{1.15}$$

where we have defined the Cartesian tensor, $\vec{\vec{A}}$, in terms of its isotropic value:

$$A_{iso} = Tr\left\{A\right\} = A_{xx} + A_{yy} + A_{zz} \tag{1.16}$$

its anisotropy:

$$\delta = A_{zz} - \frac{A_{iso}}{3} \tag{1.17}$$

and its asymmetry:

$$\eta = \frac{A_{yy} - A_{xx}}{\delta} \tag{1.18}$$

with the principal components of the tensor ordered according to the Haeberlen convention[13]:

$$\left| A_{zz} - \frac{A_{iso}}{3} \right| \geq \left| A_{xx} - \frac{A_{iso}}{3} \right| \geq \left| A_{yy} - \frac{A_{iso}}{3} \right| \tag{1.19}$$

The irreducible components of the spherical tensor in the PAS (Equation (1.15)) require rotating into the laboratory frame of reference, in order to write the

Hamiltonian of Equation (1.12). This transformation is achieved using the Wigner rotation matrix:

$$A_{lm} = \sum_{m'=-l}^{l} \rho_{lm'} D^{l}_{m'm}(\Omega_{PL}) \tag{1.20}$$

where the angles $\Omega_{PL} = \{\alpha_{PL}, \beta_{PL}, \gamma_{PL}\}$ are Euler angles that bring the PAS into coincidence with the laboratory frame via three successive rotations[16]:

1. A rotation about the *z*-axis of $\alpha_{PL} (0 \le \alpha_{PL} \le 2\pi)$
2. A rotation about the new *y'*-axis of $\beta_{PL} (0 \le \beta_{PL} \le \pi)$
3. A rotation about the new *z'*-axis of $\gamma_{PL} (0 \le \gamma_{PL} \le 2\pi)$

The Wigner rotation matrix is given as:

$$D^{l}_{m'm}(\Omega_{PL}) = e^{-i\alpha_{PL}m'} d^{l}_{m'm}(\beta_{PL}) e^{-i\gamma_{PL}m} \tag{1.21}$$

where $d^{l}_{m'm}$ are matrix elements of the reduced Wigner rotation matrix. They may be found tabulated in a number of sources, e.g. Haberlen, 1976[13]; Mehring, 1983[12]. Only those elements which are relevant for rotating secular, observable terms into the laboratory frame are given here:

$$\begin{aligned} d^{0}_{00} &= 1 \\ d^{2}_{00} &= \frac{1}{2}(3\cos^2\beta_{PL} - 1) \\ d^{2}_{\pm 2,0} &= \sqrt{\frac{3}{8}}\sin^2\beta_{PL} \end{aligned} \tag{1.22}$$

The spherical tensor operators, $\hat{T}_{l,0}$, of Equation (1.12) depend on the characteristics of each interaction, and differ depending on whether they

constitute heteronuclear spin-spin, homonuclear spin-spin, spin-field or quadratic interactions. Hence they are stated explicitly for each interaction covered here.

1.2.3.1 Chemical Shift Interaction

The chemical shift is the interaction of nuclei with local magnetic fields that arise from local electrons, which slightly shift the Larmor frequency of the nuclei. For the chemical shift:

$$C_\lambda = -\gamma \tag{1.23}$$

The spin-field spherical tensor operator components, $\hat{T}_{l,m}$, that are non-zero, and that give rise to secular terms in the Hamiltonian ($m = 0$) are:

$$\begin{aligned} \hat{T}_{00} &= \frac{-1}{\sqrt{3}} \hat{T}_{10}^{(I)} \hat{T}_{10}^{(B)} = \frac{-1}{\sqrt{3}} \hat{I}_z B_{0z} \\ \hat{T}_{20} &= \frac{1}{\sqrt{6}} 2\hat{T}_{10}^{(I)} \hat{T}_{10}^{(B)} = \frac{1}{\sqrt{6}} 2\hat{I}_z B_{0z} \end{aligned} \tag{1.24}$$

where the superscripts I and B denote spin and field tensor operators, respectively. The non-zero irreducible spherical tensor components of the spatial tensor in the PAS, excluding first-rank ($l = 1$) terms which do not contribute to the spectrum, are, according to Equation (1.15):

$$\begin{aligned} \rho_{00} &= \frac{-1}{\sqrt{3}} A_{iso} \\ \rho_{20} &= \sqrt{\frac{3}{2}} \delta \\ \rho_{2,\pm 2} &= -\frac{1}{2} \delta\eta \end{aligned} \tag{1.24}$$

Rotating these terms from the PAS to the laboratory frame, using Equation (1.21), and gathering the terms together, we obtain the secular chemical shift Hamiltonian in the laboratory frame, according to Equation (1.12):

$$\begin{aligned}\hat{H}_{CS} &= C_{CS}\sum_{l=0}^{2} A_{l,0}\hat{T}_{l,0} \\ &= C_{CS}(A_{00}\hat{T}_{00} + A_{10}\hat{T}_{10} + A_{20}\hat{T}_{20}) \\ &= -\gamma(A_{00}\hat{T}_{00} + A_{20}\hat{T}_{20}) \\ &= \omega_0 \hat{I}_z\left(\bar{\sigma} + \frac{1}{2}\delta(3\cos^2\beta_{PL} - 1) - \eta\sin^2\beta_{PL}\cos(2\alpha_{PL})\right)\end{aligned} \tag{1.24}$$

where $\bar{\sigma}$ is the isotropic chemical shift, $\bar{\sigma} = Tr\{\hat{\sigma}\}/3$. First rank ($l = 1$) terms do not contribute to the Hamiltonian to first order, since the rank 1 spin-field spherical tensor operator component, $\hat{T}_{1,0}$, is zero. From this Hamiltonian, the transition frequency for the chemical shift (i.e. the frequency by which the chemical shift perturbs the Zeeman energy levels) may be given as[14]:

$$\omega_{CS}(\alpha_{PL},\beta_{PL}) = \omega_0\left(\bar{\sigma} + \frac{1}{2}\delta(3\cos^2\beta_{PL} - 1) - \eta\sin^2\beta_{PL}\cos(2\alpha_{PL})\right) \tag{1.25}$$

1.2.3.2 Dipolar Coupling Interaction

The dipolar coupling describes the through space coupling between two proximal nuclei. Since each nucleus has a dipole moment, they each give rise to oscillating magnetic fields that affect the transition frequency of other nearby spins. There is a homonuclear and a heteronuclear case to consider. The constant in the case of the dipolar coupling Hamiltonian, derived from the classical treatment of two interacting dipoles, is given as:

$$C_\lambda = b_{IS} = -2\frac{\mu_0\gamma_I\gamma_S\hbar}{4\pi r_{IS}^3} \tag{1.26}$$

where μ_0 is the permittivity of free space, γ_I and γ_S are the gyromagnetic ratios of coupled spins, *I* and *S*, respectively, $\hbar$ is Planck's constant, and r_{IS} is the vector that joins spins I and S. The dipolar coupling tensor is axially symmetric and traceless, meaning that the isotropic component and the asymmetry are both 0. Therefore, there is a single non-zero spherical tensor component in the PAS:

$$\rho_{20} = \sqrt{\frac{3}{2}}\delta \tag{1.27}$$

Giving rise to a single secular ($m = 0$) component in the laboratory frame after rotation:

$$A_{20} = \sqrt{\frac{3}{8}}\delta(3\cos^2\beta_{PL} - 1) \tag{1.28}$$

In the homonuclear case, the secular spin-spin spherical tensor operator is given as:

$$\hat{T}_{20} = \frac{1}{\sqrt{6}}[3\hat{I}_z^{(I)}\hat{I}_z^{(S)} - (\bar{\hat{I}}^{(I)} \cdot \bar{\hat{I}}^{(S)})] \tag{1.29}$$

and in the heteronuclear case, we have:

$$\hat{T}_{20} = \frac{1}{\sqrt{6}}2\hat{T}_{10}^{(I)}\hat{T}_{10}^{(S)} = \frac{1}{\sqrt{6}}2\hat{I}_z^{(I)}\hat{I}_z^{(S)} \tag{1.30}$$

The secular homonuclear dipolar coupling Hamiltonian is given by:

$$\begin{aligned} H_D^{II} &= C_D \sum_{l=0}^{2} A_{l,0} \hat{T}_{l,0} \\ &= C_D (A_{00}\hat{T}_{00} + A_{10}\hat{T}_{10} + A_{20}\hat{T}_{20}) \\ &= b_{IS}(A_{20}\hat{T}_{20}) \\ &= -2\frac{\mu_0 \gamma_I \gamma_I \hbar}{4\pi r_{IS}^3}(3\cos^2\beta_{PL} - 1)(3\hat{I}_z^{(I)}\hat{I}_z^{(S)} - (\vec{\hat{I}}^{(I)} \cdot \vec{\hat{I}}^{(S)})) \end{aligned} \tag{1.31}$$

and in the heteronuclear case, we have:

$$H_D^{IS} = -2\frac{\mu_0 \gamma_I \gamma_I \hbar}{4\pi r_{IS}^3}(3\cos^2\beta_{PL} - 1)(\hat{I}_z^{(I)} \cdot \hat{I}_z^{(S)}) \tag{1.32}$$

The difference in the homonuclear and heteronuclear Hamiltonians is due to the different forms of the spin-spin spherical tensor operators (Equations (1.29) and (1.30)) that are the tensor product of identical and different spherical tensors, respectively.

1.2.3.3 J-Coupling Interaction

The *J*-coupling is a through-bond (electron mediated) coupling between proximal spins. It is an axially symmetric second rank tensor with an isotropic and a second rank component. However, the second rank contribution to the *J*-coupling is impossible to separate from the second rank dipolar coupling experimentally, since they transform identically under rotations, so usually only the isotropic component is considered, and the second rank contribution is considered a small perturbation of the dipolar coupling. For the *J*-coupling:

$$C_\lambda = 2\pi \tag{1.33}$$

The spherical tensor component of the *J*-tensor in the PAS is invariant

under rotations, and therefore identical in the laboratory frame:

$$A_{00} = \rho_{00} = \frac{-1}{\sqrt{3}} J_{iso} = -\sqrt{3} \times (2\pi J_{IS}) \tag{1.34}$$

The secular spin-spin spherical tensor operator is:

$$\hat{T}_{00} = \frac{-1}{\sqrt{3}} (\vec{\hat{I}}^{(I)} \cdot \vec{\hat{I}}^{(S)}) \tag{1.35}$$

for a homonuclear spin pair, and:

$$\hat{T}_{00} = \frac{-1}{\sqrt{3}} (\hat{I}_z^{(I)} \cdot \hat{I}_z^{(S)}) \tag{1.36}$$

for heteronuclear spins. Hamiltonians for the heteronuclear and homonuclear cases are given by:

$$\begin{aligned} \hat{H}_J &= C_J \sum_{l=0}^{2} A_{l,0} \hat{T}_{l,0} \\ &= C_J (A_{00} \hat{T}_{00}) \\ &= 2\pi (A_{00} \hat{T}_{00}) \\ \hat{H}_J^{II} &= 2\pi J_{II} (\vec{\hat{I}}^{(I)} \cdot \vec{\hat{I}}^{(S)}) \\ \hat{H}_J^{IS} &= 2\pi J_{IS} (\hat{I}_z^{(I)} \cdot \hat{I}_z^{(S)}) \end{aligned} \tag{1.37}$$

1.2.3.4 Quadrupolar Coupling Interaction

The quadrupolar interaction applies only to nuclear spins with $I > ½$. In this case, the nuclei possess a non-spherical charge distribution known as an electric quadrupole moment, which enables them to interact with electric field gradients (EFG) at the site of the nucleus. The EFG tensor is represented by a

traceless second rank tensor, $\ddot{V}$ meaning:

$$\begin{aligned} V_{iso} &= 0 \\ \delta &= V_{zz} \\ \eta &= \frac{V_{yy} - V_{xx}}{V_{zz}} \end{aligned} \tag{1.38}$$

The C_λ of Equation (1.12) is given as:

$$C_\lambda = \frac{eQ}{2I(2I-1)\hbar} \tag{1.39}$$

where e is the elementary charge, and eQ is the nuclear quadrupole moment. Frequently used to define the magnitude of the quadrupole coupling tensor is the quadrupole coupling constant in Hz:

$$C_Q = \frac{e^2 qQ}{\hbar} \tag{1.40}$$

along with the anisotropy, η. In the PAS of the EFG tensor, we need only to consider the second rank spatial spherical tensor components to first-order, since the zeroth rank component is zero as $V_{iso} = 0$, and the rank 1 terms are non-secular to first order. The components which are non-zero are:

$$\begin{aligned} \rho_{20} &= \sqrt{\frac{3}{2}}\delta \\ \rho_{2,\pm 2} &= -\frac{1}{2}\delta\eta \end{aligned} \tag{1.41}$$

Rotating these into the laboratory frame, we obtain:

$$A_{20} = \sqrt{\frac{3}{8}}\delta(3\cos^2\beta_{PL} - 1) - \eta\sin^2\beta_{PL}\cos(2\alpha_{PL}) \tag{1.42}$$

In this case, the secular, second-rank spin tensor operators are given by:

$$\hat{T}_{20} = \frac{1}{\sqrt{6}}(3(\hat{I}_z)^2 - \bar{\hat{I}}^{(I)}(\bar{\hat{I}}^{(I)}+1)) \tag{1.43}$$

Gathering these terms, we obtain for the first order quadrupolar Hamiltonian in the rotating laboratory frame:

$$\begin{aligned}\hat{H}_Q^{(1)} &= C_\lambda \sum_{l=0}^{2} A_{l,0}\hat{T}_{l,0} \\ &= C_\lambda (A_{00}\hat{T}_{00} + A_{10}\hat{T}_{10} + A_{20}\hat{T}_{20}) \\ &= C_\lambda (A_{20}\hat{T}_{20}) \\ &= \frac{e^2qQ}{4I(2I-1)\hbar}\left[\begin{array}{l}\left(\frac{1}{2}(3\cos^2\beta_{PL}-1) - \eta\sin^2\beta_{PL}\cos(2\alpha_{PL})\right) \\ (3(\hat{I}_z)^2 - \bar{\hat{I}}^{(I)}(\bar{\hat{I}}^{(I)}+1))\end{array}\right]\end{aligned} \tag{1.44}$$

For nuclei with large quadrupolar coupling constants (100s kHz - MHz), the secular approximation is often not a valid one, and higher order corrections to the Zeeman Hamiltonian need to be considered in order to accurately represent the rotating frame Hamiltonian. In this case, one approach is to introduce a second order quadrupolar Hamiltonian, which includes all second-rank terms, in the laboratory frame. This can be given as:

$$\hat{H}_Q^{(2)} = C_\lambda \frac{1}{\omega_0}\left(A_{21}A_{2,-1}[\hat{T}_{21}\hat{T}_{2,-1}] + \frac{1}{2}A_{22}A_{2,-2}[\hat{T}_{22}\hat{T}_{2,-2}]\right) \tag{1.45}$$

Equation (1.45) contains terms that include the product of two second rank spherical tensors. There will, therefore, be fourth rank spatial spherical tensors included in the second order quadrupolar Hamiltonian which transform with orientation dependencies which cannot be removed by MAS. From the first and second order quadrupolar Hamiltonians, the first and second order

perturbations to the Zeeman energy levels may be respectively given as[14]:

$$\omega_Q^{(1)}(\alpha_{PL},\beta_{PL}) = \frac{eQ}{4I(2I-1)\hbar}\left[\begin{array}{l}\frac{1}{2}(3m^2 - I(I+1))(3\cos^2\beta_{PL} - 1) \\ -\eta\sin^2\beta_{PL}\cos(2\alpha_{PL})\end{array}\right] \qquad (1.46)$$

and:

$$\omega_Q^{(2)}(\alpha_{PL},\beta_{PL}) = \left(\frac{eQ}{4I(2I-1)\hbar}\right)^2 \frac{2}{\omega_0} m\left[\begin{array}{l}[I(I+1) - 3m^2]D_{00}^0 \\ +[8I(I+1) - 12m^2 - 3]D_{20}^2 \\ +[18I(I+1) - 34m^2 - 5]D_{40}^4\end{array}\right] \qquad (1.47)$$

where m is the magnetic quantum number. The second order quadrupolar interaction (SOQI) is usually far smaller that the first order, being on the order of 10s kHz for a C_Q of MHz. It is apparent that the magnitude of the SOQI depends inversely on the strength of the external magnetic field, since it scaled by the Larmor frequency. Thus the magnitude of this term decreases with increasing B_0 field strength. Another important feature of the SOQI is that it has terms defined by zeroth, second and fourth rank spherical tensors. The fourth rank spherical tensors are rotated according to fourth rank Wigner rotation matrices, and these terms are not removed under MAS. In practice, this term gives rise to a second order broadening in the spectra of quadrupolar nuclei under static and MAS conditions. The term which is defined by a zeroth rank tensor is invariant under rotations and, therefore, isotropic. This means that isotropic shift in the NMR spectra of quadrupolar nuclei will have a contribution from the quadrupole coupling. This becomes very important in the NMR spectra of nuclei with large quadrupolar couplings, and particularly so when using techniques that remove the contribution of the first order quadrupolar interaction (FOQI). This isotropic term is known as the second order isotropic quadrupolar

shift, δ_Q^{iso}, and is given in Hz as[17]:

$$\delta_Q^{iso} = \left\{ -\frac{3}{40}\left(\frac{\chi_Q}{\nu_0}\right)^2 [I(I+1) - 9m(m-1) - 3] / [I^2(2I-1)^2] \right\} \tag{1.48}$$

where we have expressed the Larmor frequency in Hz as $\nu_0 = \omega_0 / 2\pi$. For ^{14}N, where $I = 1$ and m takes the values $m = 0, \pm 1$, Equation (1.48) simplifies to:

$$\delta_Q^{iso} = -\frac{3}{40}\left(\frac{\chi_Q}{\nu_0}\right)^2 \tag{1.49}$$

where χ_Q is the quadrupolar product:

$$\chi_Q = C_Q \sqrt{1 + \frac{\eta^2}{3}} \tag{1.50}$$

In this thesis, the only quadrupolar nucleus we consider is the spin-1 ^{14}N, where m may take the values $m = 0, \pm 1$. An energy level diagram for the spin-1 ^{14}N in a magnetic field is shown in Figure 1, considering the effect of the Zeeman interaction, the first and second order quadrupolar interactions, and the chemical shift. There are two allowed transitions for a spin-1 nucleus, $m = 0 \leftrightarrow m = -1$, and $m = 0 \leftrightarrow m = +1$. The Zeeman interaction splits the energy levels of these transitions equally. The first order quadrupolar interaction makes these two transitions unequal, but preserves the net energy difference between $m = +1$ and $m = -1$ states. The second order quadrupolar interaction changes the relative magnitude of the $m = 0 \leftrightarrow m = -1$, and $m = 0 \leftrightarrow m = +1$ transitions, as does the chemical shift interaction.

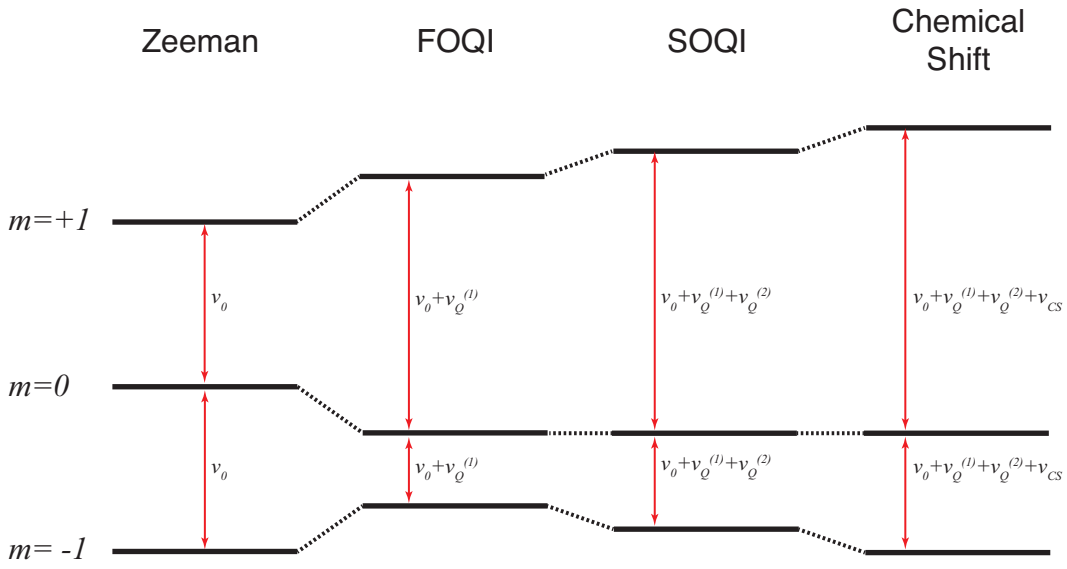

Figure 1. Diagrammatic Representation of the Energy Levels of a Spin-1 Nucleus in a Strong Magnetic Field

The Zeeman energy levels for a spin-1 nucleus and pertubations thereof for a the first and second order quadrupolar and chemical shift interactions are shown. The relative sizes of the pertubations are exaggerated to emphasise the effect of each interaction on the energy levels.

1.3 Nitrogen in Solid-State NMR

There are two NMR active isotopes of nitrogen: ^{14}N and ^{15}N. ^{14}N would potentially seem a likely candidate for routine NMR studies of organic or biological samples; not only is it extremely common in naturally occurring materials, but it is 99.63% abundant and has a reasonable gyromagnetic ratio ($\gamma_{^{14}N} / \gamma_{^{1}H} \approx 0.07$). However, it is a spin-1 nucleus, which has a quadrupole moment of $Q = 0.02044 \times 10^{-28}\ m^2$, which leads to quadrupole coupling constants, C_Q, typically on the order of 1-4MHz, with values up to ~7MHz reported[5]. Studies of ^{14}N nuclei with quadrupolar interactions of this magnitude are not straightforward. As an integer spin nucleus, there is no narrow central transition that is not affected by the first order quadrupolar interaction, and ^{14}N is not amenable to the high-resolution multi-pulse methods MQMAS[18] and STMAS[19]. The breadth of the powder pattern in a static powdered sample with even a relatively modest (for ^{14}N) C_Q of 1MHz is greater than the bandwidth of excitation for RF pulses that can be applied with standard hardware (~200kHz) meaning an accurate lineshape from a powdered sample is difficult to achieve. The first order quadrupolar interaction has an orientation dependency of $3\cos^2\beta_{PL} - 1$ and so can, in principle, be removed by MAS when the spinning rate exceeds the size of the interaction, to yield an isotopic peak. However, achievable spinning rates (currently ~110kHz maximum) do not rival the size of the first order quadrupolar interaction in most cases for ^{14}N, and a manifold of sidebands, themselves broadened by second order terms, is the best that usually can be achieved. Even then, the excitation bandwidth issue remains. In addition, for removing contributions to the ssNMR spectra from interactions as large as the FOQI, the magic angle requires setting very accurately, and spinning needs to be extremely stable to avoid introducing broadening or splitting from incomplete averaging of FOQI interaction under MAS. These issues associated with MAS, and combined with the bandwidth limitations of RF pulses mean that the shape and broadening of powder patterns in static samples, and sideband manifolds under MAS, can often be detrimentally affected by experimental limitations.

These difficulties notwithstanding, a number of experimental strategies to investigate ^{14}N by ssNMR have been presented in the literature. These methods are reviewed here, with a focus on the information which can be obtained in biological and organic samples, which this thesis investigates, and on the indirect detection method, which remaining Chapters of this thesis are primarily concerned with. For more exhaustive reviews of the ^{14}N NMR literature, the reader is referred to excellent recent reviews by O'Dell on direct ^{14}N detection[5] and ^{14}N overtone NMR[20], and by Cavadini on indirect detection of ^{14}N[21].

1.3.1 Direct Detection of ^{14}N

1.3.1.1 Single Crystals

^{14}N spectra can be relatively simply acquired from samples of arbitrary quadrupolar coupling magnitude when prepared as single crystals. In such a sample, each magnetically distinct ^{14}N site will have the same orientation of the EFG tensor, which will give rise to two resonance lines corresponding to the two fundamental transitions, $m = 0 \leftrightarrow m = -1$, and $m = 0 \leftrightarrow m = +1$. The frequency separation between the lines may be large, MHz for ^{14}N sites with low symmetry, requiring stepping of the transmitter frequency over large ranges to find the signals, but once the signals are located they may be simply observed with one-pulse or cross polarisation (CP) acquisition. Acquiring spectra at a series of different angles with respect to the magnetic field and measuring the quadrupolar splitting allows the determination of the C_Q and asymmetry of ^{14}N sites in single crystals. Acquiring with and without ^{1}H decoupling at sites where ^{14}N is adjacent to a proton splits the ^{14}N signal according to the ^{1}H/^{14}N dipolar coupling, allowing determination of the magnitude of the ^{1}H/^{14}N dipolar coupling, and hence the length of the ^{1}H/^{14}N bond. Both of these methods were demonstrated by Griffin and coworkers for a sample of N-acetyl-valine (NAV), where the asymmetric amide ^{14}N site has a C_Q=3.21MHz, η=0.32[22]. These methods have also been applied to a variety of other amino acids such as glycine, where the primary amine ^{14}N site has a smaller quadrupolar coupling of C_Q=1.18MHz, η=0.54[23] and modified amino acids such as L-histidine

hydrochloride and L-asparagine hydrochloride which contain a variety of ^{14}N sites[24, 25].

1.3.1.2 Powdered Samples

Polycrystalline, powdered samples where a distribution of EFG tensor orientations is observed for a given ^{14}N site present a more difficult situation for measurement of ^{14}N. However, a number of approaches have been developed for investigating ^{14}N in these situations for both static samples and under MAS. In static samples, for sites where the ^{14}N C_Q is small (<~200kHz), standard, rectangular pulses can generally be used to excite the entire ^{14}N spectrum and powder patterns can be obtained can be obtained from static samples. This direct detection method is the way that the deuterium nucleus (^2H), which is also a spin-1 nucleus, but with a far smaller C_Q due to a smaller quadrupole moment than ^{14}N (^2H Q = 2.86mbarn), is often studied in the solid-state. Frequently a quadrupolar echo sequence ($\pi/2_x - \tau - \pi/2_y - \tau$ – acquire) is used to acquire powder spectra. From such powder patterns, it is relatively simple to obtain information on the magnitude and asymmetry of the EFG tensor from the location of discontinuities in the ^{14}N powder pattern. Figure 2 demonstrates the features that are observed in ^{14}N powder patterns dominated by the FOQI, and how they may be related to the EFG tensor. However, this approach is generally of limited application, since C_Qs of >~200kHz are unusual for ^{14}N, and typically only observed in sites of high cubic symmetry and/or sites undergoing rapid rotation. Most samples with which this approach has been feasible are ammonium or choline salts with high cubic symmetry[5, 26].

Otherwise, specialised ‘wideline’ and ‘ultrawideline’ NMR techniques are required to excite broad static ^{14}N powder patterns. Schurko and coworkers have developed many techniques for exciting and detecting very broad powder spectra from spin-½ and quadrupolar nuclei, and applied them to static ^{14}N samples. The use of broadbanded adiabatic inversion pulses, known as WURST pulses, to increase the excitation bandwidth and record broad powder patterns was demonstrated for a series of amino acids, and discontinuities in the ^{14}N powder patterns could be used to accurately determine the ^{14}N C_Q and η in

glycine, leucine and proline[27, 28]. Shurko's group also demonstrated that WURST pulses could be introduced into QCPMG pulse trains to increase S/N in broad powder patterns with long enough T_2 relaxation[29], and that in this type of experiment, the transmitter frequency could be stepped over the width of broad powder patterns in separate experiments to acquire spectra in a 'piecewise' fashion[30]. Additionally, WURST pulses have been employed on the ^{14}N channel during ^{1}H/^{14}N CP to increase the bandwidth of CP to ^{14}N nuclei on static samples, which has been called BRAIN-CP, and can be combined with WURST-QCPMG methods to further increase S/N. Combining all these methods, piecewise acquired BRAIN-CP/WURST-QCPMG can be used to acquire static ssNMR spectra at ^{14}N sites with essentially arbitrary values of C_Q though experimental times increase linearly with the size of the C_Q[31, 32]. This method has been used to determine the C_Q and η of a series of polymorphs of glycine whose ^{15}N chemical shifts are very similar, differing by only a few ppm, but whose ^{14}N C_Qs differed by over 70kHz[32]. This demonstrates the high sensitivity of the ^{14}N EFG to local hydrogen bonding arrangements, which is generally far more so than the nitrogen chemical shift.

Under MAS, in a similar manner to static samples, if the ^{14}N C_Q is low (>~200kHz), a standard quadrupole echo pulse sequence can be relatively simply used to excite the entire frequency range of the ^{14}N signal. For spinning samples, a manifold of spinning sidebands is observed whose intensity and shape can be fit with simulations to determine the ^{14}N EFG parameters. As for static samples, this approach has not found widespread general application due to the limited range of samples that have ^{14}N sites that have high enough symmetry or motion to reduce the C_Q below ~200kHz. In order to excite and detect ^{14}N spectra under MAS with larger quadrupole couplings, specialised experimental conditions are required in terms of both MAS and RF hardware. Notably, Jakobsen *et al*. designed an extremely stable MAS setup and used a probe with a high RF bandwidth for single pulse excitation to record sideband manifolds of ^{14}N sites for a series of amino acids with quadrupolar couplings up to ~1.2MHz[33]. Fitting of simulations to these spectra allowed them to determine EFG and chemical shift parameters, even where sideband manifolds

from two sites overlapped. From the simulations, it was also possible to determine the small deviation from the magic angle which the experiments were performed at, -0.009, such is the sensitivity of interactions of this magnitude to mis-setting of the magic angle and, equivalently, stable MAS.

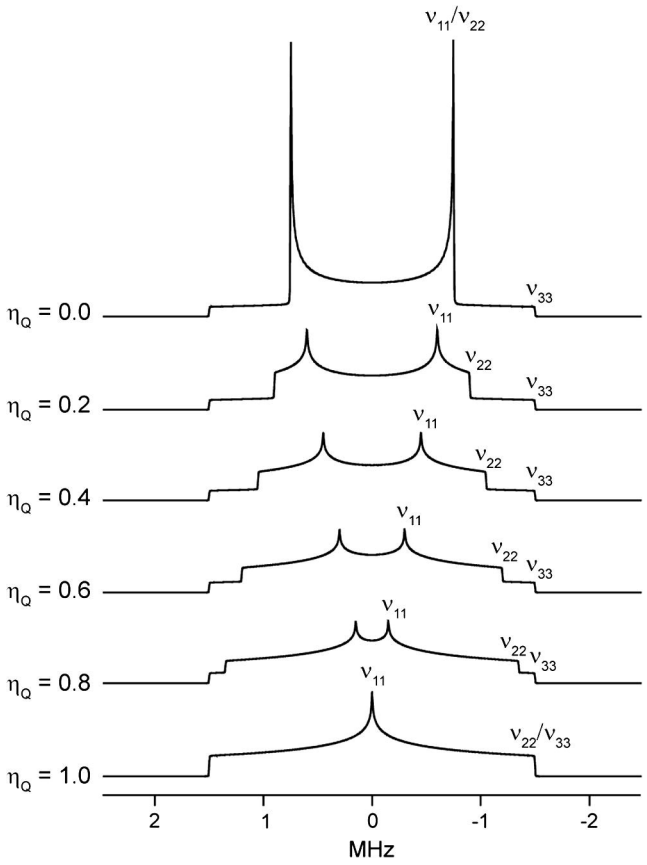

Figure 2. Simulated ^{14}N Powder Lineshapes

Lineshapes are simulated for a ^{14}N nucleus with a FOQI of C_Q=2MHz, and various values of η, as indicated. Marked are the frequencies of the discontinuities which are related to the principal components of the EFG tensor, V. The SOIQS and the chemical shift interactions are not included in these simulations. Reproduced with permission from O'Dell, 2011[5].

1.3.1.3 Biological Samples

As explained in the previous section, a particularly favourable situation for direct excitation and detection of ^{14}N exists in situations where ^{14}N is at a site of high cubic symmetry and/or where the ^{14}N site experiences rapid motion. This makes the EFG at the nucleus small, or scales the quadrupolar coupling, in the case of motions, and the entire spectrum can be excited with standard RF pulses. For instance, at the NH_4^+ site in ammonium chloride, the ^{14}N is at a site of cubic symmetry and has no quadrupolar coupling, meaning a single narrow peak is observed, even in the solid-state. In terms of biological solids, a favourable situation for direct excitation ^{14}N ssNMR exists in some common lipid-containing systems, where a ^{14}N site is found with near cubic symmetry and may be observed directly by solid-state NMR. The nitrogen nucleus present in the choline head-group of phosphatidylcholine has a tetrahedral bond geometry, with a near cubic symmetry, and rotates rapidly, meaning that the quadrupole splitting is low, generally on the order of 4-12kHz[26]. This situation reduces the difficulty associated with acquisition of ^{14}N spectra, since the spectrum is not MHz wide, and all crystallites can be excited uniformly with standard RF pulses. Spectra can be acquired in a similar manner to those acquired from deuterium, a spin-1, low-*Q* nucleus, typically by using a quadrupolar echo pulse sequence. The location of discontinuities in the static spectra, or fitting of sideband intensities under MAS allows one to extract quadrupolar parameters at individual ^{14}N sites. This favourable situation has been exploited by a number of studies to investigate the ^{14}N quadrupolar coupling at the surfaces of biological membranes, which reflects the orientation of the lipid headgroup as a function of the electrostatic potential . Changes in lipid headgroup orientation and disruption of the bilayer surface can be monitored on the addition of proteins, peptides or other membrane interacting or embedded molecules[26, 34-37].

While the direct detection methods outlined above can be used to very accurately characterise ^{14}N EFG tensor parameters, their application has some limitations. The single crystal studies, as with all investigations on single crystals, are clearly limited to that limited range of samples which can be prepared as

such. They are also limited by the number of ^{14}N sites in the molecule, compounds with more than a handful of ^{14}N sites become prohibitively difficult to assign. This precludes the general application of single crystal studies to a large number of organic and biological solids.

Considering the techniques applied to polycrystalline samples, all the direct acquisition techniques mentioned here are limited by the inherent low resolution afforded by ^{14}N powder patterns or sideband manifolds. Generally, the FOQI is far larger than the range of ^{14}N chemical shifts, leading to inevitable overlapping of resonances from different sites in molecules containing multiple ^{14}N sites. This can be a significant hindrance to interpretation of the spectra and determining the ^{14}N EFG parameters, and to date none of these methods have been applied to a sample containing more than three separate ^{14}N sites[20]. These techniques also tend to suffer from increased difficulty when investigating sites with larger ^{14}N C_Q magnitudes. Broadband pulsed methods, even those using specialised pulses, cannot excite powder patterns >~1.0MHz in a single experiment. Those methods involving stepping of the transmitter frequency over large ranges suffer from increased experimental timeframes as the C_Q increases. For these reasons, it is difficult to envision the general application of direct excitation ^{14}N NMR to samples containing many ^{14}N sites with potentially large quadrupolar couplings, such as complex biomolecules, aside from the special case of choline containing lipids discussed above.

1.3.2 ^{14}N Overtone NMR

An alternative method for obtaining ^{14}N spectra, that does not suffer from the broad linewidths and poor resolution of the ^{14}N fundamental transition, is detection of the ^{14}N overtone transition. The method involves the direct excitation and observation of the $m = -1 \leftrightarrow m = +1$ transition in ^{14}N. This is a transition forbidden by the usual selection rules, but in situations where the quadrupolar interaction is large it causes mixing of the $|+1\rangle$ and $|-1\rangle$ states, and the transition becomes weakly allowed. It is observed by pulsing at approximately twice the ^{14}N Larmor frequency, though the exact frequency of the transition is determined by the chemical shift and second order isotropic

quadrupolar shift, as seen in Figure 1. The overtone transition is not affected by the large (MHz) FOQI, only the SOQI and chemical shift interactions, meaning that ^{14}N signals are dramatically narrowed, and the pulse widths and spectral widths required to observe ^{14}N overtone signals are dramatically reduced compared to the fundamental transition. The overtone transition was first observed by NMR by Bloom & Legros[38]. Tycko and Opella subsequently demonstrated the possibility of using $^{1}H/^{14}N$ overtone CP on a static sample to determine the orientation of the peptide plane with respect to the magnetic field vector using a single crystal of NAV[39]. They also demonstrated that static ^{14}N overtone powder patterns for polycrystalline NAV were ~45 times less broad than those of the ^{14}N fundamental transition[40]. Finally, they used ^{14}N overtone separated local field experiments on a single crystal of alanyl-glycyl-glycine (AGG) to determine $^{1}H/^{14}N$ bond lengths, and demonstrated ^{14}N overtone lineshapes of NAV under spinning at various angles[41]. Wi and Frydman, and Takegoshi and coworkers independently demonstrated that the ^{14}N overtone spectrum of polycrystalline solids under MAS could by 'indirectly' reconstructed from ^{13}C lineshapes under MAS using a method known as overtone rotary recoupling (ORR)[42, 43]. Both groups observed the dephasing of ^{13}C spin echo spectra where ^{13}C nuclei were coupled to ^{14}N nuclei when pulses were applied close to the ^{14}N overtone frequency during an echo period. Plotting the dephasing fraction as a function of the ^{14}N overtone irradiation frequency allows one to construct an 'indirect' ^{14}N overtone spectrum (not to be confused with indirect detection, discussed in Section 1.4.4). Both Wi *et al*. and Takegoshi *et al*. were able to construct ^{14}N overtone spectra from the amine site in L-alanine, and the amide site in NAV that agreed favourably with simulations of the sites using previously determined values for the ^{14}N EFG tensors at these sites[42, 43]. The group of Takegoshi extended this work in a subsequent paper to investigate the EFG tensor parameters at amide ^{14}N sites in polypeptide chains of known secondary structure in order to compare the differences between EFG tensors at ^{14}N sites in an α-helix and a β-sheet[44]. The authors investigated a series of polypeptides of known secondary structure, into which they incorporated a single ^{13}C labelled residue of either alanine, leucine or valine.

They demonstrated that the on resonance ^{14}N overtone frequency (i.e. that with the highest fractional dephasing of the ^{13}C CP-MAS spectrum) was consistently greater for the same residues in an α-helical conformation than those in a β-sheet conformation, by ~200kHz. In the same study, *ab initio* calculations were performed on two of the samples investigated, one an α-helix and the other a β-sheet, both of which were geometry optimised, and the calculated differences in ^{14}N C_Q and η agreed with the observed spectra. The authors attribute the change in ^{14}N EFG parameters between the two secondary structures to the difference on the strength of hydrogen bonding. The O---H distance in the CO-NH H-bonding moiety was 0.179nm and 0.206nm in the β-sheet and an α-helix crystal structures, respectively, and the N-H distances were 0.103nm and 0.102nm, respectively. The authors posit that the longer, slightly weaker H-bond of the β-sheet alleviates the charge distribution at the ^{14}N to a greater extent than at the α-helix, meaning there is a reduced C_Q. This is the only investigation of ^{14}N EFG parameters in polypeptides of defined secondary structure in the literature, and presents important evidence that ^{14}N EFG tensor magnitude differs between the two secondary structures to an extent that they could be readily discerned in high resolution ^{14}N NMR spectra.

More recently, ^{14}N overtone studies have focussed on detection of ^{14}N overtone signals under MAS, since O'Dell and Ratcliffe's pioneering demonstration of the method in 2011[45]. In this study, it was demonstrated that the ^{14}N lineshape under MAS was narrower than that observed under static conditions, due to averaging of the chemical shift anisotropy (CSA) and dipolar couplings, and partial scaling of the SOQI. They also observed that the $+2\omega_r$ sideband of the sideband manifold was the most intense. Since then, there has been renewed interest in detecting ^{14}N overtone signals in organic and biological materials under MAS to determine ^{14}N EFG parameters. Nishiyama first demonstrated the indirect detection (discussed in greater detail in section 1.4.4) of ^{14}N overtone signals[46]. A number of polarisation transfer methods for increasing the sensitivity of the overtone transition have been proposed. CP under MAS conditions with DNP was demonstrated by Rossini and O'Dell[47], and signal enhancements from PRESTO-II[48] were demonstrated by our own

group[49]. We have furthermore demonstrated novel simulations strategies for quickly and accurately simulating ^{14}N overtone spectra[49], and demonstrated the application of double rotation (DOR) to simultaneously remove second-order quadrupolar effects and CSA and dipolar effects from the lineshapes of ^{14}N signals, recording ^{14}N overtone signals with linewidths on the order of 100s Hz rather than kHz[50].

The narrowed lines of ^{14}N overtone spectra over those of fundamental ^{14}N spectra are potentially highly beneficial in terms of increasing the spectral resolution of ^{14}N spectra. The signals from ^{14}N sites of moderate C_Q would not inevitably overlap, as is the case for the fundamental transition. Furthermore, the ^{14}N overtone shift has a contribution from the ^{14}N chemical shift, as well as the second order isotropic quadrupolar shift, meaning that the spectral range of ^{14}N overtone spectra is wider than that of the nitrogen chemical shift range which benefits resolution. However, detection of the ^{14}N overtone has drawbacks which limit its applicability in terms of complex (bio)molecules. Firstly, the overtone transition is generally very insensitive. This is a consequence of the low transition moment of this forbidden transition. Theoretically, this situation is alleviated at higher values of ^{14}N C_Q, or lower values of B_0, where the mixing of states and, therefore, the transition moment is greater, but this is not generally observed experimentally[49], and decreased resolution due to linebroadening from the increased SOQI could become an issue and offset sensitivity gains. Polarisation transfer techniques and DNP have been proposed to address this, but for complex samples with numerous ^{14}N sites with high C_Q magnitudes, experimental times are still typically prohibitive.

Secondly, overtone signals nutate at a very low rate meaning that long pulses (100s μs) are typically required for optimal excitation. Consequently, the bandwidth of excitation of these long pulses is low, and even small (<10kHz) RF offsets result in dramatically attenuated signal intensity. This makes it difficult to apply ^{14}N overtone experiments to samples that contain multiple ^{14}N sites, particularly if their ^{14}N chemical shift and C_Q are not known *a priori*.

1.3.3 Measurement of $^{13}C/^{14}N$ Internuclear Distances

While the acquisition of ^{14}N spectra in order to characterise the EFG tensor at ^{14}N sites is a particular challenge with NMR, particularly in complex materials such as biological solids, a number of methods for determining dipolar couplings between ^{14}N and other nuclei (particularly ^{13}C) exist in the literature. The dipolar coupling is directly related to the internuclear distance between nuclei, hence it can be used to determine structural restraints such as bond lengths and torsional angles. Methods for determining dipolar couplings to ^{14}N nuclei are based on the REDOR method, where rotor synchronised RF pulses are applied on the channel of one spin during a spin echo on the first spin to reintroduce the dipolar coupling between spins-½ that is averaged out under MAS[51-53]. Signal attenuation, known as dephasing, is observed in the spectrum of the first spin when the dipolar coupling is recoupled, and the dephasing fraction compared to a control spectrum ($\Delta S/S_0$) can be plotted as a function of the REDOR mixing time and fitted with simulations in order to determine the dipolar coupling, and hence the internuclear distance between the spins. This method has been extended to determine couplings between spin-½ spin->½ nuclei with methods such as TRAPDOR[54, 55], REAPDOR[56, 57] and RESPDOR[58]. The TRAPDOR method has been used to reintroduce the dipolar coupling between ^{13}C and ^{14}N spin pairs in ^{13}C labelled glycine and polyamide-6[55], and RESPDOR has been used to determine the magnitudes of dipolar couplings between all carbons and the ^{14}N nucleus in the indole sidechain of natural abundance L-tryptophan[58].

In a particularly elegant example applied to biological solids, Middleton demonstrated how TRAPDOR could be used to restrain distances between ^{13}C and ^{14}N spins in amyloid fibrils. Middleton reported the distance between ^{13}C and ^{14}N nuclei in residues on different β-strands in an amyloid fibril of the amyloidogenic fragment of human islet amyloid polypeptide ($hIAPP_{20-29}$), resolving an apparent ambiguity existing in previously published structures regarding the interdigitation of amino acid residue side groups between sheets (structures shown shown in Figure 3)[3]. This was achieved by selectively ^{13}C and ^{15}N labelling the N-terminal four residues of the peptide, and then using a

frequency selective TRAPDOR experiment to detect the $^{13}C/^{14}N$ dipolar coupling (if any) present between the ^{13}C labelled nuclei in the sidechain of a particular leucine and a ^{14}N nucleus in the asparagine side chain within the unlabelled region of an adjacent β-sheet. Detection of this coupling would only be consistent with one of the proposed structures; in the other the labelled leucine would be >8Å away, and too far to detect a dipolar coupling. The authors found that the dipolar coupling could be detected, and fitting of the TRAPDOR dephasing curve allowed them to determine the distance between the two nuclei to be 2.6Å.

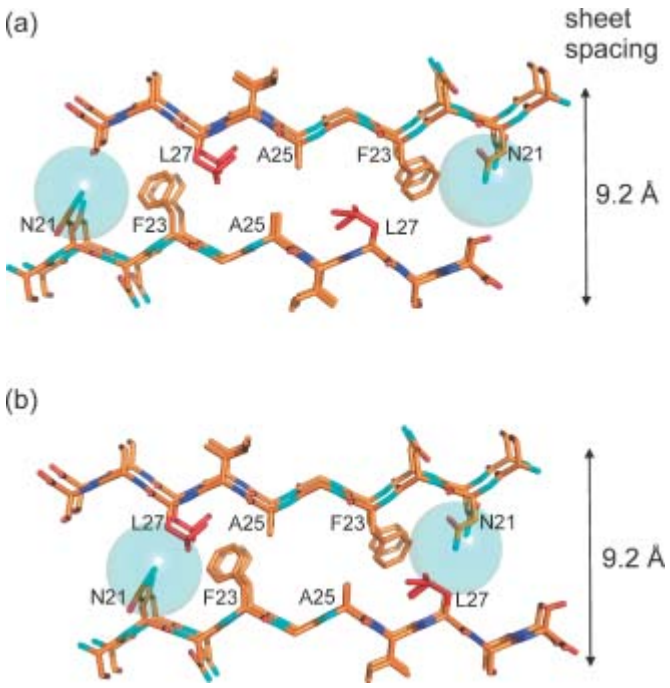

Figure 3. Alternative Structures of hIAPP$_{20-29}$ Fibrils Consistent with ssNMR Restraints

^{13}C nuclei are labelled in red, ^{15}N nuclei in dark blue, and ^{14}N nuclei in light blue. Light blue spheres represent the area within 3.5Å of the sidechain amide of asparagine-21. Only the structure in (b) would result in dephasing of the ^{13}Cα signal of leucine-21 in a ^{13}C/^{14}N TRAPDOOR experiment. Figure reproduced with permission from Middleton, 2011[3]

1.3.4 Indirect Detection of ^{14}N

1.3.4.1 The HMQC Pulse Sequence for Indirect Detection of ^{14}N

An alternative to direct detection and detection of the ^{14}N overtone for recording ^{14}N spectra for characterising ^{14}N EFG tensors in organic and biological solids is the indirect detection method. The indirect detection method was introduced independently by Gan and Bodenhausen's laboratories[1, 2], and involves the detection of ^{14}N spins via their interaction with coupled spin-½ 'spy' nuclei, either ^{13}C or ^{1}H. In this class of experiments, ^{14}N spectra are detected in the indirect dimension of a 2D experiment recorded on a spin-½ nucleus under MAS. The experiments are qualitatively and conceptually similar to liquid state heteronuclear multiple quantum coherence (HMQC) experiments[59, 60], and are referred to as HMQC experiments. The HMQC pulse sequence first demonstrated by Gan and Bodenhausen is shown in Figure 4.

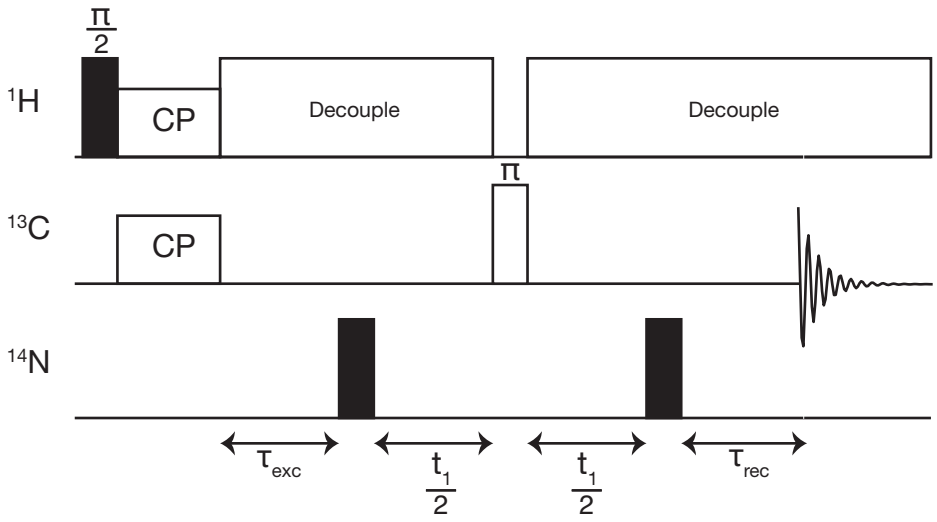

Figure 4. HMQC Pulse Sequence for Indirect Detection of Nitrogen-14
Filled rectangles on the ^{14}N channel are short (<15μs), high power (≥50kHz) pulses. t_1 evolution is centred on the ^{13}C refocussing pulse, and $\Delta t_1=\tau_r$ to ensure efficient removal of the first order quadrupolar interaction. ^{14}N SQ or DQ coherences can be selected in t_1 by phase cycling. Pulse sequence was introduced in refs [1, 2].

The ^{13}C detected pulse sequence proceeds with a $^{1}H/^{13}C$ CP step to generate transverse magnetisation on ^{13}C. This is followed by a period of free evolution, labelled τ_{exc} in Figure 4. During this period, the transverse magnetisation on ^{13}C is converted into antiphase coherence. This occurs under the evolution of not only the $^{13}C/^{14}N$ *J*-coupling, but also the $^{13}C/^{14}N$ residual dipolar splitting (RDS). The RDS is a second order quadrupole-dipole cross term that is inversely dependent on B_0, not averaged by MAS, and can be observed in the MAS spectra of ^{13}C nuclei coupled to ^{14}N as a 2:1 doublet[61]. In the alanyl-glycl-glycine (AGG) peptide studied by Gan at B_0=14.1T, the $^{13}C/^{14}N$ RDS was expected to be ≈30Hz, with $^{13}C/^{14}N$ *J*-couplings of ≤10Hz, making the RDS the dominant interaction, which is typical of organic compounds at moderate field strengths. The τ_{exc} period is set to $1/(2RDS_{CN})$ to maximise build up of

antiphase coherence. Antiphase coherence is partially converted into multiple quantum coherence on ^{14}N by the application of a high power (≥50kHz) pulse at close to the ^{14}N Larmor frequency. This evolves during a rotor synchronised t_1 period, before a second pulse on the ^{14}N channel generates antiphase coherence which evolves into in phase transverse magnetisation on ^{13}C during τ_{rec}. A π pulse on ^{13}C spins during t_1 constitutes a spin echo, and refocuses ^{13}C CSA and inhomogeneous decay of polarisation.

This method of detecting ^{14}N has a number of desirable features that make it applicable to a range of materials with multiple ^{14}N sites with various EFG tensor parameters.

Firstly, the large FOQI is removed from indirectly detected ^{14}N lineshapes so long as the magic angle is precisely set, MAS is stable and t_1 increments are rotor synchronised. This yields ^{14}N spectra whose lineshape features are dominated by the SOQI. ^{14}N peaks are centred at the sum of the ^{14}N δ_{CS}^{iso} and the ^{14}N δ_{Q}^{iso}, which is field dependent. The ^{14}N δ_{CS}^{iso} can be assumed to be the same as the ^{15}N δ_{CS}^{iso}, as the isotope effect on the nitrogen chemical shift is negligible[1, 2, 28, 32, 62]. The ^{15}N chemical shift is frequently known or can be determined from other experiments[1, 63-66]. The expression for the ^{14}N δ_{Q}^{iso} is given above in Equation (1.49), and depends on the quadrupolar product, χ_Q, given in Equation (1.50), and the Larmor frequency at the given field strength. Therefore, χ_Q can be simply determined from indirectly detected ^{14}N spectra by subtraction of the ^{14}N δ_{CS}^{iso} from the observed ^{14}N shift, to yield the δ_{Q}^{iso}, and then solving Equation (1.49) for χ_Q. An example of the type of spectra obtained from ^{14}N indirect detection methods, and the described shifts is shown in Figure 5. In this spectrum the large difference in ^{14}N δ_{Q}^{iso} for amide and amine sites is demonstrated. In order to determine ^{14}N C_Q and η more precisely than in terms of χ_Q, fitting of the ^{14}N lineshape is required. Ideally, the indirectly detected ^{14}N lineshape depends principally on the second order quadrupolar interaction which is not averaged by MAS, as well as, potentially, third order effects if the C_Q magnitude is large enough[1]. In practice, SQ ^{14}N lineshapes acquired in this

manner can be susceptible to mis-setting of the magic angle and spinning instability, which partially reintroduces the FOQI and distorts the lineshape. This, combined with the computationally demanding task of simulating such spectra to second or third order has so far precluded the fitting of indirect ^{14}N lineshapes to determine C_Q and η.

Secondly, resolution is afforded in indirectly detected experiments by the spin-½ 'spy' nuclei, which separate the indirect ^{14}N lineshapes of coupled ^{14}N spins. This also allows one to observe correlations between ^{14}N and the detected heteronuclei close in space and helps in assigning the spectra. This makes the indirect detection method more applicable to complex materials containing many ^{14}N sites than direct detection methods, where signals frequently overlap, and overtone NMR, which is limited by the bandwidth of excitation.

Thirdly, the indirect detection method does not rely on the excitation of the entire bandwidth that the ^{14}N first order quadrupolar interaction spans. It is typically sufficient to use RF pulses of ~50kHz, as measured on NH_4Cl or another ammonium salt, as demonstrated by Bodenhausen and Gan on samples with ^{14}N C_Q of ~1MHz and ~3MHz, respectively[1, 2]. While this does not fully excite all crystallite orientations, a sufficient level of SQ or DQ coherence is generated to detect 2D spectra in a few hours. Efficiencies of the experiment applied by Bodenhausen's group to samples of glycine and alanine containing only primary amines were ~15%, measured with respect to a ^{13}C spin echo without ^{14}N pulses applied[21].

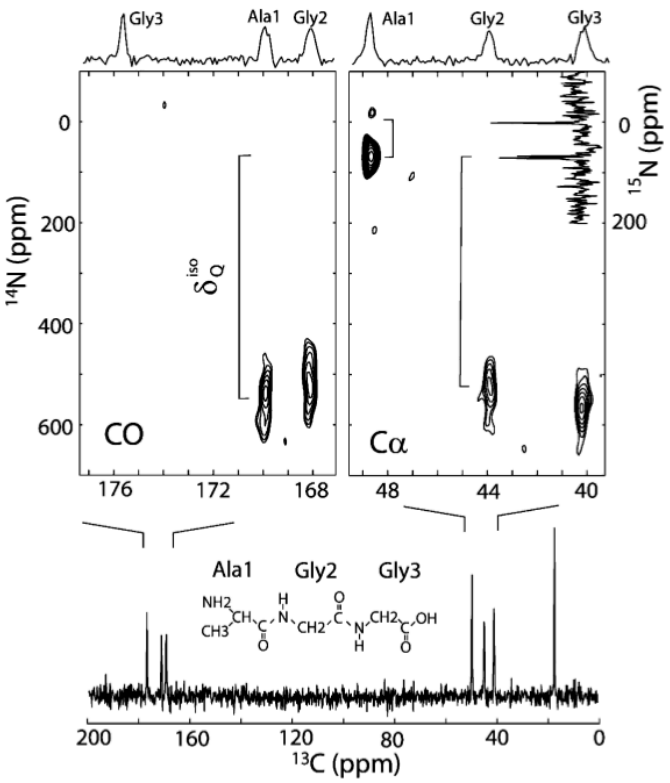

Figure 5. 2D ^{13}C/^{14}N Indirect Detection *J*-HMQC Spectrum of AGG

Acquired at 14.1T under 25kHz MAS. ^{14}N peaks are observed at the sum of the ^{14}N δ_{CS}^{iso} and the ^{14}N δ_{Q}^{iso}. The ^{14}N δ_{Q}^{iso} can be determined by subtracting the δ_{CS}^{iso}, as measured on ^{15}N. This is demonstrated by braces indicating differences in ^{14}N and ^{15}N shifts in the figure. Figure reproduced with permission from Gan, 2006[1].

1.3.4.2 Alternative and Improved Methods for Indirect Detection of ^{14}N

Since the initial application of these indirect detection pulse sequences, numerous methodological developments have been made regarding pulse sequences for the indirect detection of ^{14}N. Soon after the initial demonstration of the sequence, Bodenhausen's laboratory extended the technique to detection via ^{1}H, simply removing the CP step of Figure 4, and performing a spin echo and detecting on ^{1}H[62]. This modification has the benefits of being more sensitive due to the increased gyromagnetic ratio of ^{1}H over ^{13}C, ($\gamma_{^{13}C} / \gamma_{^{1}H} \approx 0.25$), as well as the increased natural abundance of ^{1}H, which makes the experiment highly sensitive even in unlabelled materials. The efficiency of this experiment was reduced compared to the ^{13}C detected version, however, the SQ ^{14}N filtered experiment being only 2% of a ^{1}H spin echo without ^{14}N pulses applied on glycine. Presumably, this is largely down to the reduced T'_2 of ^{1}H compared to ^{13}C and could be increased by faster MAS rates, this study being conducted at 30KHz MAS. A method involving the same interactions for coherence transfer, but more similar to a heteronuclear single quantum coherence (HSQC) experiment, where magnetisation on the spy nucleus is made to be along the *z*-axis, rather than transverse, before t_1 evolution was proposed by Bodenhausen[67] and Halpern-Manners[68] in order to obtain ^{14}N lineshapes not affected by distortion due to decay of spy nucleus magnetisation during t_1. Bodenhausen's laboratory also demonstrated that replacing the rectangular pulses on the ^{14}N channel in the HMQC sequence with trains of DANTE[69] pulses could excite a much larger bandwidth on the ^{14}N channel, resulting in increased experimental efficiencies[70, 71].

Gan's laboratory developed alternative methods for generating coherence between 'spy' nuclei and ^{14}N. The *J* and RDS couplings between ^{13}C and ^{14}N are small, requiring relatively long periods for coherence transfer during which signal is lost to T'_2 relaxation. Additionally, the RDS is inversely proportional to the applied magnetic field, and its magnitude depends on the size of the quadrupolar interaction and the relative orientation of the dipolar and quadrupolar tensors. To address these issues, Gan implemented the dipolar recoupling sequences REDOR and rotary resonance in the place of τ_{exc} and τ_{rec}

periods in the HMQC sequence[63, 64], and demonstrated a $^{1}H/^{13}C/^{14}N$ 3D spectrum using the rotary resonance recoupling[72]. Building on this idea, Cavadini *et al*. used symmetry based recoupling sequences[73] in proton detected HMQC experiments in order to recouple the $^{14}N/^{1}H$ dipolar coupling during excitation and reconversion, as well as to decoupling the $^{1}H/^{1}H$ dipolar coupling during t_1 evolution[74]. This not only increased efficiency over *J*- and RDS based HMQCs, but increased resolution in the ^{14}N dimension by increasing the ^{1}H T'_2. Nishiyama also demonstrated a HMQC emplying dipolar recoupling, applying the symmetry based recoupling sequence $SR4_1^2$ during τ_{exc} and τ_{rec} periods in a proton detected ^{14}N HMQC. Applying this pulse sequence at high MAS frequencies (up to 80kHz) and high ^{14}N RF amplitudes (up to 130kHz) 19-fold enchantments over the *J*- and RDS based HMQC performed under identical conditions were observed, with an efficiency of 9% of spin echo on a sample of glycine[75].

A number of authors have also demonstrated the extension of the indirect detection technique to indirectly detect the ^{14}N overtone transition in organic solids via both ^{1}H[46] and ^{13}C[47, 76].

1.3.4.3 Structural and Dynamic Information from Indirect Detection of ^{14}N

The Indirect detection experiments described above have been used to investigate the properties of a range of organic and biological samples from their ^{14}N spectra. The original paper of Gan demonstrated that the χ_Q at the two amide sites in a tripeptide of AGG could be calculated from the ^{14}N shift measured in the indirect dimension (as described above), without considering third order effects. The $^{13}C/^{14}N$ correlation spectrum of AGG form this paper is shown in Figure 5. It was demonstrated that the calculated χ_Q at these two sites were consistent with the ^{14}N C_Q and η determined in a previous study[1]. This study demonstrated that indirect detection experiments were possible when detected on ^{13}C, even at natural isotopic abundance, and showed that indirect detection methods could be used on complex molecules containing numerous ^{14}N sites with large (≥ 3MHz) C_Q values. The consistency of measured χ_Q values

from indirect detection experiments with known ^{14}N EFG tensor values was also demonstrated for the primary amine site in glycine[64].

^{14}N spectra can also be used to investigate dynamics in organic solids. Comparing SQ and DQ spectra of a peptide of AAG, Cavadini *et al*. were able to demonstrate that motions on a 100ns timescale were responsible for broadening of ^{14}N resonances at the amine site of alanine-1[77]. Brown et al. published a series of papers in which the properties of peptides and pharmaceuticals were studied using proton detected ^{14}N HMQC experiments with rotary resonance recoupling at high MAS frequencies[65, 66, 78, 79]. Studying a dipeptide of β-Asp-Ala, χ_Q values for the amine and amide sites were determined by indirect detection, and shown to be in good agreement with those obtained from *ab initio* calculations using an X-ray diffraction study. Furthermore, in the same peptide, they demonstrated that by varying the rotary resonance recoupling time in dipolar recoupling HMQC experiments, long range correlations between ^{1}H and ^{14}N of up to 3.5Å could be observed at long (500μs) mixing times[65]. These correlations could be potentially useful distance restraints in determining crystal packing and intermolecular contacts, particularly those involving hydrogen bonds. In a series of papers, this group used ^{1}H/^{14}N correlation spectra to investigate the molecular structure of pharmaceuticals. Studying a sample of cimetidine, a H_2 histamine receptor antagonist containing six separate ^{14}N sites, it was possible to observe ^{1}H/^{14}N correlations of increasing distance, up to 3Å, by varying dipolar recoupling times. From these correlations, relative distances of protons from ^{14}N atoms could be determined, which shed light on the hydrogen bonds which direct packing of the molecule in the solid-state[79]. The authors went on to investigate cocrystal formation in a sample of nicotinamide and palmitic acid. The formation of a hydrogen bond between the two species could be determined by indirect ^{14}N detection, where an intermolecular correlation is observed between a ^{14}N in nicotinamide, and a carboxylic acid proton in palmitic acid. Additionally, the authors note that while the ^{15}N chemical shifts of free nicotinamide and nicotinamide in complex with plamitic acid are near identical, the ^{14}N shift for nicotinamide in the two formulations are different by ~10ppm. This can be attributed to a change in the

^{14}N δ_Q^{iso}, due to the different hydrogen bonding geometries at the same nitrogen site in the two formulations, and demonstrates how sensitive the ^{14}N quadrupolar coupling is to local differences hydrogen bonding at ^{14}N sites[66]. This is in line with the ORR study of various peptides of different secondary structure conducted by Fukazawa *et.al*.[44].

1.4 Model Protein Systems for ssNMR

1.4.1 Protein G IgG-Binding Domains

Well characterised model systems are of immense use in the design and development of methods in solid-state NMR, which may be extrapolated to the study other more complex and/or less well characterised systems; particularly in terms of implementing systems for the structural elucidation of biomoleucules. Well used and extensively characterised model protein systems include the SH3 domain[80-83], ubiquitin[84] and the immunoglobulin G (IgG) binding doimains of streptococcal protein G[82, 85].

Protein G is a cell surface protein found on group G *Streptococcus*, which facilitates high-affinity binding to immunoglobulin G (IgG) mediated by three homologous domains, B1-3. Each of these domains has a highly conserved fold of ~55 amino acids which comprises a four strand β-sheet and a short α-helix[85, 86]. A solution-state NMR structure of GB3 is shown in Figure 6. These IgG-binding domains have a large amount of structural data available for them, including high-resolution X-ray and solution-state NMR data deposited in the protein data bank (PDB) [87, 88]. This allows researchers to compare structures elucidated by different methods, and evaluate the accuracy of novel techniques in resolving structural characteristics[89, 90]. These domains are also well suited to studies of protein folding and dynamics[91, 92]. The presence of the two main forms of protein secondary structure is a specific advantage for developing solid-state NMR methods to determine ^{14}N EFG characteristics, since the ^{14}N EFG tensor is thought to be highly sensitive to the different arrangements of hydrogen bonds in these two secondary structures[44]. The B1-3 domains also have desirable physiochemical characteristics in terms of their very high thermostability (T_m>85°C), without disulphide bonds, making them possible to express and purify with a high yield relatively simply[85]. They can also be prepared in a microcrystalline form, which can be an advantage for solid-state NMR. The local order found in microcrystalline samples ameliorates linebroadening from structural inhomogeneity, which can lead to very high resolution spectra, and tight crystal

packing allows one to fit more sample in the same volume, increasing sensitivity[93]. It is also possible to observe all resonances of the protein G B domains, including those from amino and carboxyl termini and loop regions at room temperature, which is unusual in proteins in general[86].

For these reasons, microcrystalline GB3 is an ideal model system for the development of ^{14}N solid-state NMR methods.

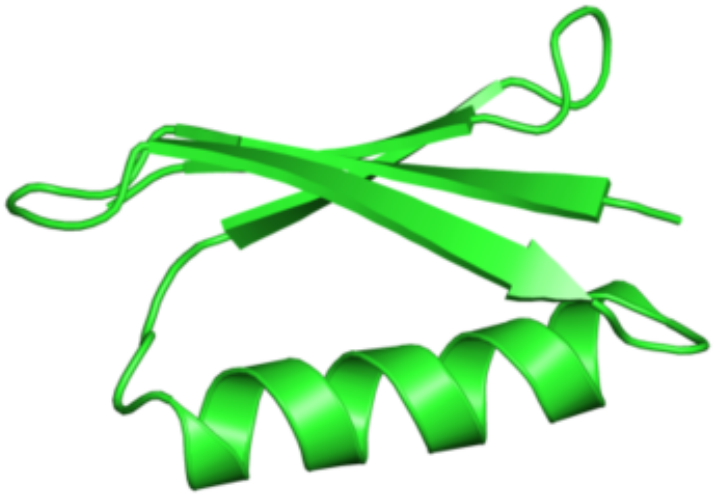

Figure 6. Solution-State NMR Structure of GB3

Solution-state NMR structure of the 56 residues of GB3. By measuring $^{13}C/^{13}C$ and $^{13}C/^{15}N$ dipolar couplings, Ulmer et. al. were able to further refine the existing 1.1Å resolution GB3 crystal-structure [39]. PDB ID: 2OED

1.5 Aims of Thesis

Solid-state NMR has, in recent years, become increasingly amenable and applicable to the structural and dynamic study of large biomolecules, such as proteins and nucleic acids. The highly naturally abundant nitrogen isotope, ^{14}N, has however remained underused in biological NMR experiments, despite nitrogen being prevalent in biological systems. This is largely due to the difficulty in working with spin-1 quadrupolar nuclei with quadrupolar couplings typically on the MHz range. This is a particular problem in large biomolecules such as proteins containing many nitrogen sites. However, the quadrupolar coupling of the ^{14}N nucleus has the potential to act as a sensitive reporter on molecular structure, and is highly sensitive to relatively small changes in local structure such as hydrogen bond geometry and length, more so than the nitrogen chemical shift. This makes it a potentially very useful reporter on secondary structure in proteins, where nitrogen is present in every amino acid residue and has specific H-bonding motifs to form different types of protein secondary structure.

Recently, there has been increased interest in the detection of ^{14}N in a range of materials, not least biological molecules, and there has been a large amount of development of direct detection, overtone and indirect detection ^{14}N ssNMR methods. However, ^{14}N NMR is yet to be established as a viable technique in biomolecular NMR, and recording of ^{14}N NMR spectra or measurement of ^{14}N EFG tensor parameters in a protein has not been achieved. In this research project, we aim to develop methods for indirect detection of ^{14}N in order to ultimately record ^{14}N ssNMR spectra in a full-length protein, GB3. This will allow, for the first time, analysis of ^{14}N EFG tensor parameters in proteins, in order to evaluate their dependence on structural features in this model system. The indirect ^{14}N detection method is the method of choice for acquiring ^{14}N ssNMR spectra in complex biomolecules, since it benefits from the sensitivity and resolution of the ‘spy’ nuclei from which signals are detected. This means that it can potentially be applied to large, complex systems containing many ^{14}N sites, while still retaining the ability to resolve signals from individual ^{14}N

sites. Correlations to 'spy' nuclei potentially facilitate the assignment of complex molecules, meaning site specific measurements can be made of ^{14}N EFG tensors in large molecules. However, the largest current limitation which has prevented the application of these methods to complex biomolecules, such as proteins, is the poor sensitivity of the method. In current HMQC sequences for detection of ^{14}N, one can typically expect to observe drops in signal intensity of ~90% as compared to an equivalent spin echo on the 'spy' nucleus. In this thesis, we focus on developing methods to increase the sensitivity of indirect ^{14}N detection methods such that they can be applicable to a model protein sample, GB3.

Specific goals of this thesis are:

- The development of methods to indirectly detect ^{14}N with increased efficiency in biological samples, using small organic molecules such as amino acids and peptides as model compounds.

- The application of these methods to provide the first measurements of ^{14}N ssNMR spectra in the model microcrystalline protein, GB3, and determine the magnitude of ^{14}N quadrupolar couplings in a full-length protein.

CHAPTER 2

EXPRESSION, PURIFICATION AND NMR RESONANCE ASSIGNMENT OF PROTEIN GB3

2.1 Introduction

In the introduction, GB3 was identified as a suitable model protein system for ssNMR method development. In this Chapter, we describe a refined protocol for the expression and purification of GB3, and provide a sequential assignment of the protein in the crystalline form that it will be used in in ^{14}N detection experiments.

GB3 is a well characterised protein in terms of its structure and dynamics, and has been investigated by a number of complementary techniques including X-ray crystallography[88], solution-state NMR[87] and ssNMR[86]. Previously published protocols for the expression, purification and crystallisation of GB3 were not found to be generally reproducible[86, 94, 95]. The purity of GB3 was found to be low after purification, typically highly contaminated with nucleic acids, and the micro-crystallisation protocols failed to produce crystalline protein. Therefore, methods to purify GB3, and to prepare GB3 in a microcrystalline form have been refined and developed, and are described herein. A number of isotope labelling strategies have been employed to produce samples suitable for both sequential resonance assignment (uniform $^{13}C/^{15}N$ labelling), and characterisation of ^{14}N sites via either ^{13}C (uniform ^{13}C labelling) or ^{1}H (no isotope labelling). We subsequently sequentially assigned the NMR resonances of GB3 in the solution-state to validate the proteins identity, and its correct folding in solution. Preparation of microcrystals suitable for ssNMR was then undertaken. The protein was finally assigned in the solid-state, as this is the form that it will be used in for ^{14}N detection experiments, which was of particular importance since we use crystals of GB3 made with a new crystallisation procedure.

It is demonstrated that expression and purification procedures produce yields of GB3 sufficient for ssNMR investigations, and that the microcrystalline material provides well resolved ssNMR spectra. The assignments of GB3 in the solid-state provide reference chemical shifts of ^{13}C and ^{15}N which are required for the implementation and interpretation of ^{14}N detection methods on GB3 later in this thesis in order to assign sites and determine the ^{14}N δ_{CS}^{iso} from ^{14}N shifts.

2.2 NMR Experiments and Assignment Strategies

In both the solution and solid-state, we use a range of 2D and 3D experiments to assign the ^{13}C, ^{15}N, and in solution, ^{1}H, resonances of GB3. These assignments are used to establish the sequential alignment residues, and the sidechain resonances of residues using experiments that provide inter and intra-residue correlations. The experiments employed, and the assignment strategies are detailed in this section.

2.2.1 Solution-State NMR Assignment

2.2.1.1 $^{1}H/^{15}N$ HSQC

The HSQC is a useful 2D experiment for detecting one-bond correlations between protons and heteronuclei[59, 96]. The experiment proceeds with a 90° pulse to protons, before polarisation is transferred to bound amide ^{15}N nuclei, using the *J*-coupling, via an INEPT transfer, where it evolves for t_1. After this time, a reverse INEPT transfers polarisation back to ^{1}H spins for detection. The resulting 2D spectrum has ^{1}H in the direct dimension and ^{15}N in the indirect. A diagram of magnetisation transfer in the HSQC experiment is shown in Figure 7A. In a typical $^{1}H/^{15}N$ HSQC, a single peak is observed for each backbone amide $^{1}H/^{15}N$ in the protein (not the N-terminal amine, which exchanges protons with the solvent too rapidly to be detectable), as well as side-chain amide $^{1}H/^{15}N$ correlations from asparagine and glutamine, and the side-chain indole group of tryptophan. This spectrum is sometimes referred to as the “fingerprint” of a protein, and can inform as to whether the protein is properly folded and whether there is any degradation or structural heterogeneity of the protein. Properly folded proteins display good peak dispersion, while poorly dispersed peaks crowded in the centre of the spectrum can indicate unfolded and/or degraded proteins. As a single resonance should be observed for each amino acid’s backbone amide; a lower number of peaks than expected combined with poor dispersion can indicate protein degradation. More peaks than expected in the HSQC can indicate the presence of multiple conformers, or degradation of the

protein, since individual peptides may have different shifts than those of the same residues in the full-length protein.

The HSQC is also an essential spectrum for resonance assignment in the solution-state, since all of the 3D solution-state experiments most frequently used for resonance assignment also evolve amide ^{15}N and ^{1}H chemical shift dimensions. This means that the ^{1}H/^{15}N HSQC can be used as a 'root spectrum' from which to trace connectivity of local nuclei.

2.2.1.2 HNCACB

The HNCACB experiment is a 3D experiment that resolves amide ^{1}H/^{15}N correlations from the protein backbone, in the manner of a ^{1}H/^{15}N-HSQC, with a third dimension that correlates each ^{1}H/^{15}N backbone amide resonance with the ^{13}C chemical shifts of the Cα and Cβ from both the corresponding residue (*i*) and the previous residue (*i*-1)[97].

As with the HSQC, this experiment is built on transfers of polarisation by INEPT transfer, exploiting the *J*-couplings between relevant nuclei. Firstly, an INEPT transfer takes place from ^{1}H to ^{15}N via the J_{HN}-coupling. A second INEPT transfer transfers polarisation from ^{15}N to ^{13}Cα. Delays are chosen so that transfer occurs both via the $^{1}J_{NC\alpha}$-coupling (~12Hz) to the ^{13}Cα of residue *i*, and, less efficiently, via the $^{2}J_{NC\alpha}$-coupling (~7Hz) to the 13Cα of residue *i*-1. Magnetisation is then transferred to the Cβ by a 90° ^{13}C pulse, where it evolves for t_1, before being transferred back to Cα by a 90° ^{13}C pulse. Simultaneous 90° pulses on ^{13}C and ^{15}N transfer magnetisation back to ^{15}N, where it evolves for t_2. A reverse-INEPT sequence transfers magnetisation back to amide protons for detection. A diagram of the polarisation transfer in the HNCACB experiment can be found in Figure 7B.

The HNCACB spectrum is extremely useful in establishing connectivity between residues in a protein, and in some favourable cases can be sufficient to entirely sequentially assign the protein.

2.2.1.3 CBCA(CO)NH

The CBCA(CO)NH is a 3D experiment frequently used as an accompaniment to the HNCACB in order to sequentially assign backbone amide resonances in a protein[98]. Similarly to the HNCACB, amide $^{1}H/^{15}N$ correlations from the protein backbone are resolved in two dimensions, with ^{13}C chemical shifts observed in the third dimension. It differs from the HNCACB in that only correlations to the Cα and Cβ of residues *i*-1 are observed in the ^{13}C dimension, since transfer between ^{15}N and $^{13}Cα$ occurs via INEPT transfer to and from the intervening ^{13}CO (though no chemical shift is evolved on the ^{13}CO). It therefore serves to remove ambiguity in the HNCACB as to whether peaks arise from intra or inter residue polarisation transfer.

Again, this experiment is built on blocks of INEPT and reverse-INEPT transfers. Initially, magnetisation is transferred from $^{1}Hβ$ to $^{13}Cβ$, where it is evolved for t_1 before being transferred to $^{13}Cα$. From here, transfer occurs via the J_{CaCO}-coupling to the adjacent carbonyl ^{13}C, and then to the backbone ^{15}N via the J_{CON}-coupling where magnetisation is evolved for t_2. A final reverse-INEPT transfers magnetisation to amide protons for detection. This sequence of polarisation transfer is shown in Figure 7C.

2.2.1.4 H(CCO)NH

The previously described HNCACB and CBCA(CO)NH experiments are used in the assignment of amide ^{1}H and ^{15}N, and $^{13}Cα$ resonances from the protein backbone. The H(CCO)NH and HCCH-TOCSY are complementary experiments that be used in order to assign aliphatic ^{1}H and ^{13}C resonances from the amino acid sidechains.

The H(CCO)NH is performed as a 3D experiment with two proton dimensions, and a ^{15}N dimension. The spectrum correlates $^{1}H/^{15}N$ resonances of a residue *i* in the first two dimensions with the resonances of the sidechain protons of residue *i*-1 in the third dimension[99]. The pulse sequence proceeds with a 90° pulse applied to protons to generate transverse magnetisation. A delay immediately follows this pulse during which the chemical shift of the

protons evolves. An INEPT transfer transfers magnetisation from the aliphatic protons to the aliphatic ^{13}C. An isotropic mixing sequence is applied to transfer polarisation between sidechain carbons, before polarisation is transferred to ^{13}CO from $^{13}C\alpha$ via an INEPT transfer. Another INEPT transfer immediately transfers magnetisation to the adjacent amide ^{15}N where a chemical shift is evolved before a final INEPT transfer generates magnetisation on amide protons, where it is detected. A cartoon of the polarisation transfer of this experiment is shown in Figure 8A.

2.2.1.5 HCCH-TOCSY

The HCCH-TOCSY is a 3D experiment with two proton dimensions and a ^{13}C dimension. The experiment begins identically to the H(CCO)NH; a 90° pulse is applied to protons, generating transverse magnetisation, a variable time t_1 period then allows chemical shift evolution, before an INEPT transfer transfers magnetisation to aliphatic ^{13}C. In the HCCH-TOCSY, a variable t_2 delay at this point allows a chemical shift to evolve on the sidechain carbons. An isotropic mixing sequence then transfers polarisation between sidechain carbons. Finally, a reverse INEPT transfer transfers polarisation back to the aliphatic protons, where it is detected[99]. The magnetisation transfer of this experiment is diagrammatically represented in Figure 8B.

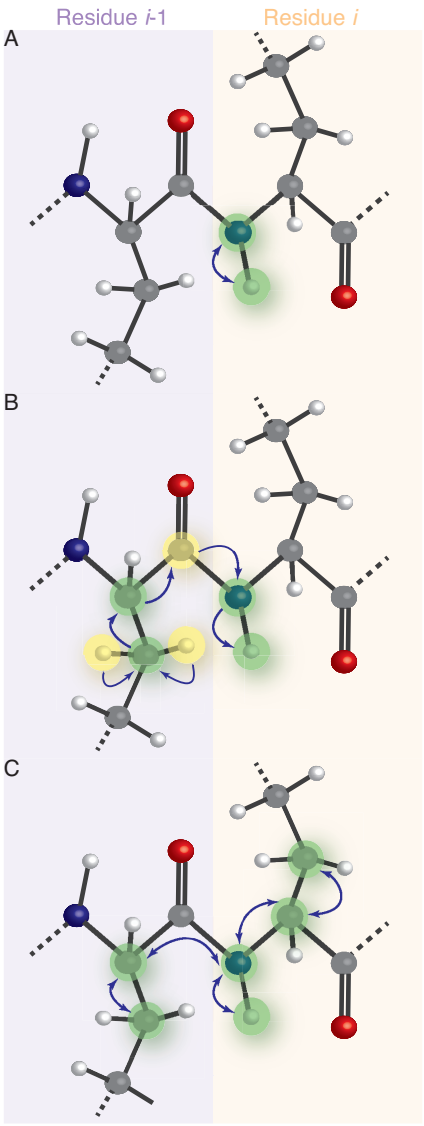

Figure 7. Transfer of Magnetisation in Solution-State NMR Experiments for Sequential Assignment of Chemical Shifts of Backbone Resides

Magnetisation transfer through two residues is diagrammatically represented for (A) HSQC (B) CBCA(CO)NH and (C) HNCACB experiments used for sequentially assigning the chemical shifts of backbone residues in a protein. Green circles indicate nuclei on which a chemical shift is evolved. Yellow circles indicate nuclei through which polarisation is transferred, but a chemical shift is not evolved. Magnetisation is transferred between nuclei via INEPT or reverse INEPT transfers in the direction indicated by blue arrows. In the case of double-headed arrows in (C), the experiment is performed as an "out-and-back" experiment. More detail on the experiments is given in the text of sections 2.2.1.1-2.2.1.3.

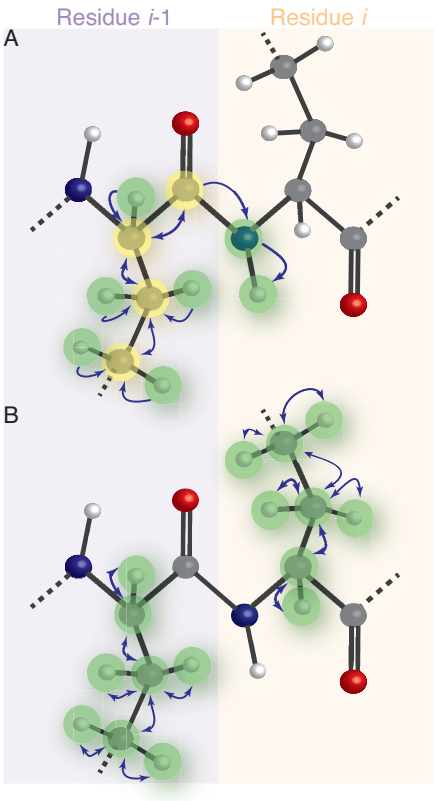

Figure 8. Transfer of Magnetisation in Solution-State NMR Experiments for Assignment of Chemical Shifts of Nuclei in Amino Acid Sidechains

Magnetisation transfer through two residues is diagrammatically represented for (A) H(CCO)NH and (B) HCCH-TOCSY experiments used for assigning chemical shifts of amino acid sidechain nuclei in proteins in the liquid state. Nuclei on which a chemical shift is evolved, or signal is detected, are marked with green circles. Nuclei on which no chemical shift is evolved, but which have magnetisation transferred via them are marked with yellow circles. Single headed arrows indicate the direction of polarisation transfer by INEPT transfer. Double headed arrows between carbon nuclei indicate isotropic mixing. Double headed arrows between protons and carbon nuclei in (B) represent the fact that the sequence begins with an INEPT transfer from protons to aliphatic carbon nuclei, and is detected on protons after reverse INEPT from the aliphatic carbons back to the bonded protons. Experiments are described further in sections 2.2.1.4 and 2.2.1.5.

2.2.2 Solid-State NMR Assignment

2.2.2.1 $^{13}C/^{13}C$ Correlation Experiments

Two-dimensional homonuclear dipolar recoupling experiments may be used under MAS to correlate $^{13}C/^{13}C$ chemical shifts in two dimensions, resolving the crowded 1D ^{13}C-spectra of proteins, and helping to assign ^{13}C chemical shifts to particular residues in proteins.

Homonuclear dipolar recoupling sequences selectively reintroduce the dipolar coupling between like spins, which is typically lost or significantly attenuated under MAS. When incorporated into 2D experiments, they give rise to cross peaks in a 2D spectrum when like spins are proximal in space. There are a large number of recoupling sequences that can be used to achieve this in biological solids, such as PDSD[100] and DARR[101] and the double quantum sequences POST-C7[102], SPC5[103] and DREAM[104, 105]. At moderate MAS frequencies (<20kHz), proton driven spin diffusion (PDSD[100]) or dipolar assisted rotational resonance (DARR[101]) experiments have become popular dipolar recoupling sequences for recording $^{13}C/^{13}C$ correlation spectra, since they are broadbanded (i.e. recoupling occurs regardless of the difference chemical shifts), and they do not suffer severely from dipolar truncation[106], making it possible to observe long range contacts useful for structural restraints in proteins. PDSD and DARR, as applied to ^{13}C in biological solids, work by coupling ^{13}C nuclei to an external 'reservoir' of strongly dipolar coupled ^{1}H nuclei to drive polarisation transfer. In the case of PDSD, during a mixing period, the $^{13}C/^{1}H$ dipolar coupling couples ^{13}C nuclei to surrounding protons. The protons are strongly coupled to each other, due to their abundance and high gyromagnetic ratio, and have a relatively small variation in their chemical shifts, meaning $^{1}H/^{1}H$ spin diffusion occurs at a higher rate. The process becomes less efficient at increased spinning speeds and field strengths, as the $^{13}C/^{1}H$ and $^{1}H/^{1}H$ dipolar couplings become better averaged, and the ^{1}H chemical shifts become better resolved. The DARR experiment works on the same principle as PDSD, but uses continuous wave ^{1}H irradiation at the rotary resonance condition ($\omega_1=n\omega_r$, where $n=1$ *or* 2) to reintroduce the $^{1}H/^{13}C$ dipolar coupling.

In both cases, in order to acquire $^{13}C/^{13}C$ correlation spectra, they are incorporated into pulse sequences where polarisation is transferred to ^{13}C by CP from protons, before an evolution period, t_1. A 90° pulse on ^{13}C stores magnetisation along the z-axis, followed by either a delay for free evolution during which $^{1}H/^{1}H$ spin diffusion occurs in a PDSD experiment, or a DARR field applied to ^{1}H spins at the MAS frequency in a DARR experiment. A final 90° pulse to ^{13}C nuclei rotates magnetisation into the *x-y* plane for detection. The length of time that the delay or DARR field is applied for (mixing time) dictates the range of $^{13}C/^{13}C$ correlations observed in the experiment. At short mixing times transfer occurs mainly between adjacent nuclei (e.g. Cα to Cβ and Cα to CO) and generally, only short-range intra-residue $^{13}C/^{13}C$ correlations are observed. At longer mixing times, long range (e.g. Cα to aliphatic sidechain ^{13}Cs) and inter-residue contacts are observed.

At higher MAS frequencies (>~20kHz), the $^{1}H/^{1}H$ dipolar couplings in proteins, necessary for spin diffusion, are attenuated significantly, to the point that broadband DARR and PDSD experiments are no longer possible, and alternative methods for dipolar recoupling must be sought. These include a number of phase and amplitude modulated sequences based on rotational resonance conditions such as PARIS[107], SHANGHAI[108], RDSD[109, 110] and MIRROR[111]. Many of these sequences are band selective (they only recouple nuclei with certain differences in chemical shift) since they rely on recoupling via second order rotational resonance conditions, $\omega_1 = n\omega_r \pm \Delta\omega_{iso}$, where n is an integer, and $\Delta\omega_{iso}$ is the difference in isotropic chemical shifts of the spins to be recoupled[109, 111]. In this work, the mixed rotational and rotary resonance (MIRROR) technique is used for experiments at MAS frequencies >20kHz, which recouples ^{13}C spins in a band selective manner[111]. The pulse sequence proceeds identically to the PDSD/DARR pulse sequence, except an RF field during the mixing period of amplitude $\omega_1 = n\omega_r \pm m\Delta\omega_{iso}$, where *n,m=0,1 or 2*, is applied. Separate experiments must be used to observe correlations between Cα and aliphatic carbons, and Cα and carbonyl carbons, due to the large chemical shift differences between these nuclei.

2.2.2.2 $^{13}C/^{15}N$ Correlation Experiments

Double cross-polarisation experiments [112, 113], where polarisation is transferred initially from protons to ^{15}N by cross polarisation, and then a second cross polarisation step transfers polarisation from ^{15}N to either $^{13}C\alpha$ or ^{13}CO spins, (commonly known as NCA and NCO experiments) correlate protein backbone nitrogens with the Cα of residue *i*, or the CO of residue *i*-1, respectively. These experiments are useful in assigning and establishing connectivity between amino acid residues in combination with $^{13}C/^{13}C$ correlation experiments. After an initial 90° pulse on ^{1}H, polarisation is transferred from ^{1}H to ^{15}N by CP, and is evolved on the nitrogen in t_1. Subsequently, CP is used to transfer polarisation to either adjacent Cα or CO nuclei, where it is detected[112]. Careful choice of offset and Hartmann-Hahn condition during the pulses applied during $^{15}N/^{13}C$ CP dictates whether polarisation is transferred to the Cα or the CO.

2.2.2.3 NCOCX

An NCOCX experiment is a useful experiment to correlate the backbone ^{15}N nuclei of a protein with the Cα of residue *i*-1 via the CO of residue *i*-1. It is helpful in establishing connectivity between residues, in conjunction with $^{13}C/^{13}C$ correlation spectra, particularly in situations where chemical shift degeneracy in NCO and/or NCA experiments make sequential assignment difficult. It can be performed as a 2D experiment, without evolution on the CO nuclei, or in three dimensions with evolution on CO, which can significantly reduce spectral crowding to reduce ambiguity and facilitate sequential assignments. After a 90° pulse on ^{1}H, CP is used to transfer polarisation from ^{1}H to ^{15}N. CP is then used to transfer polarisation to the CO of residue *i*-1, in the manner of a NCO experiment. Subsequently, a suitable dipolar recoupling sequence such as DARR or MIRROR is applied to transfer polarisation from the CO to proximal ^{13}C nuclei, where the signal is detected. The length of the dipolar recoupling time dictates the range of polarisation transfer, and the observed correlations depend on the properties of the dipolar recoupling sequence employed.

2.3 Materials and Methods

2.3.1 Expression and Purification of Recombinant GB3

2.3.1.1 Transformation of Competent Cells with GB3 Vector

Dr. Donghan Lee of The Max Planck Institute for Biophysical Chemistry, Gottingen, provided a pET expression vector, containing the 56 residues comprising the 3rd IgG-binding domain of streptococcal protein G (GB3), under the control of a T7 promoter, and including an ampicillin resistance gene. BL21 Competent cells (Promega) were transformed with the vector. To a 100µl aliquot of BL21 cells, which had been thawed on ice, was added 1µl of 10ng/µl GB3 plasmid. The solution was gently agitated, and returned to ice for 30mins. The reaction was heated for 45s at 42°C, and returned to ice for a further 2 minutes. An aliquot of 900µl SOC media, preheated to 42°C was added, before incubation in a shaker at 240rpm, 37°C for 45mins. Under sterile conditions, 200µl cells were spread onto agar plates containing 100µg/ml ampicillin, and incubated at 37°C for 14 hours. After this time, plates were stored at 4°C to arrest growth.

Glycerol stocks of competent cells were prepared from the transformed cells. A 10ml overnight culture of the transformed cells was prepared by adding a single transformed colony from the agar plate, using a sterilized inoculation loop, to 10ml LB media containing 100µg/ml ampicillin. This was incubated at 37°C for 14 hours, with shaking at 200rpm. After this time, provided the optical density of the overnight culture at 600nm (OD_{600}) exceeded 1.7, 1ml aliquots of the overnight culture were added to sterilized cyrovials containing 1ml 80% glycerol, then immediately frozen in dry ice and stored at -80°C.

2.3.1.2 Overexpression of Unlabelled GB3

Glycerol stocks of BL21 cells transformed with the GB3 vector were used to create bacterial cultures overexpressing GB3. Initially, a 10ml overnight culture of transformed cells was prepared. This contained 10ml LB media inoculated with 100µg/ml ampicillin, and a 3-4µl aliquot of GB3-transformed

cells from a frozen glycerol stock. This was incubated at 37°C with 170rpm shaking for 15 hours. The overnight culture was then transferred to 1L LB media, also containing 100µg/ml ampicillin, and continued to be incubated at 37°C with 170rpm shaking. The OD_{600} of the culture was measured at regular intervals using a Thermo Scientific Nanodrop, and once OD_{600} of the culture reached 0.7 (indicating a mid log-phase bacterial growth rate), expression of the GB3 vector was induced by addition of IPTG to a final concentration of 1mM. After 4 hours, cells were harvested by centrifugation at 6900g for 20mins at 4°C. Pelleted cells were stored at -20°C before purification.

2.3.1.3 Overexpression of Labelled GB3

GB3 was also expressed uniformly labelled with ^{13}C and ^{15}N for the purposes of characterising and assigning the protein resonances. Samples uniformly labelled with ^{13}C were produced for ^{14}N indirect detection via ^{13}C experiments. Labelled GB3 was expressed from the same glycerol stocks of GB3 described above. An overnight culture of 20ml LB was prepared, containing 100µg/ml ampicillin, and a 3-4µl aliquot of GB3 glycerol stock, and incubated at 37°C with 170rpm shaking for 15 hours. After this time, cells were pelleted by centrifugation at 4,000g for 5min and an expansion culture was prepared by transferring the cells into 100ml sterilised 1xM9 media (6g/L Na_2HPO_4, 3g/L KH_2PO_4, 0.5g/L NaCl,), enriched with 0.01mg/ml biotin, 0.01mg/ml thiamine, 4mM MgSO4, 2mM CaCl2, 2mg/ml ^{13}C-Glucose, 1mg/ml ^{15}N-NH_4Cl and 100µg/ml ampicillin, at 37°C with 170rpm shaking (1mg/ml ^{15}N-NH_4Cl was replaced with 1mg/ml unlabelled NH_4Cl for uniformly ^{13}C labelled growths). Growth of bacterial cells was monitored by regularly measuring the OD_{600} of the culture; once OD_{600}=0.7, the culture was used to inoculate a further 400ml of enriched M9 media, resulting in a 500ml final volume. GB3 expression was induced with IPTG at a final concentration of 1mM when OD_{600}=0.7. Cells were harvested after a further 4 hours by spinning at 6900g for 20mins at 4°C. As with unlabelled growths, pelleted cells were stored at -20°C before purification. Purification of GB3 was carried out identically for labelled and unlabelled cell pellets.

2.3.1.4 Purification of GB3

Attempts to express and purify GB3 by methods previously published[86, 94, 95] were made, but were not found to produce high yields of protein, or effectively remove protein and nucleic acid contaminants. Hence, a novel protocol for the expression and purification of GB3 was devised.

2.3.1.5 Enrichment of GB3 containing Fraction

The cell pellet was re-suspended in 30ml phosphate buffered saline (PBS) and lysed by sonication using a Heat Systems sonicator with a stud probe at power level 7 for 10 minutes, with a 20 seconds on/ 30 seconds off cycle. The insoluble fraction was pelleted by centrifugation at 16,000g, 4°C for 20 minutes.

The soluble protein containing supernatant was heated to 80°C for 5mins in a waterbath and then immediately chilled on ice for 10 minutes. Precipitated matter from the heating was pelleted by centrifugation at 20,000g, 4°C for 30 minutes. GB3 was isolated from the soluble fraction by ammonium sulphate ($(NH_4)_2SO_4$) precipitation at 25°C. $(NH_4)_2SO_4$ was added to the soluble fraction to 60% and stirred for 2 hours. Precipitated proteins were removed by centrifugation at 5500*g* for 20 minutes. $(NH_4)_2SO_4$ was then added to the soluble fraction up to 90% saturation with stirring for 2 hours. At this $(NH_4)_2SO_4$ concentration, GB3 was precipitated, and was recuperated by centrifugation at 5500*g* for 20 minutes, before being dissolved in 10ml 25mM Tris, pH8.0.

2.3.1.6 Anion Exchange Chromatography

Anion exchange chromatography was used to remove contamination from nucleic acids. Prior to this step, the protein solution was desalted on a GE Healthcare PD-10 desalting column to allow efficient binding to the anion exchange column. The PD-10 column was equilibrated with 5 column volumes of 25mM Tris pH8.0 (25ml total), before loading 2.5ml protein solution. The protein was then eluted with 3.5ml Tris.

A GE Healthcare HiTrap Q HP 5ml column packed with Q-sepharose on an ÄKTA Prime liquid chromatography system was used for anion exchange.

The column was equilibrated with 2 column volumes of 25mM Tris pH8.0, before loading the GB3 solution. Applying a gradient of NaCl concentration from 0-1M, in Tris buffer, over 25ml, eluted the sample.

2.3.1.7 Crystallisation of GB3

Microcrystals of purified GB3 were prepared by precipitation at 4°C. Purified GB3 was desalted and buffer exchanged into 25mM bis-Tris, pH 6.5, and concentrated to at least 20mg/ml, by centrifugal filtration in a 5,000Da MWCO Vivaspin spinning at 3700g. Hexylene glycol (Sigma) was added to the GB3 solution to a 60% volume. Microcrystals formed immediately, and were subsequently incubated for at least 3 days at 4°C. The sample was subsequently left open at 25°C for 48 hours to allow evaporation of water from the sample, and increase crystal yield.

2.3.2 Solution-State NMR Experiments

2.3.2.1 Sample Preparation and Experimental Parameters

A sample containing 500µl 1.3mM purified, uniformly $^{13}C/^{15}N$-labelled GB3 in 25mM bis-Tris, pH6.5, containing 10% D_2O, was used to acquire spectra for the characterisation and assignment of GB3 in solution. All experiments were performed on a Varian INOVA 600MHz spectrometer with a triple resonance cryo-probe at 25°C. The experiments that were performed are described in sections 2.2.1.1-2.2.1.5. The experimental parameters used for each are given in Table 1 below.

Table 1. Experimental Parameters Used in Solution-State NMR Experiments

Experiment	*^{1}H Acquisition Time (ms)*	*^{15}N Acquisition Time (ms)*	*^{13}C Acquisition Time (ms)*	*2^{nd} ^{1}H Acquisition Time (ms)*	*Mixing Time (ms)*
HSQC	107.0	32.0	-	-	-
HNCACB	107.0	16.0	53.0	-	-
CBCA(CO)NH	107.0	16.0	53.0	-	-
H(CCO)NH	107.0	22.0	-	7.8	-
HCCH-TOCSY	107.0	-	53.0	97.0	15.5

2.3.3 Solid-State NMR Experiments

2.3.3.1 Sample Preparation and Experimental Parameters

Microcrystalline $^{13}C/^{15}N$ labelled GB3 (9mg) was packed into a 1.6mm Agilent rotor for ssNMR assignment. All experiments were performed on a Varian DD2 600MHz spectrometer, equipped with a 1.6mm fast-MAS triple resonance probe. Experiments were carried out with VT gas temperature regulated at -10°C. Spinning frequencies of between 35-40kHz were used for experiments on the 1.6mm probe. The experiments performed are described in sections 2.2.2.1-2.2.2.3.

2.3.3.2 $^{13}C/^{13}C$ Correlation Experiments

For $^{13}C/^{13}C$ correlation experiments using MIRROR mixing, MAS frequencies of 40kHz were used, ^{1}H to ^{13}C CP was for 1ms, the ^{13}C field was matched to a 100kHz proton field with a tangential ramp of the CP condition from 95%-105%. A proton field of either 37kHz or 24kHz was used during MIRROR mixing periods in order to recouple $^{13}C\alpha$ and aliphatic ^{13}C, or $^{13}C\alpha$ and ^{13}CO spins, respectively, by matching the $n=1$, $m=-1$ MIRROR condition. SPINAL-64[114] decoupling was applied at 140kHz during t_2 acquisition. Experiments were performed with 50 and 100ms MIRROR mixing times in order to observe one bond and longer range contacts, respectively. In each experiment, 8 scans per t_1 increment, and 481 complex increments were acquired using the States-TPPI method[115]. Evolution times were 26.4ms in t_2 and 13.63ms in t_1, which allowed the signal to decay completely. The recycle delay was 3s. Data were processed in NMRPipe[116], where -10Hz negative linebroadening and a 5Hz Gaussian window function with a maximum at the last point of the FID were applied in F_1 and F_2, before zerofilling to 4096 points in both dimensions, and 2D Fourier transform.

2.3.3.3 $^{13}C/^{15}N$ Correlation Experiments

NCO and NCA experiments were performed at 35kHz MAS. For ^{1}H to ^{15}N CP, a ^{15}N field was matched to a 63kHz ^{1}H field for 1.1ms, with a tangential

ramp of the CP condition from 90%-110%. For both NCA and NCO, CP from ^{15}N to ^{13}C was achieved with a ^{15}N field of 52kHz (1.5xω_r), and a ^{13}C field of 87kHz (2.5xω_r), applied for 7.5ms. The condition was tangentially ramped from 95%-105%. SPINAL-64 decoupling was applied at 140kHz during t_1 evolution and acquisition. 128 complex points with 16 scans per increment were acquired for each experiment, using States-TPPI[115], with a recycle delay of 2.5s. Evolution times were 14.5ms and 22.3ms in F_1 and F_2. Data were processed in NMRPipe, with -10Hz negative linebroadening and a 5Hz Gaussian window function with a maximum at the last point of the FID applied in F_1 and F_2; data were zerofilled to 4096 points in both dimensions, and 2D Fourier transformed.

2.3.3.4 NCOCX

Two dimensional NCOCX experiments with a MIRROR mixing times of 50ms were conducted to detect nearest neighbour ^{13}CO/^{13}CX contacts. ^{1}H to ^{15}N and ^{15}N to ^{13}C CP parameters were identical to those of the NCO experiment. A ^{1}H field of 19kHz was applied during the mixing period, which is equal to the n=1, m=*-1* MIRROR condition for CO and Cα at this field and spinning rate. Decoupling during evolution times and acquisition was using SPINAL-64 at 140kHz. 96 complex points were acquired using States-TPPI[115], with 64 scans per increment and a recycle delay of 2.5s. Experiments were processed with -5Hz negative linebroadening in either dimension, before both dimensions were zerofilled to 2048 points and Fourier transformed. A 3D experiment was performed under the same conditions, with a MIRROR mixing period at 16kHz for 100ms, in order to detect both nearest neighbour and longer range contacts. The experiment was performed with 64, 32 and 16 points in t_3, t_2 and t_1, corresponding to evolution times of 12.6ms, 6.2ms and 25.4ms respectively in t_1, t_2 and t_3. States-TPPI was used to acquire complex points, and the recycle delay was 2.75s. Data were processed in NMRPipe. A sine-bell window function was applied in all dimensions. Dimensions 1 and 2 were zerofilled to 1024 points before Fourier transformation. In F_1 (^{15}N), linear prediction was applied to double the number of points, before zerofilling to 64 points and Fourier transformation.

All ssNMR spectra were referenced to DSS using a secondary reference of adamantane, where the downfield peak was referenced to 40.48ppm[117]. ^{15}N dimensions were indirectly referenced to DSS using the downfield resonance of adamantane, according to [118].

2.4 Results and Discussion

2.4.1 Expression and Purification of GB3

Soluble unlabelled and labelled ($^{13}C/^{15}N$ and ^{13}C) GB3 has been expressed, purified, and converted into a microcrystalline form, suitable for solid-state NMR studies. Figure 10A shows the results of SDS-PAGE analysis of the expression and purity of GB3 for pre and post induction cell lysates, and sonication and heating stages of the purification procedure. Following induction with IPTG, a band at a molecular weight of ~6kDa, consistent with the molecular weight of GB3 is observable in the post-induction cell lysate, which is not visible in the pre-induction lane. A very intense band at a molecular weight of ~12kDa is present in the post sonication supernatant, but not the pre-sonication pellet. This band may be consistent with anomalous running due to the gel being overloaded, or GB3 forming a dimer in these buffer conditions and concentrations; all purification stages frequently enrich this band at ~12kDa. Similarly to the sonication step, the supernatant lane after heating the cell lysate shows an intense band at ~12kDa, while the pellet from after the heating step does not contain this band, suggesting the cell lysis, heating and centrifugation steps effectively precipitate and remove solid matter from the cell lysate, thereby enriching GB3, which remains in solution.

Figure 10B shows the SDS-PAGE analysis of enrichment of GB3 using $(NH_4)_2SO_4$ precipitation. The 60% $(NH_4)_2SO_4$ pellet lane demonstrates that at this $(NH_4)_2SO_4$ saturation, many of the high molecular weight contaminants present in the post heating supernatant are precipitated and removed from solution, and the intense band at ~12kDa is not present. The lane of precipitated and pelleted material after the $(NH_4)_2SO_4$ saturation is increased to 90% shows a very intense, overloaded band running at ~12kDa with fewer protein contaminants than found in the post heating supernatant, indicating that the enrichment of GB3 by this method is effective. The final lane containing supernatant after the 90% $(NH_4)_2SO_4$ cut has no bands of protein visible, demonstrating that all the protein material has been precipitated in the $(NH_4)_2SO_4$ cuts, and no GB3 remains in solution.

The final stage of protein purification was to use anion exchange chromatography to separate nucleic acid contaminants from GB3 after the $(NH_4)_2SO_4$ precipitated pellet was solubilised and desalted. While the protein is relatively pure after precipitation with 90% $(NH_4)_2SO_4$, it remains contaminated with nucleic acids after resolubilisation. This is evidenced by the A_{260nm}/A_{280nm}, which was typically ≥ 1.7 when measured at this stage of purification. The chromatograph from anion exchange chromatography is shown in Figure 9. Observed are two peaks, one which is split into a doublet and elutes ~0.2M-0.45M NaCl, and a broad peak that elutes >0.6M NaCl. The first peak to elute had an $A_{260nm}/A_{280nm} \leq 0.6$, indicating most absorbance arose at 280nm from absorption of protein, rather than nucleic acids. This peak contained purified GB3, which ran on a gel at ~12kDa. This was used for solution-state NMR, and crystallisation for ssNMR, after desalting and buffer exchange. The second peak, at >0.5M NaCl concentration, was determined to contain solely nucleic acids, demonstrated by its typical $A_{260nm}/A_{280nm} \geq 2.0$ and no observed protein on an SDS-PAGE gel, and was discarded.

Yields of GB3 were approximately 90mg/L growth on unlabelled LB media, and approximately 45mg/L when grown labelled on M9 media.

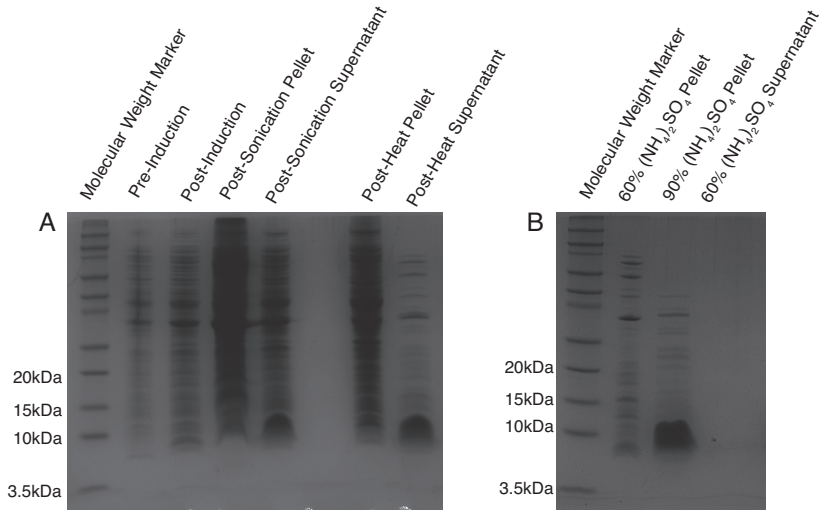

Figure 10. SDS-PAGE Results of GB3 Expression and Purification

(A) SDS-PAGE gel showing protein composition of pre and post induction E.coli cell lysates, and the pellets and supernatants obtained after sonication of the lysate and heating the lysate to 80°C for 5 minutes. (B) SDS-PAGE gel of material precipitated by addition, to the heated cell lysate, of NH_4SO_4 to 60% and 90% saturation, and the supernatant after $(NH_4)_2SO_4$ was added to 90%.

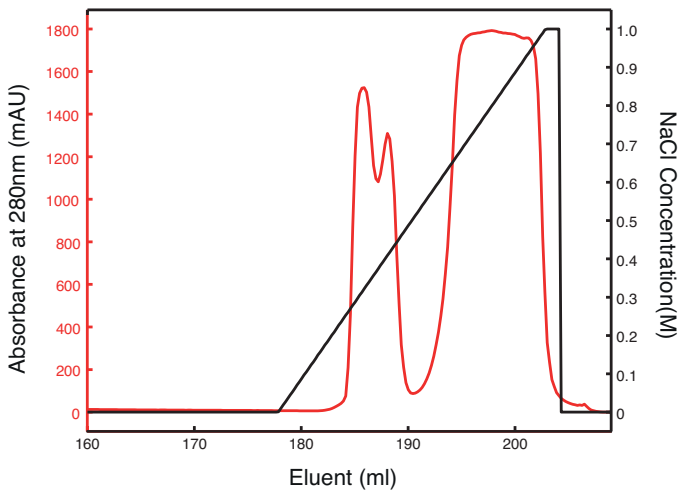

Figure 9. Anion Exchange Chromatograph of GB3

Material precipitated from the GB3 cell lysate (from a 1L growth) was dissolved in 10ml 25mM Tris, pH8. This was loaded on to a 5ml HiTrap HP column packed with Q-sepharose, equilibrated with 25mM Tris, pH8. The concentration of NaCl (black trace) was increased from 0M-1M over 25ml. The A_{280} (red trace) was measured as the NaCl concentration was ramped.

2.4.2 Resonance Assignment of GB3 in the Solution-State

2.4.2.1 $^{1}H/^{15}N$ HSQC

The $^{1}H/^{15}N$ HSQC spectrum of GB3 was recorded in order to assess folding and stability of the recombinantly expressed protein. The assigned spectrum is shown in Figure 11. The HSQC spectrum of GB3 has well dispersed peaks with no doubling of peaks, indicative of a well-folded protein without any observable structural polymorphism.

54 peaks arising from backbone amides are present in the HSQC, which is one less than would be expected for the 56-residue construct; the N-terminal methionine residue is not expected to be observed, as the positively charged amine NH_3 of the N-terminal residue rapidly exchanges protons with water in solution, and so is not observed. However, a peak from glutamine-2 is also not observed, suggesting that the N-terminal methionine has been cleaved during purification, meaning that glutamine-2 is the new, not observed, N-terminus. An additional peak, not arising from the protein backbone, is observed at a ^{1}H shift of 10.5ppm. This resonance is assignable to the indole group of the sole tryptophan amino acid in GB3. Since the observed proteolysis does not cause any apparent misfolding of the protein, or give rise to any structural heterogeneity, the removal of the N-terminal methionine does not present an issue for use of GB3 as a model system for characterisation and NMR method development.

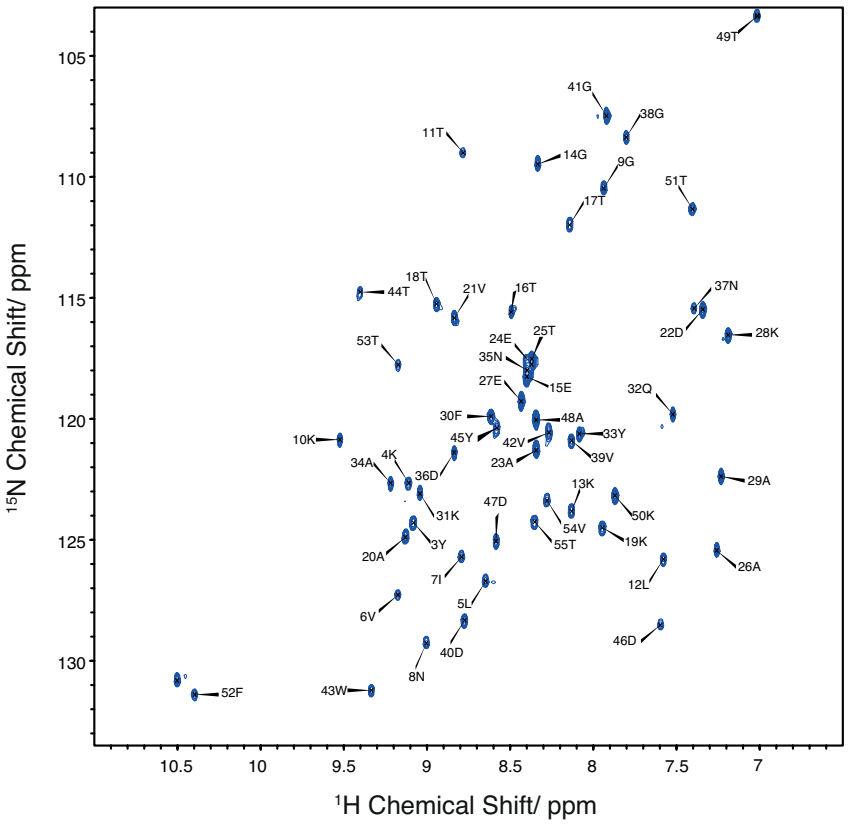

Figure 11. Assigned ^{1}H/^{15}N HSQC Spectrum of GB3

Acquired with a sample of 1.3μM GB3 in 25mM bis-Tris at pH6.5, 25°C. Assigned peaks are labelled with the amino acid type and order in the protein sequence. Sequential assignment was achieved using 3D HNCACB and CBCA(CO)NH experiments.

2.4.2.2 Sequential Assignment of GB3 Backbone Resonances

Three-dimensional, triple-resonance experiments were conducted in order to sequentially assign to the backbone ^{1}H, ^{15}N and ^{13}Cα resonances of GB3, and to subsequently assign ^{1}H and ^{13}C resonances to the sidechains of amino acids in the sequence.

HNCACB and CBCA(CO)NH spectra, shown in Figure 12, were used to sequentially assign backbone ^{1}H, ^{15}N and ^{13}Cα resonances. The HNCACB is principally used in order to establish sequential connectivity between resonances. When displayed as a strip plot of spectra for different ^{15}N planes, such as that in Figure 12, when at the ^{15}N plane of a residue *i*, four ^{1}H/^{13}C correlations are observed in each plot which are cross peaks between the Hα of residue *i*, and the Cα and Cβ of both the same residue (*i*), and the previous residue (*i*-1). The Cα peaks appear with positive contours, and Cβ peaks with negative contours, and they can be told apart. Furthermore, peaks arising from the Cα and Cβ of residue *i*-1 typically appear with less intensity than those from residue *i*, as magnetisation has been transferred to the ^{13}C via INEPT transfer which is optimised for the one bond transfer from residue *i*, rather than the two bond *J*-coupling from the ^{15}N of residue *i*-1. Sequential assignment is achieved by aligning two strip plots for different ^{15}N planes such that the intense Cα and Cβ peaks in one plot align with the weak Cα and Cβ peaks, respectively, in another plot. The dashed lines in Figure 12 guide the eye to illustrate this process.

The CBCA(CO)NH spectrum, also shown in Figure 12 in green contours, aids in sequential assignment of backbone resonances of the protein when it is overlaid onto the HNCACB spectrum. For the ^{15}N plane of residue *i*, only ^{1}H/^{13}C cross peaks from residue *i*-1 are observed in the CBCA(CO)NH, since polarisation transfer proceeds initially from amide ^{15}N nuclei via the adjacent ^{13}CO of residue *i*-1 before Cα and Cβ transfer. It is simply used to remove any ambiguity in the HNCACB spectrum regarding which ^{15}N/^{13}C cross-peaks arise from the Cα and Cβ of residue *i*, and which from residue *i*-1.

Using this assignment strategy, it has been possible to unambiguously assign 96% of backbone amide ^{1}Hα and ^{15}N chemical shifts, and 98% of ^{13} Cα

and ^{13}Cβ chemical shifts of the 56 residue protein. Carbon and amide nitrogen chemical shifts may be found in Table 2, and the chemical shifts of amide protons are found in Table 3. Only methionine-1 and glutamine-2 ^{1}Hα and ^{15}N shifts are not observed. In terms of ^{13} Cα and ^{13}Cβ chemical shifts only those of methionine-1 are not observable. This supports the theory that methionine-1 has been cleaved from the protein, and confirms that glutamine-2 is not observable in the HSQC due to rapid exchange of the amine with the solvent as discussed previously.

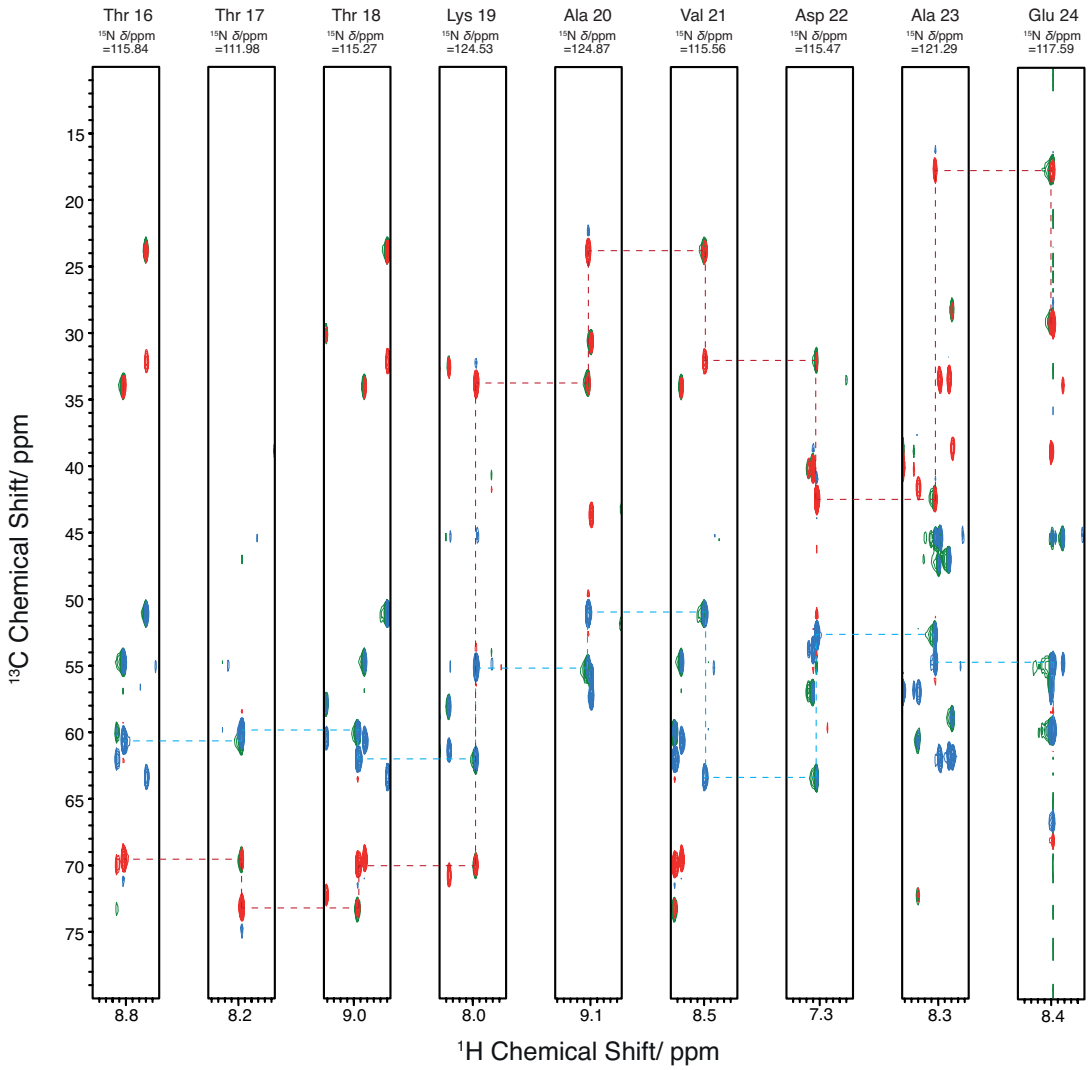

Figure 12. Strip plot of Aligned Section of HNCACB and CBCA(CO)NH Spectra of GB3

Arranging the spectra in a strip plot allows sequential assignment of resonances. Each plot shows the Cα (blue) and Cβ (red) of both the residues i (strong intensity) and i-1 (weak intensity). Aligning residues of one plot's intense Cα and Cβ peaks to the weak Cα and Cβ peaks of another plot facilitates sequential assignment. Such alignments are indicated by blue dotted lines (Cα) and red dotted lines (Cβ) between peaks of adjacent strip plots. The CBCA(CO)NH spectrum is also plotted in green contours. In this spectrum, only inter-residue correlations are observed, which can help to remove ambiguity when the intensities of i and i-1 contacts are similar.

2.4.2.3 Assignment of GB3 Sidechain Resonances

Having sequentially assigned backbone amide ^{1}H and ^{15}N, and $^{13}C\alpha$ and $^{13}C\beta$ resonances using the HNCACB and CBCA(CO)NH spectra, the H(CCO)NH and HCCH-TOCSY spectra were used to assign ^{1}H and ^{13}C resonances from the sidechain groups of GB3.

The H(CCO)NH and HCCH-TOCSY spectra are shown in Figure 13. Assignment of sidechain residues proceeds by initially using the H(CCO)NH spectrum to identify proton sidechain chemical shifts in a residue specific manner. The H(CCO)NH spectrum is three dimensional, and is typically displayed with the two proton dimensions along *x* (directly detected amide protons) and *y* (indirectly detected aliphatic protons), and amide ^{15}N along the *z*-axis. Once the backbone has been assigned, the assigned $^{1}H/^{15}N$ HSQC is used to identify the amide $^{1}H/^{15}N$ resonance of each residue. One may then view a particular amino acids side chain ^{1}H resonances in the H(CCO)NH at the corresponding ^{15}N plane of the spectrum. At the amide proton chemical shift of the residue *i* along *x* on this plane, peaks will be observed that correspond to a correlation with the aliphatic protons of residue *i*-1. The chemical shift of these protons may be read from the *y*-axis. An example of this process is demonstrated in the H(CCO)NH spectrum in Figure 13.

Knowledge of the sidechain proton chemical shifts gleaned from the H(CCO)NH is then used with the HCCH-TOCSY spectrum in order to assign ^{13}C chemical shifts of amino acid sidechains. The 3D HCCH-TOCSY spectrum is arranged with proton chemical shifts in the first two dimensions, and ^{13}C in the third. Marking the sidechain proton chemical shifts obtained from the H(CCO)NH in both proton dimensions of the HCCH-TOCSY spectrum allows one to filter through ^{13}C planes of the experiment, finding planes where correlations between all marked proton shifts are observed, with one on the diagonal. The proton shifts on the diagonal belong to the same proton in both ^{1}H dimensions, and the chemical shift of the ^{13}C to which this proton is bound is that of the plane where it is found. This process is illustrated in the HCCH-TOCSY spectrum shown in Figure 13.

Using the H(CCO)NH and the HCCH-TOCSY, is has been possible to assign 54% of Hα, 50% of Hβ, 51% of other sidechain proton resonances, and 46% of ^{13}C sidechain resonances, excluding previously assigned Cα and Cβ resonances. Proton resonances can be found in Table 3, and those of the sidechain carbons in Table 2.

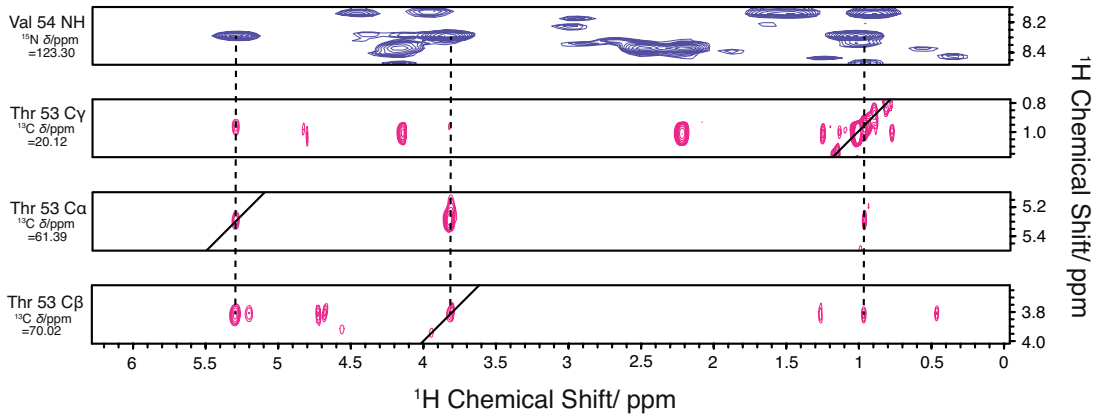

Figure 13. Strip Plots of H(CCO)NH and HCCH-TOCSY Spectra of GB3

The H(CCO)NH spectrum (purple contours, top panel) correlates the backbone amide $^{1}H/^{15}N$ chemical shifts of a residue (obtained from the assigned HSQC) with the ^{1}H sidechain chemical shifts of the residue i-1. In the strip shown, the $^{1}H/^{15}N$ resonances (along y and z-axes) from valine 54 is correlated with the α, β and γ proton chemical shifts of threonine 53 on the x-axis. These chemical shifts may be used to find the α, β and γ ^{13}C chemical shifts using the HCCH-TOCSY spectrum (pink contours, lower panels). Dashed lines mark the proton sidechain shifts of threonine 53 on x and y-axes. ^{13}C planes where cross-peaks are observed where the marked peaks intersect reveal the ^{13}C chemical shifts of aliphatic carbons bound to these protons. On planes where identical proton chemical shifts are correlated on the diagonal, the ^{13}C chemical shift is that of the carbon bound to that proton. The ^{13}C chemical shift values of threonine 53 sidechains are marked on each strip of the plot. Three strips are found with peaks where all the proton shifts intersect; these are at the aliphatic ^{13}C chemical shifts of threonine 53.

Table 2. Chemical Shifts of Assigned Amide Nitrogen and Carbon Nuclei in GB3 in the Solution-State. Values are given in ppm.

Residue	N_{amide}	*Cα*	*Cβ*	*Cγ*	*Cδ*	*Cε*
1 Met						
2 Gln		55.84	34.77	30.6		
3 Tyr	124.27	57.29	43.61			
4 Lys	122.8	55.29	36.3	25.83	29.17	
5 Leu	126.7	52.86	42.59	27.38	25.95, 25.12	
6 Val	127.27	62.09	32.91	20.93		
7 Ile	125.71	60.64	39.3	18.49, 28.49	14.32	
8 Asn	129.26	50.96	37.93			
9 Gly	110.46	44.59				
10 Lys	120.88	59.11	32.54	25.83	29.16	42.5
11 Thr	108.99	62.08	69.83	21.47		
12 Leu	125.8	55.1	43.5	27.67	24.35, 25.18	
13 Lys	123.8	54.09	34.84	24.87		43.2
14 Gly	109.48	45.31				
15 Glu	118.26	54.7	34.11	35.42		
16 Thr	115.84	60.62	69.54	19.21		
17 Thr	111.98	60.02	73.25	20.92		
18 Thr	115.27	62.06	69.99	19.24		
19 Lys	124.53	55.13	33.79	24.63	28.8	42.13
20 Ala	124.87	51.05	23.83			
21 Val	115.56	63.42	32.1	19.61		
22 Asp	115.47	52.67	42.46			
23 Ala	121.29	54.88	17.74			
24 Glu	117.59	59.89	29.38	37.71		
25 Thr	117.55	66.74	68.06	21.4		
26 Ala	125.4	54.86	17.65			
27 Glu	119.25	59.79	29.18			
28 Lys	116.53	59.73	32.47	25.57	29.74	42.24
29 Ala	122.37	54.78	17.59			
30 Phe	119.93	56.76	37.69			
31 Lys	123.08	59.88	31.65			
32 Gln	119.83	58.99	28.3	33.46		
33 Tyr	120.6	61.82	38.58			
34 Ala	122.66	56.28	17.98			
35 Asn	117.85	56.86	38.88			
36 Asp	121.38	57.09	40.14			
37 Asn	115.44	53.82	40.15			

38 Gly	108.37	47.01				
39 Val	120.93	61.78	33.4	21.3		
40 Asp	128.34	52.88	43.02			
41 Gly	107.5	45.43				
42 Val	120.6	62.08	33.57			
43 Trp	131.23	57.82	30.16			
44 Thr	114.84	60.45	72.32	22.39		
45 Tyr	120.35	57.06	41.6			
46 Asp	128.49	51.86	43.17			
47 Asp	125.04	56.41	42.24			
48 Ala	120.06	55.19	18.4			
49 Thr	103.34	60.49	70.17			
50 Lys	123.18	56.53	29.39	25.26		43.6
51 Thr	111.33	62.29	72.11			
52 Phe	131.4	57.38	42.6			
53 Thr	117.86	61.42	70.92	20.08		
54 Val	123.37	58.05	32.53	20.27		
55 Thr	124.22	61.41	70.82	20.25		
56 Glu	133.75	58.28	32.81			

Table 3. Chemical Shifts of Assigned Protons in GB3 in the Solution-State. Values are given in ppm.

Residue	H_{amide}	*Hα*	*Hβ*	*Hγ*	*Hδ*	*Hε*
1 Met						
2 Gln		5.08	2.27	1.93, 2.05		
3 Tyr	9.1					
4 Lys	9.12	5.25	2.05, 1.91	1.35, 1.50	1.61	
5 Leu	8.65	5	1.14, 0.77	0.89	0.55, 0.51	
6 Val	9.18	4.2	2.09	0.86		
7 Ile	8.8	4.12	1.57	1.33, 0.73, 0.61	0.77	
8 Asn	9.01	5.3	2.55, 2.99			
9 Gly	7.94	4.47, 4.09				
10 Lys	9.53	4.1	1.84	1.49	2.99	
11 Thr	8.79	4.39	4.22	1.15		
12 Leu	7.58	4.46	1.41, 1.60	1.36	0.9	
13 Lys	8.13	5.14	1.91, 1.75	1.46		2.99
14 Gly	8.34	4.27, 4.17				
15 Glu	8.4	5.63		2.19		
16 Thr	8.84	4.74	3.95	0.48		
17 Thr	8.15	5.82	4.3	1.22		
18 Thr	8.94	4.69	3.8	0.46		
19 Lys	7.95	5.36	1.76	1.54, 1.28	1.68	2.98
20 Ala	9.13	4.83	1.36			
21 Val	8.5	4.15	2.22	1.01		
22 Asp	7.34					

23 Ala	8.35					
24 Glu	8.38	3.89	1.84, 2.00	2.21		
25 Thr	8.37	3.71		1.26		
26 Ala	7.26					
27 Glu	8.44					
28 Lys	7.2	3.7	1.78	1.34, 1.55	1.6	2.88
29 Ala	7.24					
30 Phe	8.62					
31 Lys	9.04	4.12	1.55,			
32 Gln	7.52	4.07	2.22	2.43		
33 Tyr	8.08					
34 Ala	9.22					
35 Asn	8.39					
36 Asp	8.84					
37 Asn	7.39					
38 Gly	7.81					
39 Val	8.13	4.17	1.78	0.68		
40 Asp	8.78					
41 Gly	7.93					
42 Val	8.27					
43 Trp	9.34					
44 Thr	9.41	4.84	4.29	1.2		
45 Tyr	8.59					
46 Asp	7.6					
47 Asp	8.59					
48 Ala	8.35					
49 Thr	7.02					
50 Lys	7.87	4.22	2.03	1.24, 1.42		-2.97
51 Thr	7.41					
52 Phe	10.4					
53 Thr	9.18	5.29	3.81	0.97		
54 Val	8.28	4.49	0.32	0.38		
55 Thr	8.36		3.83	0.97		
56 Glu	7.91					

2.4.2.4 Secondary Structure Analysis of GB3 in Solution

The secondary structure of GB3 in solution was analysed using the chemical shift index (CSI) method in order to confirm the correct folding of the protein.

The CSI method is a simple method to characterise the secondary structure of a protein by analysing the chemical shift differences of residues with respect to a set of random coil values. It is based on the observation that the ^{1}Hα, ^{13}Cα, ^{13}Cβ and ^{13}CO chemical shifts of individual amino acids in the two major types of secondary structure found in proteins (α-helices and β-sheets) have consistent characteristic deviations from the chemical shift values of the same amino acids found in a random coil conformation [119, 120]. The CSI method is estimated to be typically a 75%-80% accurate predictor of a proteins secondary structure[121]. For ^{1}Hα nuclei in a protein, if the observed chemical shift is less than (upfield) or greater than (downfield) the random coil value (±0.1ppm), it is assigned a score of -1 or +1, respectively. It is assigned 0 if it is within range. Regions in which a group of 4 or more '+1's are observed without an intervening '-1' are interpreted as a β-strand, while a group of four or more '-1's are interpreted as an α-helix. Remaining regions are designated as in a random coil conformation. The CSI procedure proceeds identically for ^{13}Cα and ^{13}CO chemical shifts, except that the shifts are in the opposite direction from the reference values, so a stretch of 4 '+1's are interpreted as a region of α-helix, and '-1's, β-strand. The reference random coil values from which chemical shift deviations were calculated were taken from reference [122]. From ^{13}Cβ chemical shifts, only regions of β-strand can be reliably identified, and ^{13}Cβ chemical shifts deviate from random coil values in the opposite direction from ^{1}H and other ^{13}C nuclei so that regions of β-strand are predicted by a region of four or more '+1's. Where the CSIs of more than one type of nucleus are available for a protein, a consensus CSI can be calculated from the sum of the available CSIs. Bar charts plotting the of the deviation of ^{13}Cα, ^{13}Cβ, and ^{1}Hα chemical shifts (Δδ) from the random coil values for each residue of GB3, and the consensus CSI for the assigned nuclei in GB3, are shown in Figure 14. It predicts a secondary structure for the protein of two β-strands at the N-terminus (residues

2-9 and residues 11-20) and one at the C-terminus of the protein (residues 38-55), flanking an α-helical region (residues 22-36). This is in good agreement with structures of GB3 reported in the literature[87, 88], and exemplified by the GB3 structure shown in Figure 6. This finding suggests that the expressed GB3 is correctly folded, and therefore lends weight to the assignment of GB3 in the solution-state. The presence of the two main different types of protein secondary structure is a useful feature for a model system for ^{14}N studies of proteins. ^{14}N NMR is very sensitive to changes in H-bond length and the H-bonding arrangement of nitrogens should differ in regions of different secondary structure in proteins.

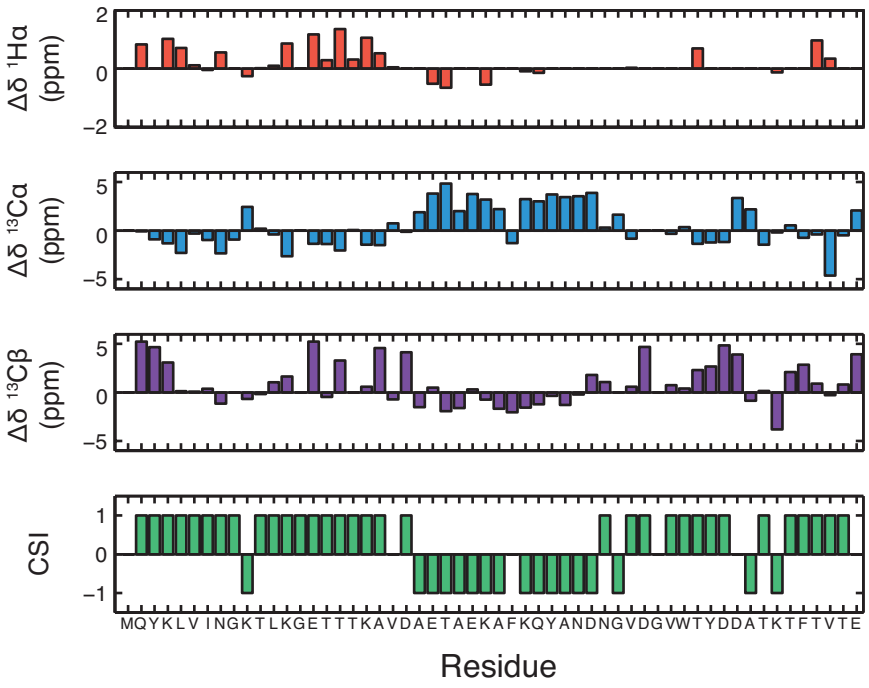

Figure 14. CSI Plot of Assigned Residues of GB3 in Solution

Chemical shift deviations (Δδ) for assigned ^{1}Hα, ^{13}Cα and ^{13}Cβ nuclei for each amino acid in GB3 are plotted as bar charts. Residues with a negative deviation for the ^{1}Hα are likely to be in an α-helix, and those with a positive deviation are likely to be involved in a β-sheet. For the ^{13}Cα, a negative Δδ is likely to be a β-sheet, and a positive Δδ is likely to be an α-helix. The consensus CSI is plotted in green, where positive indicates a residue in a β-sheet conformation, and negative a residue in an α-helix. The CSI predicts residues 2-9, 11-20, and 38-55 to be in a β-sheet conformation, and residues 22-36 to be α-helical.

2.4.3 Resonance Assignment of GB3 in the Solid-State

A series of solid-state NMR experiments were conducted on a sample of U-^{13}C/^{15}N microcrystalline GB3 in order to sequentially assign ^{13}C and ^{15}N chemical shifts of residues in the solid-state.

Two-dimensional ^{13}C/^{13}C (MIRROR) and ^{13}C/^{15}N (NCA and NCO) correlation spectra were used as a starting point for assignment in the solid-state. Representative MIRROR spectra correlating Cα and aliphatic carbons, and CO and Cα carbons, and NCA and NCO double CP spectra are shown in Figure 15. The assignment strategy using these 2D experiments is indicated by dashed lines connecting resonances in the spectra in Figure 15, and proceeds as follows: Firstly, peaks in the ^{13}C/^{13}C spectra can be attributed to particular types of amino acid on the basis of their chemical shifts. For specific types of amino acid, it is possible to identify peaks at characteristic chemical shifts corresponding to correlations between the ^{13}Cα and/or aliphatic ^{13}C nuclei on the sidechain of the same amino acid. Once spin systems have been defined in this manner, the NCA and NCO spectra can be employed to establish connectivity between adjacent amino acids. A given ^{13}Cα nucleus will give rise to a peak in the NCA spectrum corresponding to a correlation between itself and the amide ^{15}N of the same residue (*i*). In the NCO spectrum, the nitrogen shift of residue *i* will correlate with a ^{13}CO resonance from the previous residue (*i*-1). Taking this ^{13}CO chemical shift, and using the ^{13}C/^{13}C correlation spectra, one can observe a correlation between the ^{13}CO, and the ^{13}Cα of residue *i*-1. This method is demonstrated by dotted lines for two separate spin systems of three amino acids in Figure 15.

Despite the excellent resolution and sensitivity of the 2D spectra, chemical shift degeneracy, particularly in the poorly dispersed ^{13}C carbonyl region, and the ^{15}N dimension, precluded the unambiguous sequential assignment of more than a handful of residues by this method. In order to complete the assignment of microcrystalline GB3, a 3D NCOCX spectrum using *n=1,m=-1* MIRROR dipolar recoupling mixing period of 100ms was used to sequentially assign backbone ^{13}C and ^{15}N residues. Strip plots from this spectrum demonstrating the sequential assignment strategy are shown in

Figure 16. On the ^{15}N plane corresponding to the backbone amide ^{15}N chemical shift of a given residue, *i*, two ^{13}C/^{13}C correlations are observed in the ^{13}Cα/^{13}CO region of the spectrum. One, which correlates the Cα and CO of residue *i*-1, is more intense than the other, which correlates the Cα of reside *i* with the CO of residue *i*-1. The reduced intensity of the *i*-1 resonance reflects the increased distance between these nuclei, and hence the reduced magnitude of the dipolar couplings between these sets of spins. By arranging the strip plots such that the weak peak in one plot aligns with the intense peak in the next, spin systems can be sequentially assigned. Using this method, 100% of observable ^{15}N amide and ^{13}Cα chemical shifts (excluding the cleaved Met-1) have been sequentially assigned. A proportion of sidechain aliphatic carbon atoms have also been assigned from cross peaks with assigned ^{13}Cαs in the aliphatic ^{13}C/^{13}C correlation MIRROR spectra. The solid-state assignments are given in Table 4.

Incorporation of the band selective MIRROR recoupling into the 3D NCOCX scheme has the very useful feature that it resolves both inter and intra-residue ^{13}CO/^{13}Cα contacts for each amino acid, allowing complete assignment of a protein in a single 3D experiment. This type of strategy has been demonstrated previously in NCOCX experiments in the moderate spinning frequency regime (<20kHz), using long ^{13}C/^{13}C DARR mixing periods to observe inter and intra residue contacts, including sidechain nuclei, after ^{15}N/^{13}C transfer[123]. Here, we demonstrate the same principle in a faster MAS regime. At fast MAS frequencies (> 20kHz), use of a band selective homonuclear recoupling sequence such as MIRROR during mixing dictates the CO/CX transfers that are observed. These band-selective sequences will only recouple dipolar couplings between CO and CX nuclei with a certain chemical shift differences. Using the band-selective MIRROR sequence with a condition that specifically promotes ^{13}CO/^{13}Cα transfer, only ^{13}CO/^{13}Cα contacts will be observed, but the long range intra-residue ^{13}CO(*i*)/^{13}Cα(*i*+1) dipolar coupling is selectively reintroduced as well as the ^{13}CO(*i*)/^{13}Cα(*i*) dipolar coupling, and this long range coupling is less susceptible to dipolar truncation by larger short range couplings than it may be with a broadbanded recoupling sequence[106]. The spectrum presented here therefore allows one to fully assign the backbone

resonances of a protein from a singe spectrum since on the ^{15}N plane of each amino acid, a contact with the ^{3}Cα of residue *i* and *i*-1 is observed meaning that time consuming extra 3D experiments such as the NCACO, which can be employed to remove ambiguity in the NCOCX spectrum, are not necessary.

Using the assignment strategies outlined above, 100% of the expected backbone ^{15}N, ^{13}Cα and ^{13}CO residues of GB3 have been sequentially assigned. The solid-state assignments are presented in Table 4. Consistently with the solution-state data on GB3, resonances from methionine-1 are not observable in any correlation spectra, suggesting it has been cleaved during purification. The carbonyl of the C-terminal glutamate-56 is also not observable in any ^{15}N/^{13}C spectra, since all spectra with a CO dimension have polarisation transferred from the ^{15}N of residue *i*+1. No doubling of peaks or unassignable peaks are observed in any of the presented spectra, suggesting that the microcrystals formed by GB3 using the protocol described have a highly uniform structure and packing. The ssNMR assignments given here agree with those given in the published assignment of microcrystalline GB3 to within ±0.2ppm[86]. This suggests that we have the same crystalline form as that of the published assignment. The solid-state and solution-state NMR assignments determined here for the Cα and, where available, Cβ, nuclei agree to within ±0.3ppm for >80% of nuclei. However, differences of up to 1ppm are observed between the solution and solid-state assignments of Cα chemical shifts. For instance, the Cα chemical shifts of residues asp-47 and ala-48 both differ by ~1ppm between solution and solid-state assignments. The difference in this case may be due to the residues being in an unstructured loop region, as predicted by the observed CSI, and by known GB3 X-ray and solution and solid-state NMR structures[86-88]. In other cases, such as glutamate-15, which is not expected to be in a disordered region, and whose Cα chemical shift differs by 0.92ppm between solution and microcrystalline forms, the large difference in chemical shifts potentially indicates the location of intramolecular crystal packing interactions.

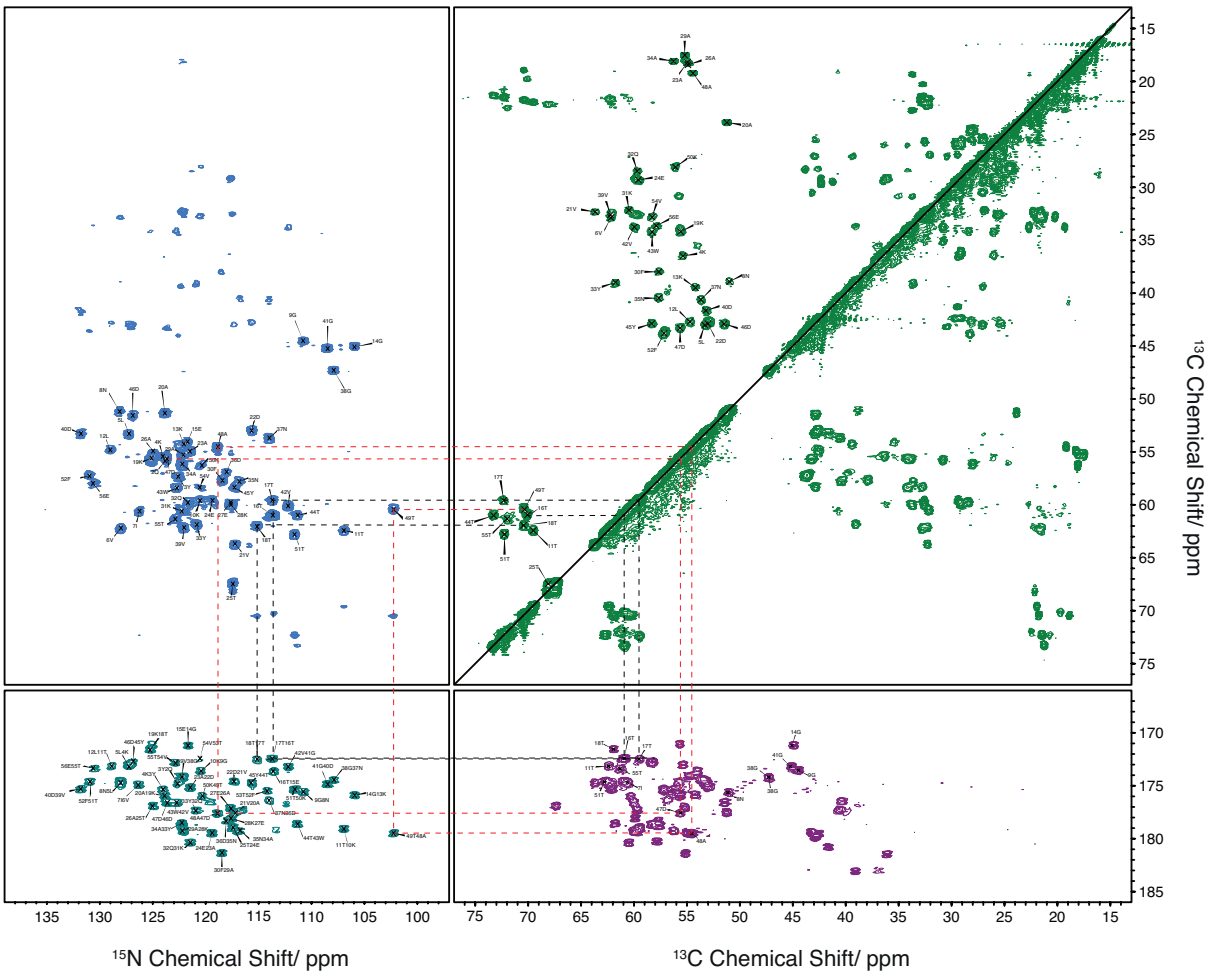

Figure 15. 2D Solid-State ^{13}C/^{13}C and ^{13}C/^{15}N Correlation Spectra of GB3

Cα-Aliphatic (light green) and Cα-CO (purple) ^{13}C/^{13}C correlation spectra recorded at 40kHz MAS with 100ms mixing with ^{1}H irradiation at the n=1, m=-1 MIRROR condition and Cα (blue) and CO (dark green) regions of ^{13}C/^{15}N correlation spectra recorded at 35kHz MAS. Dotted red and black lines guide the eye through the sequential assignment of residues Asp-47-Thr-49, and Thr-16-Thr-18, respectively.

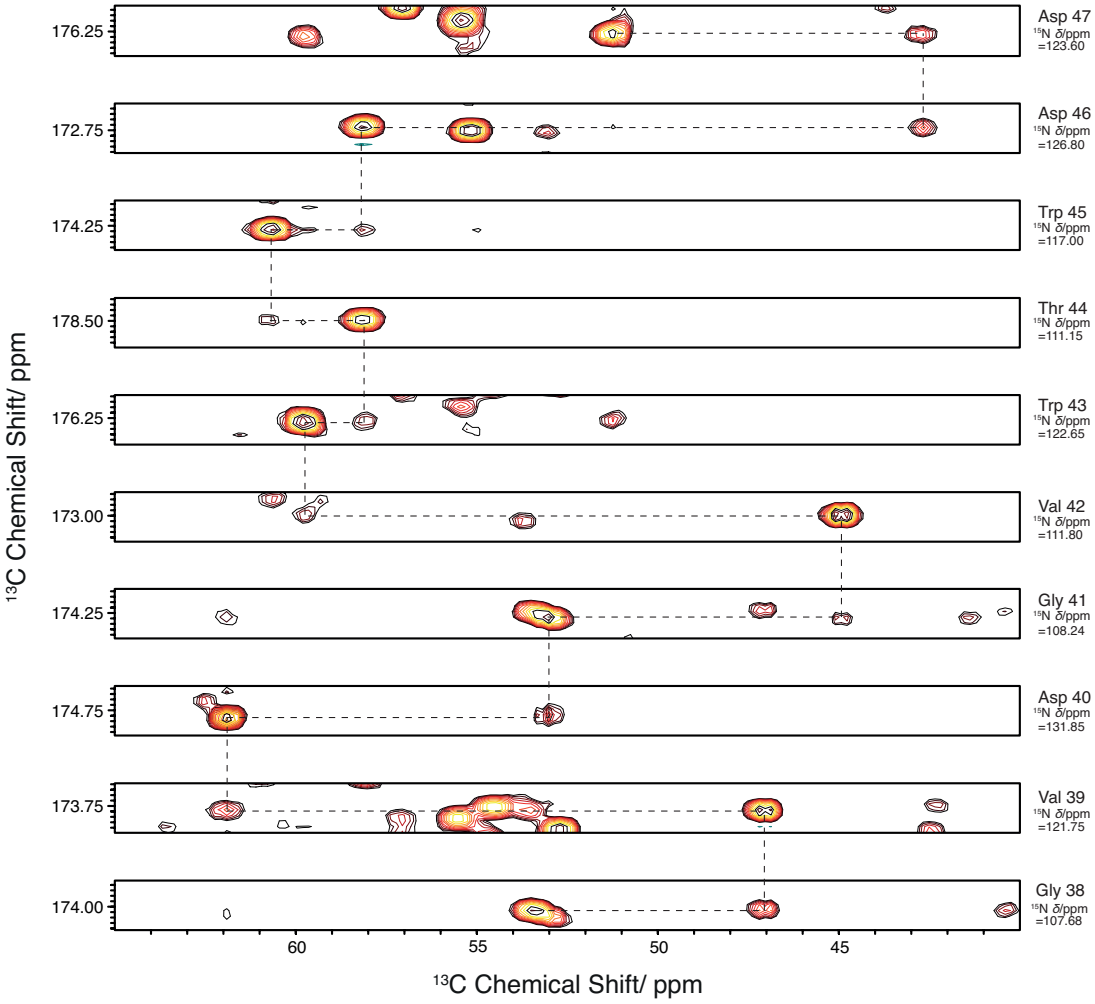

Figure 16. Strip Plot From 3D NCOCX Spectrum of GB3

Representative assigned strip plot from 3D NCOCX Spectrum of GB3. On each nitrogen plane, two ^{13}CO/^{13}Cα cross peaks are observable. The more intense of these peaks is a correlation between the ^{15}N of residue i and the ^{13}Cα of residue i-1 via the ^{13}CO of residue i-1. The less intense correlates ^{15}N of residue i with the ^{13}Cα of residue i via the ^{13}CO of residue i-1. Alignment of less intense peaks with more intense peaks between ^{15}N planes can establish connectivity between residues, as demonstrated by the black dashed lines.

Table 4. Chemical Shifts of Assigned Amide Nitrogen and Carbon Nuclei in Microcrystalline GB3. Values are given in ppm.

Residue	N_{amide}	*CO*	*Cα*	*Cβ*	*Cγ*	*Cδ*	*Cε*
1 Met							
2 Gln	123.83	174.31	55.50				
3 Tyr	122.34	175.08	57.11	43.80			
4 Lys	124.06	172.50	55.40	36.50			
5 Leu	127.22	174.31	53.12	43.01			
6 Val	128.01	174.82	62.05		22.16		
7 Ile	126.26	174.31	60.36				
8 Asn	128.04	175.34	50.81	38.91			
9 Gly	110.77	173.27	44.24				
10 Lys	120.26	178.70	59.48				
11 Thr	106.93	172.76	62.22	69.51			
12 Leu	128.92	173.79	54.73	42.71			
13 Lys	122.01	175.34	53.69				
14 Gly	105.67	170.69	44.77				
15 Glu	121.68	173.27	53.78				
16 Thr	113.37	171.72	60.66				
17 Thr	113.80	172.50	59.27	72.31			
18 Thr	114.79	171.20	61.73	70.44			
19 Lys	124.97	175.86	55.43	30.92			
20 Ala	123.64	177.41	51.09	23.88			
21 Val	117.24	174.57	63.52	32.32			
22 Asp	115.39	175.08	52.74	42.89			
23 Ala	121.50	179.48	54.74	18.20			
24 Glu	119.25	177.15	59.37	29.30			
25 Thr	117.64	176.64	67.25	68.10			
26 Ala	124.95	176.64	55.00	18.44			
27 Glu	117.61	180.00	59.59				
28 Lys	117.22	178.96	59.95				
29 Ala	122.06	181.03	54.92	17.52			
30 Phe	118.42	178.19	57.49	37.99			
31 Lys	122.01	179.77	60.36	32.17			
32 Gln	121.66	176.89	59.49				
33 Tyr	120.83	178.19	61.67	38.98			
34 Ala	122.16	178.96	56.07	18.09			
35 Asn	116.56	177.93	57.50	40.48			
36 Asp	117.89	175.86	56.36				
37 Asn	113.75	174.05	53.39	40.61			
38 Gly	107.63	173.79	47.05				
39 Val	121.77	174.82	61.89	32.60			
40 Asp	131.79	174.57	53.01	41.65			
41 Gly	108.30	172.90	44.94				
42 Val	111.85	176.62	59.76	33.77			

43 Trp	122.65	178.70	58.11		
44 Thr	111.16	174.31	60.69	73.26	
45 Tyr	117.04	172.76	58.15	42.88	
46 Asp	126.80	176.14	51.24	42.89	
47 Asp	123.65	177.19	55.40	43.29	
48 Ala	118.47	179.22	54.29	19.23	
49 Thr	101.89	175.86	60.06	70.34	
50 Lys	120.31	175.08	55.88	28.18	
51 Thr	111.16	174.31	62.55	72.23	
52 Phe	130.97	175.34	56.89	43.56	140.00
53 Thr	113.67	172.24	60.83		
54 Val	120.09	178.62	58.09		
55 Thr	122.64	173.08	61.09	71.99	
56 Glu	130.43		57.72		

2.5 Conclusion

In this Chapter, the expression, purification and crystallisation of GB3 for ssNMR have been described. The expression protocol for this protein is efficient, and produces a high yield of protein which can be prepared as microcrystals that give rise to very high resolution ssNMR spectra. Proteins have been produced variously unlabelled and uniformly labelled with ^{13}C, and ^{13}C and ^{15}N in order to characterise the protein using a variety of different methods, and in order that samples with specific labels for various ^{14}N correlation experiments can be prepared.

The assignment of GB3 has been undertaken in both soluble and microcrystalline preparations. This was in order to initially validate the identity and correct folding of the protein in solution, ensuring its suitability for use as a model protein system for ^{14}N NMR method development. The observed presence of two different secondary structures (α-helix and β-sheet) found in GB3 make it potentially a highly suitable system in which to investigate ^{14}N C_Qs in proteins, since the magnitude of the C_Q is strongly dependent on the H-bonding arrangement at the ^{14}N site, which arrangement differs between α-helices and β-sheets. GB3 was sequentially assigned in the microcrystalline form it will be used in ^{14}N indirect detection experiments. The acquired spectra were well resolved, which is useful for a model sample. Furthermore, the assignment in the solid-state should be able to inform data interpretation of ^{14}N NMR experiments on the same protein, prepared in the same manner. In particular, a full ^{13}C assignment in the solid-state is important for assignment of residues in $^{13}C/^{14}N$ spectra, and knowledge of ^{15}N chemical shifts is necessary in order to be able to separate contributions from chemical shift and isotropic second order quadrupole shift in the ^{14}N shift of indirectly detected ^{14}N spectra.

CHAPTER 3

INDIRECT DETECTION OF ^{14}N VIA ^{13}C IN ORGANIC COMPOUNDS

3.1 Introduction

As outlined in Chapter 1, a number of methods have been proposed and implemented for the characterisation of the quadrupolar coupling at ^{14}N sites in organic and biological molecules. One of the most promising of these techniques for potential application to complex biomolecules containing many ^{14}N sites, such as proteins, is indirect detection of ^{14}N via spin-½ nuclei such as protons or ^{13}C. These experiments are promising in terms of application to biological solids, since they exploit the resolution and sensitivity of the 'spy' nucleus. They also do not require the uniform excitation of the entire ^{14}N spectrum. This means that they may potentially be applied to numerous ^{14}N sites in the same experiment, circumventing the resolution issues encountered when directly detecting ^{14}N, which are unavoidable since the small nitrogen chemical shift dispersion compared to the FOQI causes extensive overlap of resonances in the spectra of compounds with numerous ^{14}N sites. Indirect ^{14}N detection methods also have greater potential to be applied to complex materials than methods to detect the ^{14}N overtone transition. Overtone experiments suffer from intrinsic sensitivity issues, as the transition moment of this forbidden transition is small. Furthermore, ^{14}N overtone signals have low nutation frequencies, requiring long pulses with correspondingly low bandwidths to excite, which makes it difficult to detect multiple ^{14}N sites simultaneously.

Currently, the greatest drawbacks to indirect detection methods, which have so far precluded their application to complex biomolecules, are the poor ^{14}N transfer efficiencies and the linear increase of transfer efficiency with ^{14}N RF amplitudes. In this Chapter, we begin to address these issues and present a robust and efficient method of recording ^{14}N spectra using ^{13}C as a spy nucleus that is easy to implement and is optimally efficient with moderate (<60kHz) ^{14}N RF amplitudes. One common feature of ^{14}N-HMQC type sequences in the

literature so far is the inclusion of either a delay for excitation and reconversion of coherence either side of pulses to ^{14}N, or a dipolar recoupling sequence, such as rotary resonance, to recouple the spy nucleus/^{14}N dipolar coupling, in its place. These experiments are usually referred to as *J*-HMQC and D-HMQC experiments, respectively. In *J*-HMQC experiments, the delays allow the evolution and reconversion of the *J*-coupling and the RDS, a high order cross-term between the dipolar and quadrupolar couplings, between spy and ^{14}N nuclei[1, 2, 21, 62, 67, 124]. D-HMQC experiments use a dipolar recoupling sequence such as rotary resonance or symmetry based recoupling sequences to recouple the heteronuclear dipolar interaction between the spy nuclei and ^{14}N nuclei[63-65].

In the novel pulse sequence for correlation of ^{13}C/^{14}N spins presented in this Chapter, shown in Figure 17, delays for excitation and reconversion in the pulse sequence are absent, as are established dipolar recoupling techniques such as rotary resonance or symmetry based pulse sequences. Instead, coherence between ^{14}N and ^{13}C is generated under the direct action of long (ms), moderate amplitude (≤30-60kHz) pulses to ^{14}N. This Chapter presents the application of the ^{13}C detected pulse sequence to a number of organic samples. These samples include amino acids, both ^{13}C labelled and natural abundance, as well a simple tripeptide, which has ^{14}N quadrupolar coupling parameters similar to those found in proteins. The robustness and efficiency of the pulse sequence is evaluated, and the methods potential applicability to more complex biological systems such as proteins is assessed. Furthermore, numerical simulations of the pulse sequence have been performed for spin systems representing the samples under study in order to characterise optimal parameters and compare agreement between experimental and simulated lineshapes in order to characterise the quadrupolar coupling at ^{14}N sites. Results from the studies described in this Chapter have been published previously in Ref [4].

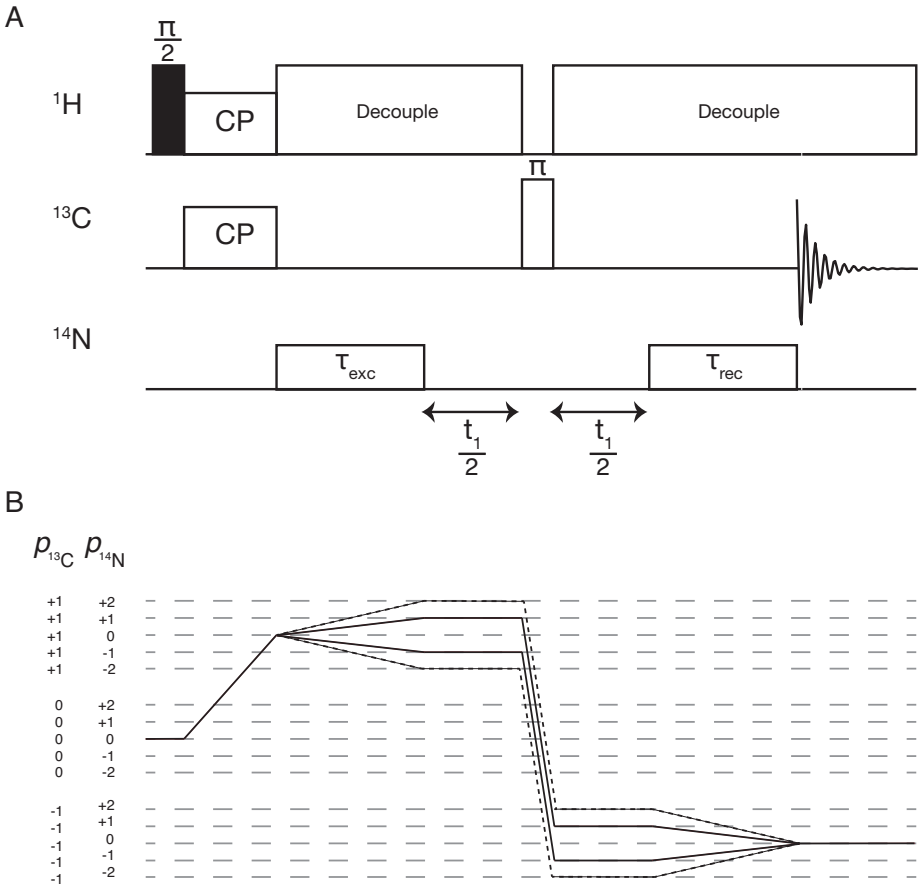

Figure 17. Pulse Sequence for Indirect Detection of ^{14}N via ^{13}C

(A) Pulse sequence, and (B) coherence transfer pathway (CTP) for $^{13}C/^{14}N$ correlation experiment. (A) Note the absence of delays preceding excitation (τ_{exc}) and following reconversion pulses (τ_{rec}) on ^{14}N; transfer of polarisation occurs directly under these pulses. These pulses are 100s µs – ms long and rotor synchronised. The π pulse to ^{13}C constitutes a spin echo, which refocuses the chemical shift of ^{13}C while signal evolves on ^{14}N during t_1. The spin echo period should be rotor synchronised, as should acquisition of t_1 increments, with $\Delta t_1=\tau_r$, to remove contributions from the first-order quadrupolar interaction. (B) Solid lines and dashed lines denote the CTPs for single or double quantum evolution, respectively, on ^{14}N during t_1. Selection of SQ coherence on ^{14}N can be achieved with a phase cycle that alternates the phase of either ^{14}N pulse as x, -x with a receiver phase of x, -x. A phase cycle of x, y, -x, -y on either ^{14}N pulse with a receiver phase of x, -x selects the ^{14}N DQ signal. Figure adapted with permission from Jarvis et. al., 2013, [4].

3.2 Indirect ^{13}C-Detected Studies of Simple Biological Molecules

The pulse sequence shown in Figure 17 has been used to record ^{13}C/^{14}N correlation spectra on a number of amino acids and a short peptide. In all experiments, SQ ^{14}N signals were selected in t_1 by phase cycling. The samples investigated were 1,2-^{13}C glycine, unlabelled L-histidine (L-His), 1,2-^{13}C alanine and gly-(^{13}C-gly)-gly (referred to as triglycine hereafter). Chemical structures of the compounds investigated in this Chapter are shown in Figure 18. Labelled samples were purchased from Cambridge Isotope Laboratories (CIL) and unlabelled samples from Sigma, and used without further purification.

Experiments were carried out on a Bruker Avance-II 600 MHz spectrometer, with a 2.5mm triple resonance MAS probe modified in house to tune to the ^{14}N Larmor frequency, 43.5MHz, on the third channel. Spectra were acquired at 25kHz MAS frequency. The spectral width was set equal to the spinning rate in t_1 to ensure folding of sidebands into the centre of the spectrum, and removal of contributions from the first order quadrupolar interaction. Magnetisation was transferred from ^{1}H to ^{13}C by cross-polarisation, linearly ramped on the proton channel to match a 50kHz spin lock field on ^{13}C. τ_{exc} and τ_{rec} pulse lengths were 2ms. On all compounds this value was found to give optimal efficiency, with RF amplitudes optimised experimentally to 35kHz for amino acids and 50kHz for triglycine. SPINAL-64[114] decoupling was employed on the proton channel at an RF field of 115kHz during ^{14}N pulses, where efficient decoupling is vital to achieve good transfer efficiencies, and at 100kHz during t_1 and t_2 periods. The States method was used to acquire complex points in the indirect dimension[125]. This, in combination with stroboscopic acquisition, is essential in order that all folded sidebands have the same phase and can be co-added. Points were recorded until the signal decayed to noise, requiring no more than 64 increments on any of the samples studied here. Information on the number of points collected for each experiment is given in figure legends. Data were processed in MatNMR[126], where 30Hz linebroadening was applied in both dimensions, before zero-filling to 1024 data points in each dimension and 2D Fourier transform. All spectra are referenced to solid NH_4Cl at 39.3ppm in the ^{14}N dimension, corresponding to liquid ammonia

at 0ppm[127]. The ^{13}C dimensions are referenced to 0.5% DSS in $CDCl_4$, using the the upfield peak in the adamantane ^{13}C spectrum at 40.48ppm as a secondary reference[117].

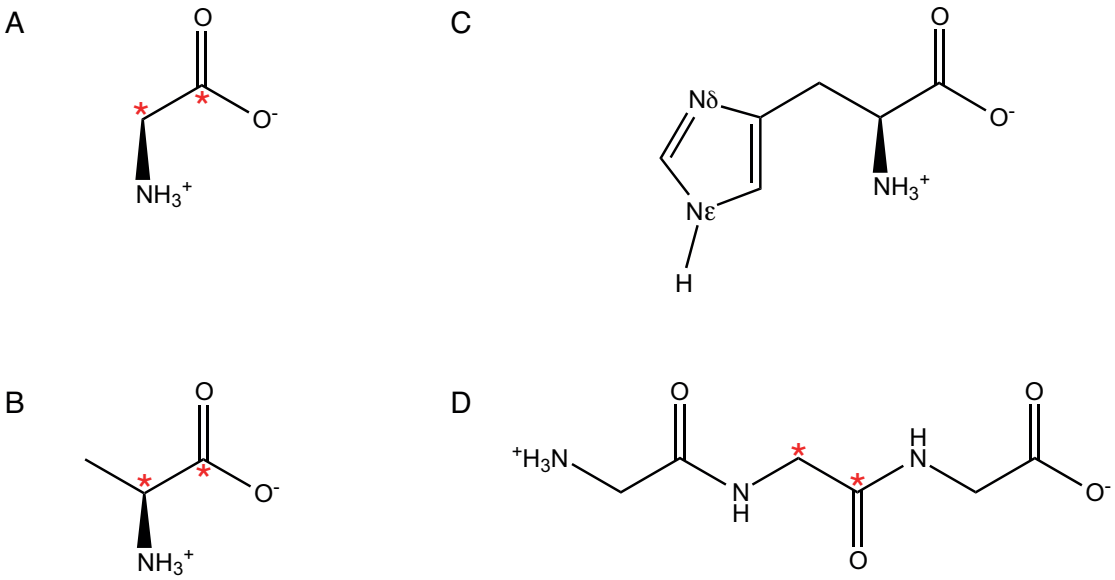

Figure 18. Chemical Structures of Compounds Used in $^{13}C/^{14}N$ Correlation Experiments

(A) 1,2-^{13}C glycine (B) 1,2-^{13}C alanine (C) L-histidine (D) gly-(^{13}C-gly)-gly. Sites labelled with ^{13}C are marked with red asterisks.

3.4 Numerical Simulations of ^{13}C Detected ^{14}N Indirect Detection Experiments

To characterise the effects of experimental parameters on the spectra acquired from the indirect detection experiment, numerical simulations of the experiment were designed and run using SPINEVOLUTION [128]. A simulation of the pulse sequence (Figure 17) to generate the 2D spectrum was written, and its accuracy verified by comparison with experimental spectra. The simulations approximated experimental spin systems with two spins, a dipolar coupled ^{13}C and ^{14}N pair, in order to maximise time efficiency. Protons, and therefore CP and decoupling were omitted to save time. Simulations were conducted for the C-N amine site in glycine and alanine, and C-N amide in triglycine. The magnitude, anisotropy and relative tensor orientations of quadrupolar and dipolar couplings were adjusted to represent the different systems. Values for the magnitudes and orientations of the interactions used in the simulations, which were taken from the literature, are shown in Table 5. Simulations were performed in the interaction frame of the $^{13}C/^{14}N$ dipolar coupling. Relative orientations of the dipolar and quadrupolar tensors in the case of the amine bond in single amino acids were assumed to be coincident (i.e. $\{\alpha,\beta,\gamma\}=\{0,0,0\}$) due to the high symmetry and mobility of the NH_3^+ group, which is sp3 hybridised at this ^{14}N site. ^{13}C chemical shift anisotropy (CSA) was also omitted, since the primary amine Cα site has a low CSA (<1ppm)[23], and it was found not to contribute to any observed spectral features. The nitrogen CSA at was also omitted as it is typically small at amine sites (~10ppm)[23, 129] and did not contribute to the simulated spectra. For the triglycine amide bond, an average value for the quadrupolar tensor's relative orientation with respect to the $^{13}C/^{14}N$ dipolar tensor in peptides, of $\{\alpha,\beta,\gamma\}=\{0,90,120\}$, was taken from Ref [130]. A ^{13}C CSA=11.47kHz, $\eta_{CSA}=0.98$, with and orientation of $\{\alpha,\beta,\gamma\}=\{0,90,120\}$ was used to simulate the ^{13}C carbonyl site in triglycine[130]. Parameters such as frequency offsets, RF field amplitudes and spinning speed were taken from the NMR experiments being simulated. *J*-couplings between the two spins were omitted. A 3D zcw angle set of 6044 orientations was used for all 2D simulations, having

verified spectra converged with this angle set[131]. Having simulated the experimentally observed spectrum, subsequent simulations were run in order to gauge the effects of experimental parameters on spectral features, and determine the robustness and efficiency of the pulse sequence when applied to various samples with different characteristics representative of ^{14}N sites in biological molecules. Specifically, simulations were designed to investigate optimal values for ^{14}N pulse length and RF power, respectively, in systems with different ^{14}N quadrupolar couplings.

Table 5. Interaction Parameters Used in HCN Simulations

Simulation	C_Q *(MHz)*	η	*Bond Length (Å)*	*Dipolar Coupling (kHz)*	*Ref.*
Alanine (Amine)	1.18	0.54	1.45	-0.716	[132]
Glycine (Amine)	1.25	0.46	1.45	-0.716	[133]
Triglycine (Amide)	3.01	0.48	1.33	-0.928	[133]

3.5 Results and Discussion

In order to assess experimental efficiency, preliminary ^{14}N indirect detection experiments were conducted with the pulse sequence shown in Figure 17. The samples were chosen such that a representative range of quadrupolar coupling, CSA and dipolar coupling magnitudes could be investigated to evaluate the efficacy of the pulse sequence across a range of spectral properties.

3.5.1 2D $^{13}C/^{14}N$-Correlation Experiments: 1,2-^{13}C-Glycine and 1,2-^{13}C-Alanine

Figure 19A and 9B show well-resolved $^{13}C/^{14}N$ correlation spectra of amino acids 1,2-^{13}C-glycine and 1,2-^{13}C-alanine, respectively. In both of these molecules, there is only one ^{14}N atom, which is found in an NH_3^+ group bonded to the α-carbon. In both cases, the sp3-hybridisation state of the ^{14}N in the amine bonded to the α-carbon has a relatively high symmetry, leading to a relatively small perturbation of the EFG, and thus modest C_Q values of 1.25MHz and 1.18MHz for glycine and alanine, respectively, as reported[132, 133]. A single peak is observed in the glycine spectrum at 113ppm in the ^{14}N dimension, corresponding to a correlation between the amine group and the α-carbon of glycine. An analogous single peak is observed in the alanine spectrum at 106ppm in the ^{14}N dimension. The contribution to the ^{14}N shift of the ^{14}N δ_{CS}^{iso} and the ^{14}N δ_{Q}^{iso} can be simply separated from each other in these spectra by subtraction of the ^{15}N δ_{CS}^{iso} (which can be assumed to be the same as the ^{14}N δ_{CS}^{iso}) from the ^{14}N shift, yielding the ^{14}N δ_{Q}^{iso}. The ^{15}N δ_{CS}^{iso} of glycine, measured in a separate ^{15}N CPMAS experiment, was determined to be 41.2ppm, which subtracted from the observed ^{14}N shift of 108ppm, gives a ^{14}N δ_{Q}^{iso}=66.8ppm. This is in excellent agreement with the expected value of the ^{14}N δ_{Q}^{iso} at this site of 66.6ppm, calculated from Equation (1.49) using literature values of C_Q=1.25MHz, η=0.46. Similarly, the ^{14}N δ_{Q}^{iso} of alanine is determined to be 63.5ppm, by subtraction of the ^{15}N δ_{CS}^{iso}=42.5ppm (measured by CPMAS) from the ^{14}N shift of 106ppm. This value is consistent with the expected value of the

^{14}N δ_Q^{iso}=60.8ppm, calculated from Equation (1.49) using literature values of C_Q=1.18MHz, η=0.54. This demonstrates that the position of indirectly detected ^{14}N peaks may be used to simply and accurately characterise the ^{14}N δ_Q^{iso}, however as is apparent from Equation (1.50), the ^{14}N C_Q and η are correlated and both contribute to the ^{14}N δ_Q^{iso}. The consequence of this is that the values of C_Q and η may only be constrained to a certain range of values that are consistent with ^{14}N δ_Q^{iso}. The range of values that these parameters can be constrained to is governed by the size of the magnetic field, and the magnitude of the quadrupolar interaction at a site. For more precise determination of the quadrupolar interaction parameters C_Q and η, the lineshape, which also depends on C_Q, η and the relative orientations of quadrupolar and dipolar tensors, would need to be fit to simulations. The accuracy with which these parameters can be defined from the ^{14}N shift, and the ^{14}N lineshape of indirectly detected ^{14}N spectra is discussed further in Chapter 4.

A good agreement is observed between numerical simulation and experiment for both alanine and glycine, as shown by projections through the ^{14}N-dimension for both experiments and numerical simulations in Figure 19. The numerical simulations describe a quadrupolar tensor of 1.25MHz, η=0.54 for glycine, and 1.18MHz, η=0.46 for alanine, collinear with the dipolar tensor in both cases. The good agreement confirms that the experiment is able to quantitatively characterise relatively small quadrupolar couplings at ^{14}N sites, and demonstrates that a simulation fitting approach to determining quadrupolar coupling parameters more precisely is potentially viable.

The experimental efficiency of the ^{14}N detection pulse sequence as applied to 1,2-^{13}C-glycine was analysed; the results are shown in Figure 20. Here, the intensity of the observed peaks in the first slice of the ^{13}C dimension from the ^{13}C/^{14}N correlation experiment were compared with the intensity of corresponding peaks from 1D cross-polarisation (CP) and spin-echo experiments acquired under the same conditions. In the spin echo experiment, echo times equal to the length of ^{14}N pulses applied in the ^{14}N filtered experiment were applied. The intensity of the Cα resonance in the ^{14}N filtered experiment is 10% with respect to CP, and 17.5% with respect to the spin-echo

experiment. The spin echo experiment has reduced intensity compared to the CP experiment due to the inherent loss of ^{13}C signal during a spin-echo primarily due to T_2'-decoherence due to residual dipolar couplings, as well as other effects such as evolution of homonuclear ($^{13}C/^{13}C$) *J*-couplings. Hence, the comparison of ^{14}N filtering efficiency with respect to the spin echo is a more relevant measure of performance, taking into account site-specific relaxation effects that affect both experiments. This is the measure used to express ^{14}N transfer efficiency throughout this thesis. The efficiency of the ^{14}N filtered experiment applied to glycine is typical of that observed performing this experiment at amine sites in amino acids enriched with ^{13}C at multiple sites, and is sufficient to acquire good quality spectra in a few hours. This compares favourably with ^{14}N-HMQC experiments in the literature, where maximal efficiencies at primary amine sites in amino acids of 15% with respect to spin echo have been reported[21].

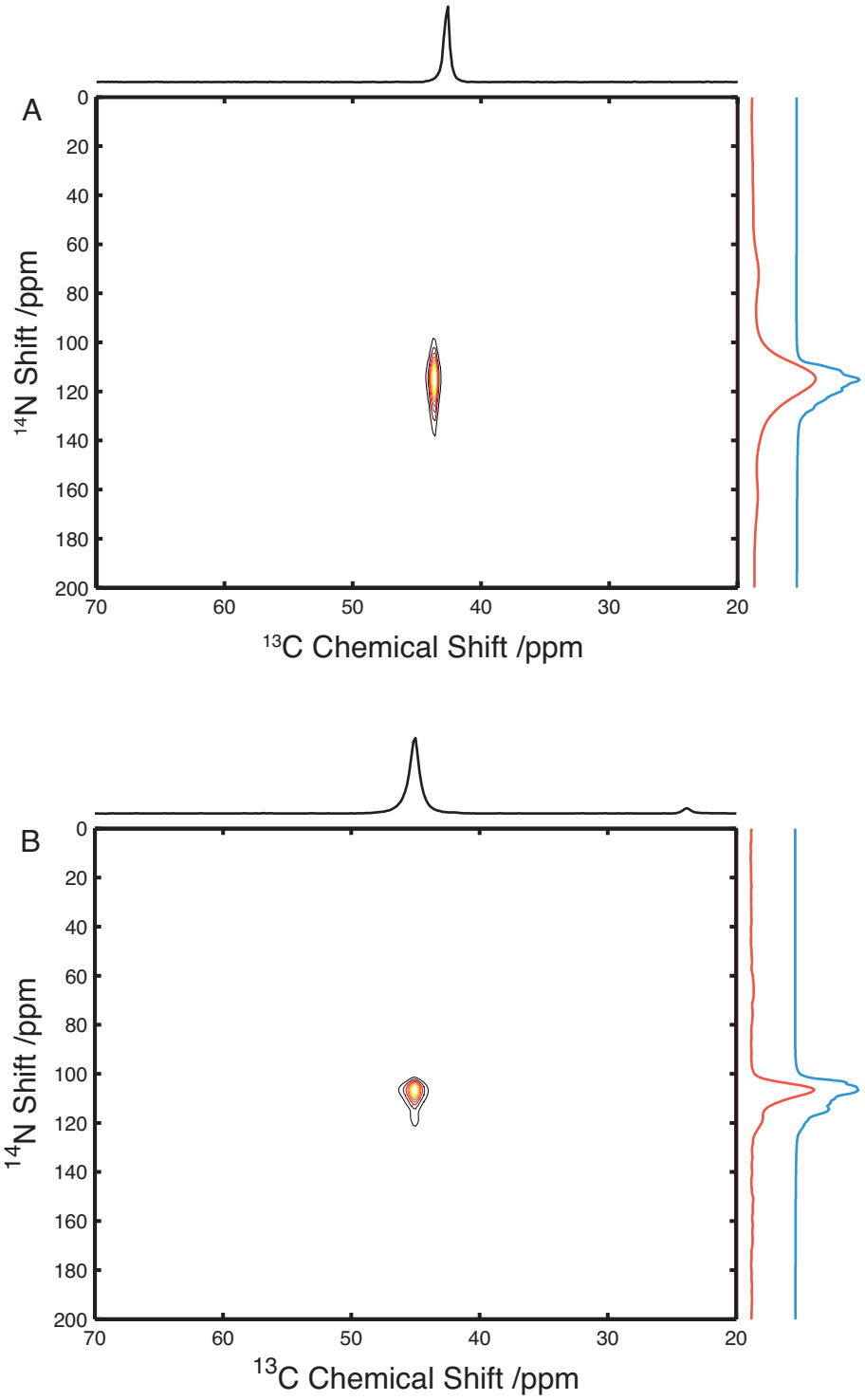

Figure 19. 2D ^{13}C/^{14}N Correlation Spectra of 1,2-^{13}C-Glycine and 1,2-^{13}C-Alanine

(A)1,2-^{13}C-Glycine. (B) 1,2-^{13}C-Alanine. Both spectra acquired with 16x64 scans, 2.5s recycle delay. Experimental (red) and simulated (blue) projections of the ^{14}N dimension are shown for both spectra. Figure adapted with permission from Jarvis et. al., 2013, [4].

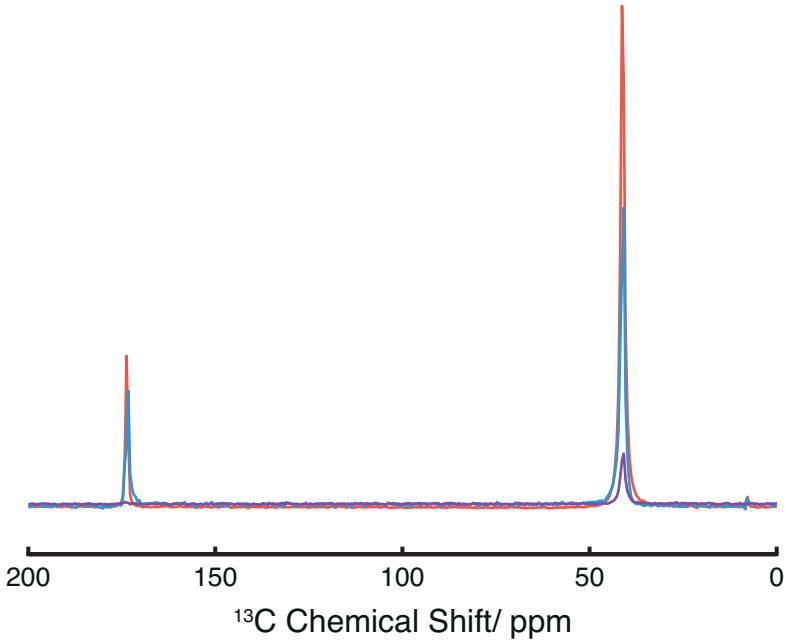

Figure 20. Efficiency of ^{14}N Transfer in ^{13}C Detected ^{13}C/^{14}N Correlation Spectra in 1,2-^{13}C-Glycine

^{13}C CPMAS (red), ^{13}C CP spin echo with τ_{echo}=2ms either side of refocusing π pulse (blue) and ^{14}N filtered ^{13}C spectrum with τ_{exc} and τ_{rec} ^{14}N pulses applied for 2ms at 35kHz (purple) spectra. Efficiency of the ^{14}N filtered experiment at the Cα site (43ppm) is 9% with respect to CP, and 17.5% with respect to spin echo. Figure adapted with permission from Jarvis et. al., 2013, [4].

3.5.2 2D $^{13}C/^{14}N$-Correlation Experiments: Histidine and Gly-(1,2-^{13}C-Gly)-Gly

To assess how the pulse sequence's performance was affected applying it to compounds with a variety of ^{14}N quadrupolar coupling magnitudes and asymmetries, compounds with a variety of different amine, amide and heterocyclic nitrogen sites were investigated.

3.5.2.1 Histidine

Unlabelled histidine was chosen as a model sample to assess the applicability of the indirect detection experiment to materials without ^{13}C enrichment, as well as investigating the experiment's ability to characterise quadrupolar couplings at ^{14}N sites that differ from the primary amine sites found in glycine and alanine. The $^{13}C/^{14}N$ correlation spectrum of histidine is shown in Figure 21. The amine site in L-His is similar to equivalent sites found in glycine and alanine. The ^{14}N C_Q at this site has been determined by NQR to be 1.25MHz, with η=0.11[134]. The two nitrogens in the imidazole ring on the sidechain of L-His have reduced symmetry and mobility compared to the primary amine site, giving rise to quadrupolar couplings of 3.36MHz, η=0.31 (Nδ) and 1.44MHz, η=0.92 (Nε), neither of which are coincident with the $^{13}C/^{14}N$ dipolar tensor[134]. The particularly asymmetric distribution of electron density at the Nδ site, which forms double and single bonds to adjacent carbons, results in a large (>3MHz) quadrupolar coupling at this site. At the Nε site, the EFG is alleviated somewhat by two adjacent single bonds formed with carbons, and the formation of a H-bond. It is desirable for the experiment to be able to detect the various sites with different ^{14}N EFG tensor parameters in L-His simultaneously if it is to be generally applicable to the study of small organic molecules, and potentially applied to complex samples such as peptides and proteins. In the spectrum shown in Figure 21A, three resonances are observed at ^{14}N shifts of 97ppm, 420ppm and 445ppm, which are assigned to the NH_3^+, Nε and Nδ sites, respectively. This confirms that the indirect detection sequence is able to simultaneously excite ^{14}N sites with a range of quadrupolar parameters,

including those with large C_Q magnitudes, despite the ^{14}N pulses applied not being of a sufficient bandwidth to uniformly excite the entire ^{14}N spectrum. This result also demonstrates that this $^{13}C/^{14}N$ correlation experiment is sensitive enough to permit the analysis of entirely unlabelled materials via natural abundance ^{13}C, with the spectrum in Figure 21A taking 20h to acquire on ~12mg sample. In fact, the ^{14}N transfer efficiency with respect to spin echo of the Cα in unlabelled histidine was 22%, compared to 16% at the Cα site in U-^{13}C-glycine. This increased efficiency we attribute to the lack of a $^{13}C/^{13}C$ *J*-coupling in the unlabelled compound; meaning signal is not attenuated by evolution of the *J*-coupling during the ^{13}C spin echo period. The experimental time for acquisition of $^{13}C/^{14}N$ correlation spectra on unlabelled histidine compares favourably with ^{15}N CPMAS experiments on unlabelled small organic molecules, where 2D correlation methods at natural abundance would generally be prohibitively time consuming. Recently, the first example of a natural abundance $^{13}C/^{15}N$ correlation experiment on an organic molecule (a deoxyguanosine derivative) was demonstrated under DNP at 100K, which required 25 hours of signal averaging with a DNP enhancement factor ($\varepsilon = S_{on} / S_{off}$) of 11, translating into a timesaving factor of >100 compared to performing the experiment at room temperature without DNP[135]. The $^{13}C/^{14}N$ correlation method presented here could potentially provide a viable alternative route to the investigation of nitrogen sites via ^{13}C in naturally abundant samples, or those which cannot be ^{15}N enriched, without requiring specialised experimental equipment. This method would have the additional advantage, over acquisition of ^{15}N spectra, of containing information on the ^{14}N EFG tensor at nitrogen sites.

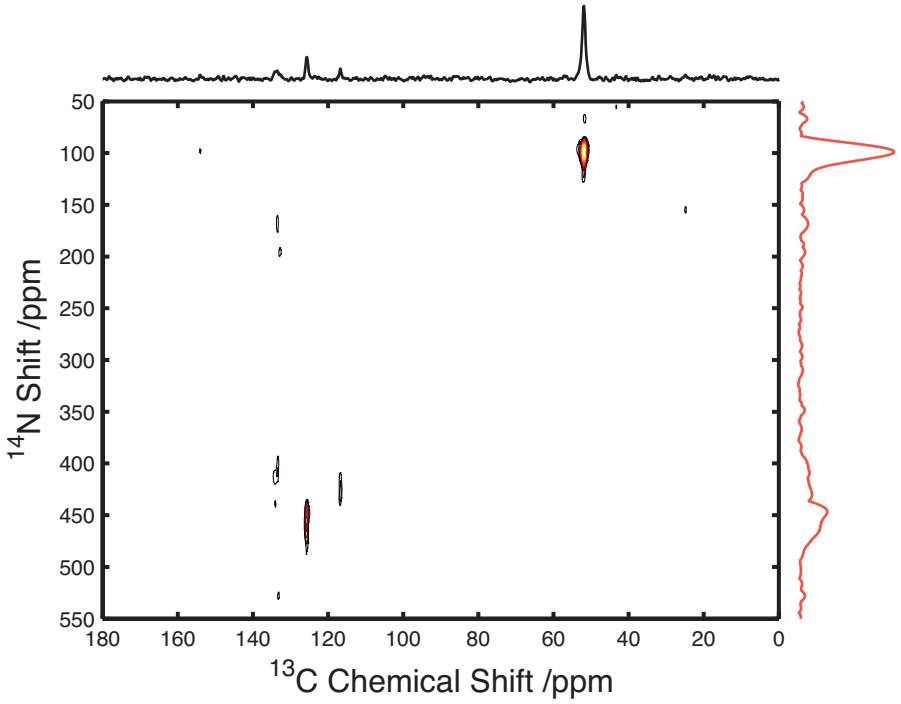

Figure 21. 2D ^{13}C/^{14}N Correlation Spectrum of Natural Abundance Histidine

Acquired with 192x32 scans. A projection of the ^{14}N dimension is plotted in red. Figure adapted with permission from Jarvis et. al., 2013, [4].

3.5.2.2 Triglycine

Partially labelled triglycine was chosen as a model sample since the quadrupolar coupling at the N-2 and N-3 sites in triglycine is several MHz larger than those found in free amino acids, and are more representative of those found in peptides and proteins due to their structural similarity. In the peptide bond, both the nitrogen and carbon are found in an sp^2-hybridised state due to a resonance that allows the lone pair of electrons on the nitrogen to form a partial double bond with the carbonyl carbon, with a pair of electrons from the C=O double bond being donated to the oxygen atom. The sp^2-hybridisation at the nitrogen site gives a trigonal planar bond geometry, which has a less symmetrical distribution of electron density about the nitrogen than in the sp^3-hybridised case. This results in a greater quadrupolar coupling owing to greater perturbation of the EFG at these sites. The reported C_Q at the N-3 site in triglycine is 3.01MHz, η=0.48[133]. The orientation of the quadrupolar tensor, unlike in the amine case, is not aligned with the N-Cα bond, the *z*-axis of the C_Q tensor lies perpendicular to the peptide plane, with the *x*- and *y*-axes of the C_Q tensor in the peptide plane. [130, 133]. The spectrum of triglycine (Figure 22) has peaks arising from correlations between the two labelled ^{13}C nuclei in residue *i*=2 of gly-(1,2-^{13}C-gly)-gly to adjacent ^{14}N nuclei. The labelled carbonyl ^{13}C is correlated with and the ^{14}N of the *i*=3 glycine residue (N-3), and the labelled α-carbon in residue *i*=2 is correlated with the adjacent nitrogen in the same residue (N-2). The peak corresponding to the amide N-3 is centred at 440ppm in the ^{14}N dimension with a breadth of ~250ppm. An additional resonance is observed at 125ppm with a breadth of ~80ppm. The peak arising from the amide bond of N-2 is found at 410ppm with intensity over ~190ppm. These peaks are significantly broadened in the ^{14}N dimension, compared to the spectra of free amino acids. While it is expected that second order quadrupolar broadening arising from fourth rank terms in the quadrupolar Hamiltonian that are not averaged, only scaled by MAS, should be larger in samples with increased C_Q magnitudes, it was noted while working on the simulations that spectral features were present in the simulations that did not appear in the experimental spectra. It was found that when detecting the SQ ^{14}N spectrum the

pulse sequence is extremely sensitive to deviations from the magic angle, and the poor agreement and excessive linebroadening observed here are attributable to the magic angle being mis-set, which is particularly pronounced for materials with large quadrupolar couplings. Since the contribution of the large (MHz) FOQI to the to the ^{14}N lineshape is removed in the indirect dimension by MAS and rotor synchronisation of t_1 increments, this is expected, and it has been previously observed that indirect detection experiments are extremely sensitive to the setting of the magic angle, and the stability of spinning [1, 64]. This also explains why the effect is more pronounced for ^{14}N sites with larger quadrupolar couplings, and therefore larger first order quadrupolar interactions. To investigate further the severity of this effect at a ^{14}N site with a large quadrupolar coupling, numerical simulations off the magic angle were conducted on a spin system representing the N-3 site of tricylcine at 14.1T, 25kHz MAS, in order to ascertain the limits by which the magic angle is required to be set experimentally. Figure 23 demonstrates the effect of being off magic angle on the ^{14}N dimension of a simulated triglycine spectrum. Deviations from the magic angle of ±0.05° caused changes in the simulated lineshape with features shifting by 10s ppm. From comparison of these simulations to the experimental lineshape, it was deduced that the angle experimental spectra were acquired at deviated from the magic angle by -0.025°, and highlighted the importance of very accurately setting the magic angle of the probe when conducting future measurements. The comparison of the experimental slice through the carbonyl resonance with that of a simulated spectrum shown in Figure 22 shows the simulation that was found to agree best with the experiment, with a deviation from the magic angle of -0.05°. A more accurate routine for doing so than maximising the length of rotor echoes in the ^{79}Br FID of KBr should be employed in future. By maximising the intensity of a 1D ^{87}Rb STMAS experiment on $RbNO_3$, one can set the magic angle to within ±0.004°[136, 137]. Despite the off-angle effects in the spectra, the agreement between numerical simulation and experiment eventually achieved shows that, as for the case with the low C_Q value of glycine, the ^{14}N lineshapes are of

sufficient quality to characterise the quadrupolar coupling at high-C_Q ^{14}N sites found in peptides and proteins.

The position of the peaks in the ^{14}N dimension is reasonably consistent with that expected for amide ^{14}N site at this field, however the broadening introduced into the ^{14}N lineshape from the mis-set of the magic angle introduced a large error into the experimental ^{14}N δ_Q^{iso} value. A typical ^{15}N δ_{CS}^{iso} for a glycine residue in a peptide is ~120ppm[130], which when subtracted from ^{14}N shift at the centre of gravity of the ^{14}N lineshape of the N-3 site of 440ppm gives a measured ^{14}N δ_Q^{iso}=320ppm at this site. This is somewhat lower than the expected ^{14}N δ_Q^{iso} value, as calculated from the literature values of 3.01MHz, η=0.48 using Equation (1.49), of 388ppm. This inaccuracy may be attributable to error in the estimation of the ^{15}N δ_{CS}^{iso}, rather than measuring it in a ^{15}N CPMAS experiment, as well as the increased error associated with measuring the centre of gravity of this significantly broadened peak from the ^{14}N spectrum. However it is expected from this calculation, which includes a large error margin, as well as the results on materials with similar C_Q magnitudes presented in Chapter 4, that the ^{14}N indirect detection method can be used to characterise the large ^{14}N quadrupolar couplings that are expected in peptides and proteins, as well as the small couplings found at amine sites. It also shows that large ^{14}N δ_Q^{iso} shifts (>300ppm) are expected between ^{14}N sites with small (~1MHz) and larger (~3MHz) quadrupole couplings at this field. Typically, far smaller differences of ~70ppm in the ^{15}N chemical shift would be expected from primary amine and amide sites in amino acids and peptides, demonstrating the increased sensitivity of ^{14}N spectra to molecular structure over ^{15}N spectra.

Efficiency of the ^{14}N transfer was analysed for triglycine in the same manner as for glycine, in order to compare the transfer efficiency in ^{13}C labelled compounds at ^{14}N sites with different magnitudes of C_Q. Spectra of the ^{14}N filtered ^{13}C spectrum, ^{13}C CP spectrum and spin echo spectrum of triglycine are shown in Figure 24. In this case, intensities of the signals in trigylcine were 7% and 9% of CP and spin-echo experiments, respectively, at the carbonyl (N-3) site, and 4% and 4.5% of CP and spin-echo experiments at the Cα (N-2) site. While lower than the efficiencies observed for glycine, which is attributed to the

increased C_Q at this amide site, the efficiency of the experiment as applied to triglycine is comparable to that observed by other HMQC experiments reported in the literature [2, 124]. This is sufficient to collect spectra of materials with larger quadrupolar couplings within around half a day of experimental time. However, it is possible to increase the efficiency of this experiment applied to labelled samples by suppression of homonuclear *J*-couplings which contribute to loss of signal in U-^{13}C labelled samples, since they are not refocused by the spin echo on the ^{13}C channel. One can increase the efficiency using a selective spin echo applied to either the carbonyl or Cα site during ^{14}N evolution in order to refocus the ^{13}C/^{13}C *J*-coupling during the spin echo (data not shown). This makes the experiment selective for whichever spin is pulsed and refocused, though, meaning overall sensitivity is reduced if all sites are of interest. This finding is supported by the transfer efficiencies of unlabelled histidine, where a low abundance of ^{13}C resonances results in an effective absence of ^{13}C/^{13}C *J*-coupling. The transfer efficiencies of unlabelled histidine are found to be greater than those in both 1,2-^{13}C-glycine and triglycine, with a ^{14}N transfer efficiency of 22% observed at the Cα site in histidine, compared to 17% at the same site in U-^{13}C glycine. It is also apparent from this finding that application to sites with lower ^{13}C T_2' decoherence rates will yield inherently more sensitive spectra, as less signal will be lost during the spin echo. For the same reason, faster MAS rates would also benefit the performance of this experiment by increasing the ^{13}C coherence lifetime.

Collectively, the spectra presented demonstrate that the presented pulse sequence can accurately examine both the relatively small (<1.5MHz) quadrupolar couplings present in free amino acids, as well as those found at heterocyclic nitrogen sites, and the large couplings (>3MHz) that are usually found in peptides and proteins without any more than routine optimisation of experimental parameters (such as RF amplitude and offset on recoupling pulses) on the spectrometer. The sequence also has the required sensitivity to detect ^{13}C/^{14}N correlations in unlabelled materials, as well as those enriched with ^{13}C. Projections of the ^{14}N dimension from experimental spectra generally agree well with those obtained from numerical simulation of the pulse sequence, for all spin

systems investigated, as shown in Figure 19 and 21. The observed good agreement between experiment and simulation means that quadrupolar couplings at ^{14}N sites may be characterised precisely, through the fitting of simulated lineshapes to those which are determined experimentally. It also allowed additional simulations investigating the effect of various experimental parameters on the observed spectra to be performed.

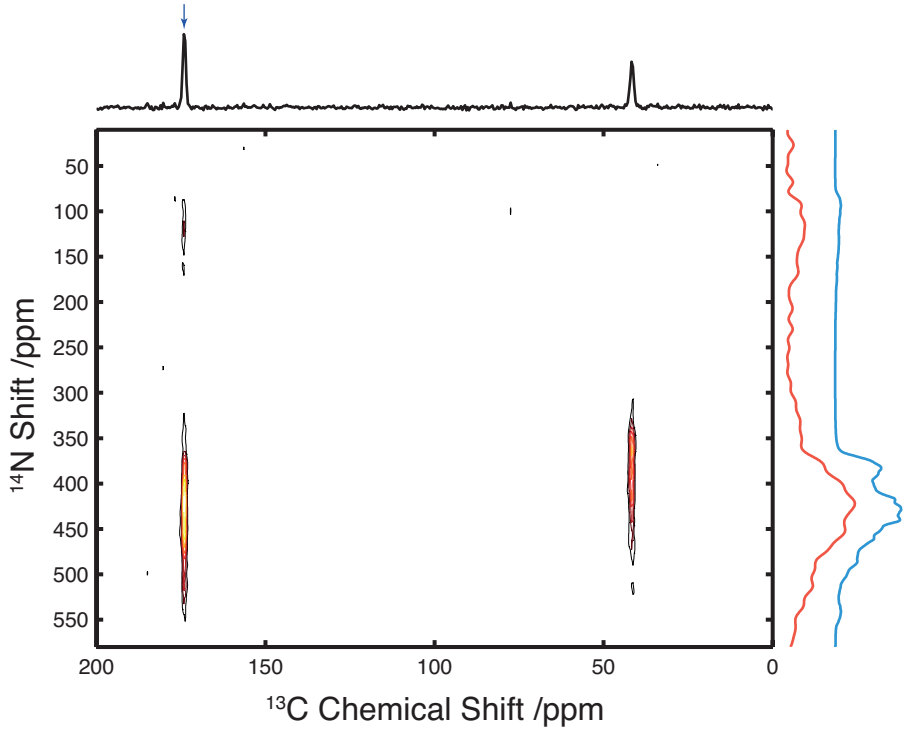

Figure 22. 2D $^{13}C/^{14}N$ Correlation Spectrum of gly-(^{13}C-gly)-gly

Acquired with 512x32 scans. Experimental (red) and simulated (blue) slices through the ^{14}N-axis at the amine resonance indicated by the blue arrow are plotted. The simulation was performed -0.05° off the magic angle, and fit best to the experimental data. Figure adapted with permission from Jarvis et. al., 2013, [4].

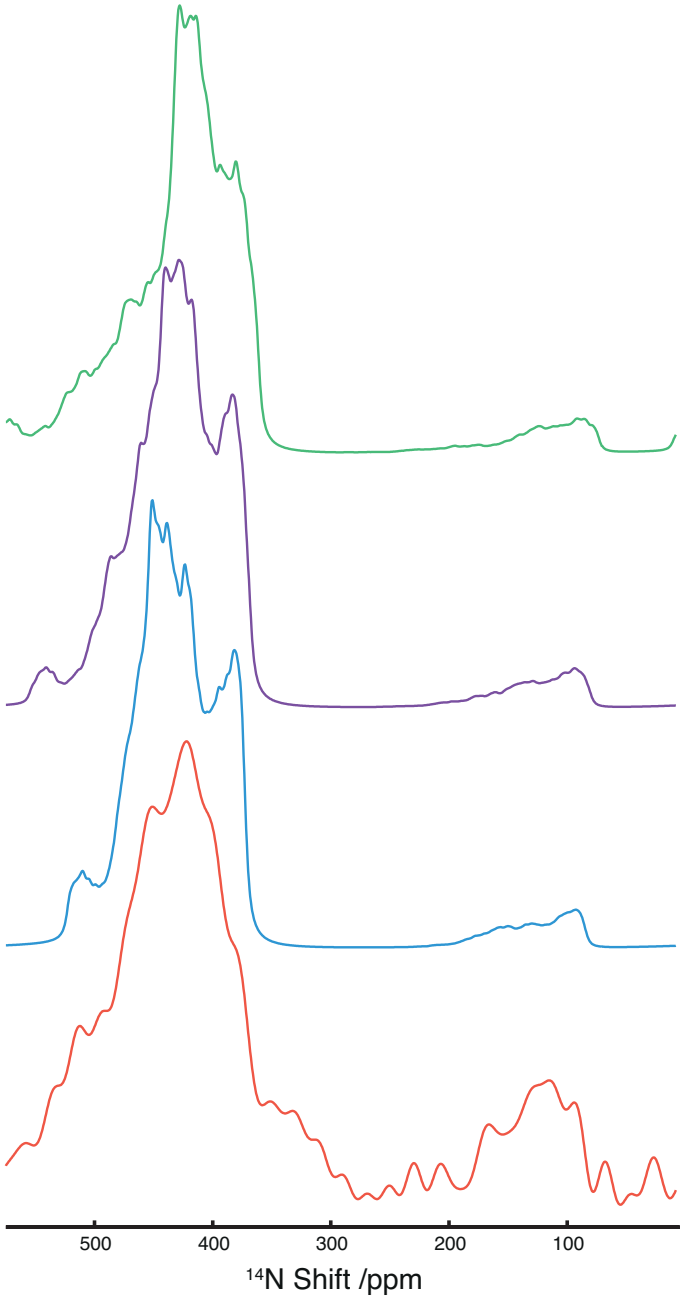

Figure 23. Effect of Deviation From the Magic Angle on Simulated ^{14}N Lineshape for gly-(2,3-^{13}C-gly)-gly

The slice through the ^{14}N-axis at the amide resonance of the experimental 2D HNC filtered spectrum of gly-(^{13}C-gly)-gly is plotted in red. A simulated slice at the Magic angle (54.7356°) is plotted in blue. Purple and green plots represent simulated slices at -0.025° and -0.05° deviation the magic angle, respectively.

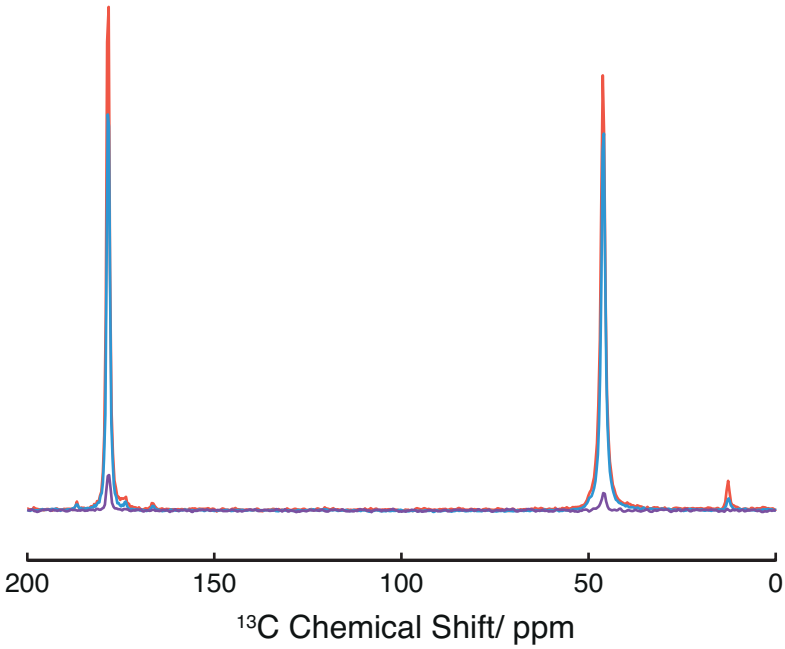

Figure 24. Efficiency of ^{14}N Transfer in ^{13}C Detected ^{13}C/^{14}N Correlation Spectra in Triglycine

^{13}C CPMAS (red), ^{13}C CP spin echo with τ_{echo}=2ms either side of refocusing π pulse (blue) and ^{14}N filtered ^{13}C spectrum with τ_{exc} and τ_{rec} ^{14}N pulses applied for 2ms at 50kHz (purple) spectra. Efficiency of the ^{14}N filtered experiment at the CO site (178ppm) is 7% with respect to CP, and 9% with respect to spin echo. Efficiency of the ^{14}N filtered experiment at the Cα site (45ppm) is 4% with respect to CP, and 4.5% with respect to spin echo. Figure adapted with permission from Jarvis et. al., 2013, [4].

3.5.3 Effect of ^{14}N Pulse Lengths and RF Field Amplitude

In order to determine the sensitivity of the experimental parameters to the characteristics of different nitrogen environments, and determine optimal parameters for performing ^{13}C/^{14}N correlation experiments on different systems, numerical simulations have been performed for spin systems representative of amino acids and small peptides. The simulated signal intensity was measured as a function of both the pulse length of applied ^{14}N RF fields, and their RF amplitude. The amplitude of the RF field was varied between 0-100kHz and the ^{14}N pulse length between 0-4ms, reflecting parameters attainable with the 3.2mm and 2.5mm MAS probes available in our laboratory. The results are displayed as contour plots in Figure 25A and 12B. In Figure 25A, the simulated spin system is representative of the primary amine site in glycine, where ^{14}N C_Q=1.25MHz, η=0.46. The ^{13}C CSA was omitted, and the dipolar coupling= -0.716kHz, representing a 1.45Å bond length. Figure 25B was performed for a spin system representative of the N-3 amide site in the peptide backbone of triglycine, which has ^{14}N EFG tensor parameters representative of those found in peptides generally[138]. The simulation parameters were: ^{14}N C_Q=3.01MHz, η=0.46, ^{13}C CSA=11.47kHz and the dipolar coupling= -0.928kHz, for a 1.33 Å bond length. In both cases, relaxation was neglected in the simulations and the simulations were carried out with a B_0 field of 14.1T (600MHz for ^{1}H) and 25kHz MAS. It is important that the experiment is robust enough to be applicable to ^{14}N sites with various spectral properties, in terms of quadrupolar coupling magnitude and orientation, CSA and dipolar coupling magnitude, simultaneously and without optimisation of experimental parameters within extremely fine limits in order for it to be relatively simply applied to complex systems which may contain multiple nitrogen sites in different environments with different ^{14}N EFG tensor parameters.

As shown by Figure 25A, for a system with a moderate quadrupolar coupling (C_Q=1.25MHz, η=0.46), optimal signal intensities are observed for excitation and recoupling pulses in excess of 1.5ms with RF amplitudes of between 30 and 70kHz. There is a broad condition to be met in order to achieve efficient transfer of polarisation in terms of both ^{14}N pulse length and RF power,

meaning the experiment should be robust with respect to these parameters when applied to materials with moderate quadrupolar interactions. Experimentally, optimal transfer efficiencies were not observed >2ms, despite optimal transfer efficiencies being simulated at ~3ms. This we attribute to the simulations not accounting for dipolar dephasing or relaxation of either ^{13}C or ^{14}N, with T_2' of the former and $T_{1\rho}$ of the latter being critical for achieving optimal efficiencies. However, ~90 – 95% of the simulated maximum may be achieved with 2ms pulses at 35kHz, which is what parameters were optimised to experimentally. Experimental build up curves of the ^{14}N filtered experiment as a function of both ^{14}N RF amplitude and pulse width, acquired on 1,2-^{13}C-glycine, are shown in Figure 26. These demonstrate that the efficiency of ^{14}N transfer deteriorates at ^{14}N pulse widths greater than ~2ms under the experimental conditions used here. The observed optimal amplitude of the ^{14}N pulses of ~35kHz is in agreement with the simulations, however. The required RF amplitudes for maximal transfer efficiency are readily attainable with our available RF hardware and most commercially available probes of rotor sizes up to 3.2mm, and it is an advantageous feature that maximum observed efficiency does not increase with ^{14}N power, meaning that the pulse sequence can be applied with a range of hardware under conditions where high (>60kHz) RF amplitudes cannot be achieved. This is contrast to *J*-based HMQC experiments, where ^{14}N transfer efficiency tends to increase linearly with the ^{14}N RF field[75].

For the peptide spin system simulated in Figure 25B, a maximum intensity is observed at pulse lengths above 3.5ms, with RF power between 75 and 100kHz. However, as observed for the amine simulations, relaxation effects were not taken into account in these simulations, and experimental optimisation indicates that at pulse lengths >2ms, the signal intensity begins to decline. This suggests that at this field and spinning frequency, this maximum cannot be realised experimentally due to relaxation. Another broad region of high of signal intensity (~70% of maximum) is observed at pulse lengths of between 1.5ms-3ms, with RF power of between 40kHz and 70kHz. These experimental parameters are far more achievable; these field strengths are readily attainable with standard probes up to 3.2mm, and the pulse length is suitable to generate

an efficient transfer of polarisation without losing excessive amounts of signal due relaxation. The high intensity at this region on the simulated contour plot is also in line with the observed maximum experimental efficiencies for triglycine, where pulses of 2ms at 50kHz were employed. The fact that this maximum is broad, encompassing 1.5ms and 20kHz, means that the experiment is robust with respect to ^{14}N pulse width and field strength at larger quadrupolar couplings as well as smaller ones; it is well compensated for small changes in pulse length and RF amplitude in both cases examined, and excessively precise optimisation to achieve signal is not necessary.

The maxima observed in the two plots, as well as both being relatively broad, both occur at similar pulse lengths and RF amplitudes, with only slight increases in RF amplitude optimal for higher C_Q systems. Practically, this makes the pulse sequence relatively insensitive to the magnitude of the quadrupolar interaction in terms of generating coherence with ^{14}N, and easy to implement on samples with various quadrupolar coupling magnitudes, without very precise re-optimisation of ^{14}N pulse parameters. It is robust enough to investigate samples of varying quadrupolar couplings with relatively simple optimisation of RF field and pulse length for a particular sample. This observation is in line with the experimental observations for L-His, where all three ^{14}N sites with distinct ^{14}N quadrupole coupling parameters could be observed in a single experiment. The experiment also has the benefit of not having the experimental efficiency strongly correlated with ^{14}N RF power, even when the C_Q is high, as is typically the case for other ^{14}N *J*-HMQC sequences[75], meaning that the sequence is suitable for implementation with a wide range of probes with different sample volumes and coil sizes. The high-RF maxima in Figure 25B does, however indicate that there may be experimental conditions beyond the scope of the simulations performed here, where very high (>100kHz) ^{14}N RF amplitudes could potentially provide increases in efficiency that could be achieved with dedicated ^{14}N probes, microcoil probes[139, 140], or new generation ultrafast-MAS probes, where ^{14}N RF amplitudes could potentially reach or exceed 400kHz, and the coherence lifetime of the ‘spy’ nucleus should also be increased at higher MAS rates.

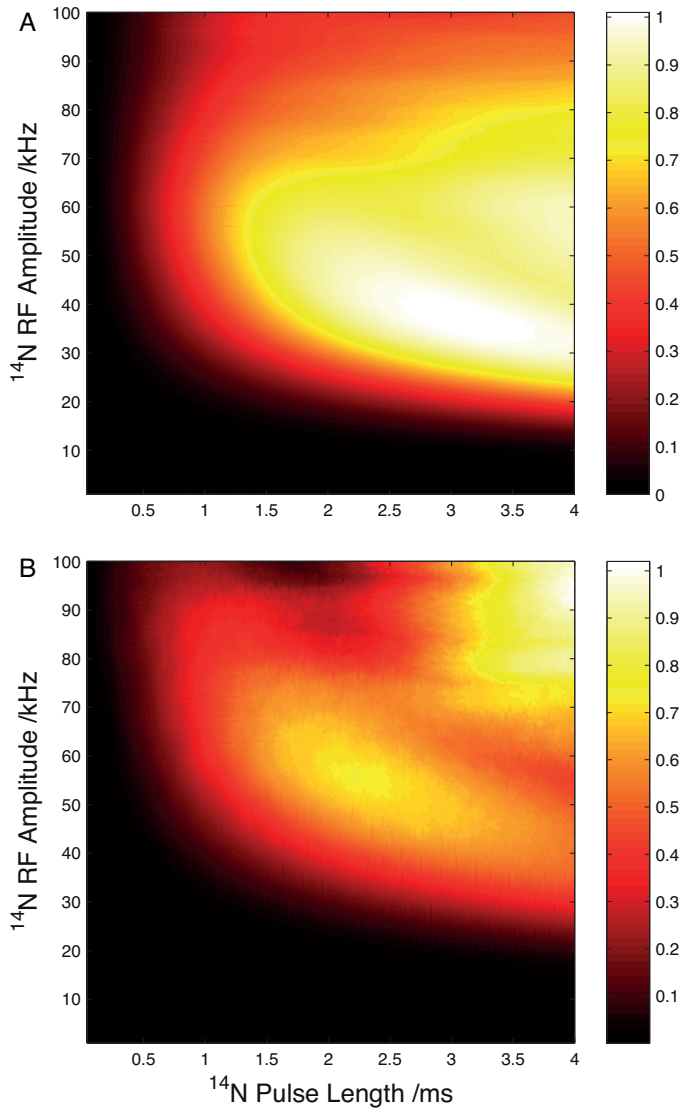

Figure 25. Contour Plots of Signal Intensity as a Function of ^{14}N Pulse Length and RF Amplitude

Simulated signal intensity is plotted against ^{14}N pulse length and the RF amplitude of ^{14}N pulses for a spin system representing (A) the primary amine site in glycine, C_Q=1.25MHz, and (B) a spin system representing the N-3 amide site of triglycine, C_Q=3.01MHz. Simulations performed at a B_0 field of 14.1T (600MHz ^{1}H Larmor frequency) and MAS frequency of 25kHz. In each plot, the data are normalised to 1. Figure adapted with permission from Jarvis et. al., 2013, [4].

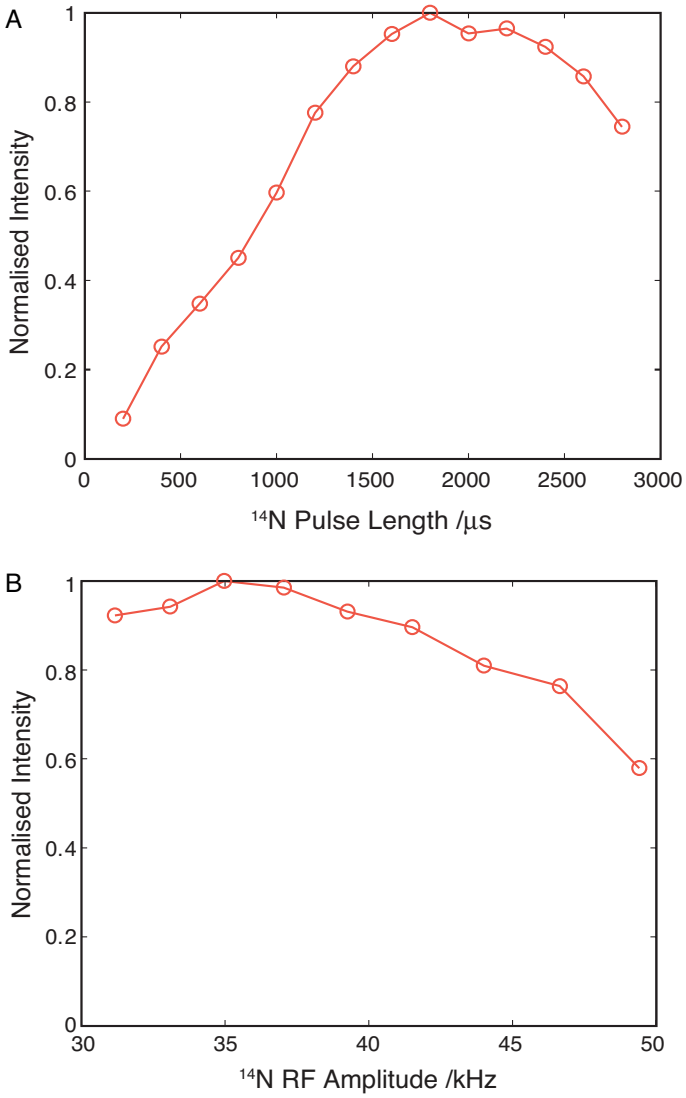

Figure 26. Plots of Experimental ^{14}N Transfer Efficiency as a Function of ^{14}N Pulse Length and Amplitude for 1,2-^{13}C-Glycine

The ^{14}N transfer efficiency of 1,2-^{13}C-Glycine is plotted as a function of ^{14}N pulse length in (A), and RF amplitude in (B), at B_0=14.1T, MAS rate=25kHz. Transfer efficiency is reported as the intensity of the ^{14}N filtered signal, expressed as a percentage of ^{13}C spin echo signal, normalised to the maximum observed transfer efficiency of 17.5%. Figure adapted with permission from Jarvis et. al., 2013, [4].

3.4 Conclusion

In this Chapter, a novel ssNMR method for investigating ^{14}N sites in biological molecules via proximal ^{13}C nuclei has been presented, and it has been applied to a variety of simple biological molecules to assess its properties. This method differs from previous methods for the indirect detection of ^{14}N in biological and organic molecules by the transfer and reconversion of coherence between ^{13}C and ^{14}N under periods of long (ms), moderate field strength (30-50kHz) RF irradiation on the ^{14}N channel.

This sequence has been assessed experimentally, and is demonstrated to be robust with respect to ^{14}N pulse width and amplitude, and applicable to samples containing a range of ^{14}N quadrupolar couplings, with only routine optimisation of the offset and length of pulses to ^{14}N, making the sequence relatively simple to implement. The sequence is also efficient, with efficiencies of up to 22% with respect to the CP signal reported for some ^{14}N sites; as well as sensitive, being able to record spectra from natural abundance material in a matter of hours. It has been shown that the position of peaks in indirect ^{14}N spectra, a function of both the ^{14}N δ_{CS}^{iso} and the ^{14}N δ_{Q}^{iso}, agree well with the expected values, calculated from previously determined ^{14}N quadrupolar coupling parameters. The spectra also have an increased spectral range compared to ^{15}N MAS spectra, whose peaks are a function solely of the ^{15}N δ_{CS}^{iso}. Crucially, this resolution is a function of the structure at ^{14}N sites, reflected by the quadrupolar coupling. It is demonstrated with simulations and calculations that relatively small (100s kHz) changes in C_Q magnitude at ^{14}N sites are easily discernible in ^{14}N spectra, manifested as shifts in peak position of 10s ppm at 14.1T. This demonstrates that the ^{14}N C_Q, and ^{14}N spectra, are extremely sensitive to small changes in molecular structure that could be associated with differences in bond lengths or geometries at ^{14}N sites. Changes in ^{14}N C_Q associated with the difference in H-bonding geometry between an α-helix and a β-sheet could potentially dominate the ^{14}N spectrum of a large peptide or protein.

The qualities of the method have also been assessed with numerical simulations. Simulations have been able to accurately represent the ^{14}N lineshapes from the 2D experiments presented here, which allows the accurate characterisation of ^{14}N sites in the molecules investigated. This also suggests that there may be a potential for precise characterisation of ^{14}N C_Q parameters, or even *de novo* determination of unknown ^{14}N quadrupolar couplings by using a simulation fitting approach. Furthermore, investigations of the effects of ^{14}N pulse durations and RF amplitudes by numerical simulation have confirmed the experimentally observed robustness of the sequence, since good efficiencies are observed for a relatively large range of pulse lengths and fields, at sites with a number of quadrupolar coupling magnitudes.

CHAPTER 4

INDIRECT DETECTION OF ^{14}N VIA ^{1}H IN ORGANIC COMPOUNDS

4.1 Introduction

In Chapter 3, a novel method for indirect ^{14}N detection via ^{13}C was presented. This method was based on the established HMQC sequence for detection of ^{14}N, but differs in the use ^{14}N RF irradiation to generate coherence between ^{13}C and ^{14}N spins, rather than periods of free evolution. It is possible to extend the indirect detection pulse sequence presented in Chapter 3 in order to detect signals on ^{1}H, to record ^{1}H/^{14}N correlation spectra. This has the potential to provide a substantial increase in the absolute sensitivity of ^{14}N indirect detection experiments over those detected via ^{13}C, since the ^{1}H gyromagnetic ratio is around 4x greater than that of ^{13}C, and the sensitivity of detection improves as $(\gamma)^{3/2}$ if all other variables are kept equal[141, 142] meaning an 8x sensitivity increase is expected for ^{1}H detection over ^{13}C detection. Moreover, the natural abundance of ^{1}H is >99.8%, and that of ^{13}C is 1.1%, meaning naturally abundant samples should additionally be almost 100x more sensitive using ^{1}H as a 'spy' nucleus, rather than ^{13}C. The only requirement is that the ^{14}N site be protonated, which is commonly the case in organic and biological samples. Furthermore, the dipolar coupling, RDS and *J*-couplings between ^{1}H and ^{14}N are expected to be larger than those between ^{13}C and ^{14}N, meaning that interactions that drive coherence transfer in all versions of HMQC type experiments are greater, resulting in further increased sensitivity. However, direct detection of protons in the solid-state is typically more challenging than detection of lower γ nuclei such as ^{13}C. This is due to the strong homonuclear dipolar couplings between abundant protons leading to rapid T_2' decoherence and homogeneously broadened resonances.

A number of methods to alleviate the issues with proton detection are available. There are many pulsed homonuclear decoupling methods to supress

strong $^{1}H/^{1}H$ homonuclear dipolar couplings in solids, such as DUMBO, Lee-Goldberg sequences and symmetry based recoupling sequences[143]. Alternatively, one may modify the chemistry of the molecules under study in order to dilute the strong homonuclear decoupling network. This is typically achieved with perdeuteration of samples, followed by exchange of labile deuterons with protons. For organic molecules, this is realised by synthesis with deuterated reagents before exchanging deuterons for protons. In proteins, this is typically achieved by expressing the protein on fully deuterated media and exchanging amide deuterons with protons in water. Deuteration and back exchange at 95% of exchangeable sites can reduce linewidths from almost 2kHz to a few hundred Hz small peptides at less than 10kHz MAS at a field of 500MHz[141]. In some perdeuterated microcrystalline samples linewidths of less than 20Hz are achievable at 24kHz MAS at a field of 400MHz with 10% of exchangeable amide sites protonated, leading to well resolved ^{1}H detected 2D correlation spectra, at the expense of the sensitivity of a fully protonated sample[144]. A third technique for suppressing the strong $^{1}H/^{1}H$ couplings that give rise to poor resolution and short ^{1}H coherence lifetimes is fast MAS. The resolution of strongly coupled, high-γ , protons increases approximately linearly with increasing MAS rate[145], along with the ^{1}H coherence lifetime[146, 147]. At a MAS rate of 29kHz, one starts to resolve individual sites from ^{1}H chemical shifts in fully protonated small organic molecules such as amino acids[148]. At MAS rates of 40kHz it is possible to record well resolved 2D ^{1}H correlation spectra in proteins in some exceptional cases[149]. Recent hardware development of MAS probes has seen the advent of ultra-fast MAS probes capable of spinning at frequencies of 60kHz - 111kHz. In this regime, the ^{1}H spectrum of entirely protonated small organic molecules can be fully resolved in a 1D experiment[150-153]. When, studying proteins with ultrafast MAS, polarisation transfer and protein assignment strategies akin to those employed in the solution-state are possible[146, 154]. One can also study fully protonated proteins more routinely[147, 155], though proton detected spectra on proteins typically still benefit from use of perdeuterated samples with some extent of amide proton back-exchange in this ‘ultrafast’ regime[154, 156].

In this Chapter, fast MAS (35kHz - 78kHz) is used is used to increase the resolution and T_2' of protons in $^{1}H/^{14}N$ detected correlation spectra recorded using a modified version the RF-driven indirect ^{14}N detection experiment presented in Chapter 3. Experiments were performed using the pulse sequence shown in Figure 27 on samples of natural abundance amino acids, an acetylated amino acid, and a tripeptide. All experiments were performed with fast MAS without sample deuteration or ^{1}H homonuclear decoupling as it has emerged as arguably the simplest, and one of the most effective methods for recording high-resolution ^{1}H spectra in the solid-state. Experimental $^{1}H/^{14}N$ transfer efficiencies in excess of 25% are observed in organic molecules, which is in excess of those typically reported for ^{1}H detected ^{14}N correlation spectra using either *J*- and RDS, or dipolar recoupling based HMQC methods. The increased efficiencies allow us to record higher quality indirect detection data, with well-defined features in the ^{14}N lineshapes. This permits us to address a current shortcoming of indirect ^{14}N detection experiments. As observed in Chapter 3, and previously demonstrated in the literature[1, 65, 79], it is possible to simply determine the quadrupolar product, $\chi_Q = C_Q\sqrt{1+(\eta^2/3)}$, of a ^{14}N site from the ^{14}N shift if the δ_{CS}^{iso} at the site is known. However, since C_Q and η are correlated, one cannot simply determine the value of either from the peak position. To achieve this, a fitting of the lineshape to simulated spectra to determine η would be required. This has not previously been attempted largely due to poorly resolved spectral features in ^{14}N lineshapes due to low S/N[1, 64]. In this Chapter, using simulations performed in the *Spinach* library[157], a simulation fitting approach to attempt to define ^{14}N EFG tensor parameters, C_Q and η, more precisely than has previously been achieved in indirect ^{14}N detection experiments is presented, and the results, limitations and potential improvements for the approach are discussed.

4.2 Indirect ^{1}H-Detected Studies of Biological Samples

Proton detected ^{14}N correlation experiments, using the pulse sequence shown in Figure 27, have been carried out on polycrystalline samples of the amino acids L-histidine (L-His), L-histidine hydrochloride monohydrate (His.HCl.H_2O) and N-acetyl valine (NAV). The pulse sequence is analogous to the ^{13}C detected experiment presented in Figure 17, but detected on ^{1}H. Also omitted is the ^{1}H decoupling during t_1. This was omitted here, as experiments were performed in a MAS regime (≥35kHz) where ^{1}H/heteronucleus dipolar couplings are more efficiently removed by MAS, and are not as important for preserving the coherence lifetime of the detected nucleus for ^{1}H detected experiments as for ^{13}C detected experiments. Additionally, with fast MAS, high power proton decoupling (100kHz – 120kHz) can be problematic due to the potential for broad rotary resonance recoupling conditions at 100kHz – 120kHz depending on the exact spinning frequency[158, 159]. In all experiments, SQ signals were detected on ^{14}N in the indirect dimension, selected by phase cycling, as described in Figure 17. Experiments were carried out at two different fields and a number of spinning speeds to assess the effects of these parameters on experimental efficiency and spectral features.

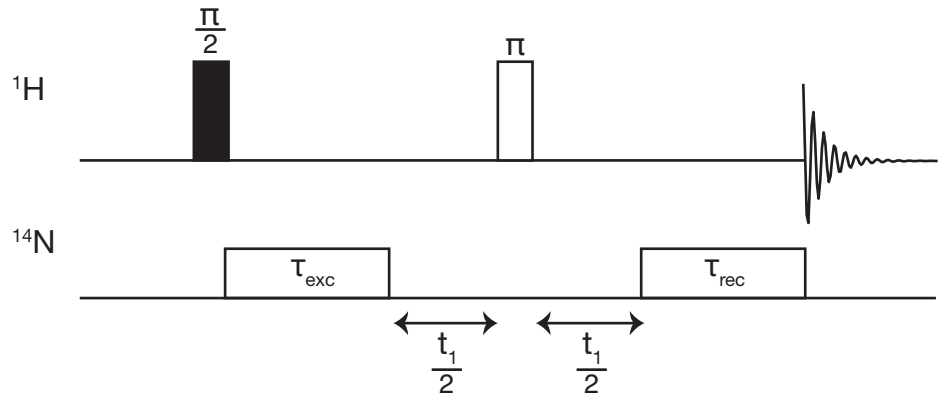

Figure 27. Pulse Sequence for Indirect Detection of ^{14}N via ^{1}H

Pulses on the ^{14}N channel, τ_{exc} and τ_{rec}, are applied for 100s μs. In common with other indirect ^{14}N detection sequences, t_1 evolution must be rotor synchronised, with $\Delta t_1=\tau_r$, to ensure removal of contributions from the first order quadrupolar interaction to the spectrum.

4.2.1 Materials

L-His, His.HCl.H_2O, and NAV were purchased from Sigma-Aldrich UK and used without further purification. The chemical structure of each small molecule used in the experiments described in this Chapter are shown in Figure 28.

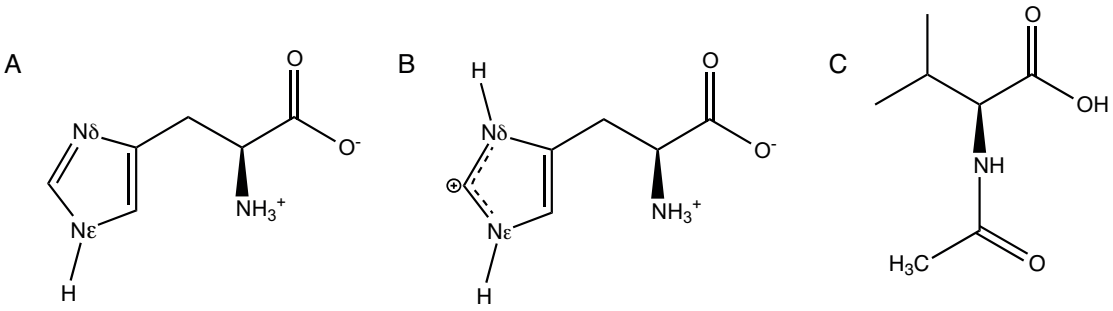

Figure 28. Chemical Structures of Compounds Used in $^{1}H/^{14}N$ Experiments

(A) L-Histidine (B) L-Histidine hydrochloride monohydrate (C) N-Acetyl-Valine.

4.2.1 Solid-State NMR

Experiments were carried out on L-His and His.HCl.H_2O on a Varian DDR2 Spectrometer operating at 14.1T (Larmor frequencies of 600MHz and 43.3MHz for ^{1}H and ^{14}N respectively), equipped with a 1.6mm triple resonance MAS probe tuned to ^{1}H and ^{14}N Larmor frequencies in double resonance mode. Samples were spun at 35.5kHz and 40kHz MAS rate, respectively, for His.HCl.H_2O and L-His samples. Proton π/2 and π pulse lengths were 1.55μs and 3.10μs on both samples. ^{14}N excitation and reconversion pulses were used with amplitudes of 30kHz for 390μs and 62.5kHz for 575μs, for His.HCl.H_2O and L-His, respectively, as measured on solid NH_4Cl. In both cases 2D spectra were recorded in a phase sensitive manner using States-TPPI[115], with 254 increments recorded with 64 scans per increment for His.HCl.H_2O, and 160 increments recorded with 112 scans per increment for L-His. Both spectra were recorded with a recycle delay of 2s. The spectra were zerofilled to 2048 points in both dimensions, before 2D Fourier transform without a window function applied.

Further experiments were carried out on NAV. NMR experiments were performed on a Bruker Avance II spectrometer operating at a field of 20.0T (proton and ^{14}N Larmor frequencies of 850MHz and 61.4MHz), and a JOEL 1mm double resonance probe tuned to ^{1}H and ^{14}N frequencies. MAS frequency was 78kHz, ^{1}H π/2 and π pulse lengths were 0.9μs and 1.8μs, and the ^{14}N pulses were applied for 370μs, at an amplitude of 72kHz, as measured on solid NH_4Cl. Data were recorded with 128 complex t_1 increments with 136 scans per increment, using States-TPPI[115]. The recycle delay was 3s. Data were zerofilled to 2048 points in each dimension before 2D Fourier transform.

Experiments carried out at both fields were processed with NMRPipe[116]. All spectra are referenced to solid NH_4Cl at 39.3ppm for the ^{14}N dimension, corresponding to liquid ammonia at 0ppm[127]. The proton dimensions are referenced to DSS at 0ppm, indirectly referenced according to [118], from a ^{13}C spectrum of adamantane, where the downfield peak is referenced to 40.40ppm[117].

4.3 Numerical Simulations of ^{1}H Detected ^{14}N Indirect Detection Experiments

4.3.1 *Spinach*

All simulations were performed using the *Spinach* library, version 1.6.2[157]. The simulations were written by Prof. Ilya Kuprov, University of Southampton. Spin systems for each nitrogen site investigated in each molecule were approximated as ^{1}H/^{14}N spin pairs comprising the ^{14}N nucleus and its closest bound proton. CSA parameters were specified for each nucleus in the simulations, and dipolar and *J*-couplings were included between the two nuclei, as well as the ^{14}N quadrupolar coupling. Initial values for ^{14}N quadrupolar couplings were taken from literature values, or from CASTEP calculations performed by Mr. Richard Bounds, University of Southampton. A summary of the sites simulated and the initial parameters used for each simulation are given in Table 6. In order to compare simulations to experimental lineshapes, a 9x9 grid of values for C_Q and η was constructed for each site simulated, with values of each within ±0.25MHz for the C_Q magnitude, and ±0.25 for η. The RMSD between the simulated and experimental spectra was calculated for each simulation in the grid, according to:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2} \qquad (4.1)$$

Where x_i and y_i are points in the experimental and simulated spectra, respectively. All simulations neglected the effects of relaxation, and had an apodization functions applied to the FIDs to match as closely as possible the effective T_2 at each site in both direct and indirect dimensions. Simulations of each site in His.HCl.H_2O and L-His were performed with 770 α and β orientations using the Lebedev method, having verified this was sufficient for convergence in each case. The maximum Floquet rank used was determined by the size of the quadrupolar interactions at the ^{14}N site and the MAS frequency;

for NH_3^+ and Nε sites in L-His and His.HCl.H_2O, a rank of 45 was used, and for the Nδ in His.HCl.H_2O, a rank of 55 was used. For simulations of NAV at 75kHz MAS, 1454 α and β orientations were required for convergence, and a Floquet rank of 55 was used. In all cases it was verified that the number of powder points and the Floquet rank was sufficient for convergence of the simulations. Simulations were run on high-memory nodes of the IRIDIS4 supercomputer, each with 32 cores and 256GB RAM, and each simulation for histidine sites took ~1 hour, and simulations of NAV with an increased Floquet rank took ~1 hour 45 minutes.

Table 6. Spin System Parameters Used in Simulations of $^1H/^{14}N$ Correlation Spectra

The ^{14}N δ_{CS}^{iso} used in simulations was as measured in a separate ^{15}N CPMAS experiment on each sample. Values for the ^{14}N C_Q and η in NAV are from [22], with orientations from [130]. Values for the ^{14}N C_Q and η are from [160] for L-His, with an orientation from[24]. These parameters were taken from a CASTEP calculation in the case of His.HCl.H_2O.

Sample	*Site*	δ_{CS}^{iso} *(ppm)*	C_Q *(MHz)*	*η*	^{14}N C_Q *Orientation (°,{α,β,γ})*
His.Hcl.H_2O	NH_3^+	47.3	1.32	0.06	0,0,0
	Nε	177.2	1.33	0.94	0,90,0
	Nδ	189.8	1.53	0.23	0,90,0
L-His	NH_3^+	41.6	1.22	0.15	0,0,0
	Nε	168.9	1.44	0.92	0,90,0
NAV	N1	127.7	3.21	0.32	13.2,90,103

4.4 Results and Discussion

To evaluate the efficiency of proton detected experiments on a range of ^{14}N sites, as well as to asses the viability of a simulation fitting approach to determine ^{14}N EFG parameters from fitting of numerically simulated data to experimental spectra, experiments have been carried out on a number of organic samples containing a range of nitrogen sites with various NMR interaction parameters. Data was collected on a number of amino acids containing a number of different amine and heterocyclic ^{14}N sites, as well as NAV, which contains an amide ^{14}N site with a large (>3MHz) quadrupolar coupling typical of those found in peptides and proteins.

4.4.1 2D $^{1}H/^{14}N$-Correlation Experiments: L-His and His.HCl.H_2O

Samples of L-His and His.HCl.H_2O were studied in order to gauge the effects of small changes in local geometry and hydrogen bonding arrangement between the samples on the spectral features of indirectly detected ^{14}N spectra. $^{1}H/^{14}N$ correlation spectra for both these compounds are shown in Figure 29. In the spectrum of L-His, two peaks are observed, corresponding to the NH_3^+ and Nε. These resonances are observed at ^{14}N shifts of 100ppm and 278ppm, respectively, measured from the centre of gravity of the signal. In the spectrum of His.HCl.H_2O, the NH_3^+ and Nε correlations are also observed, as well as a peak from the Nδ site. In His.HCl.H_2O, unlike in L-His, the Nδ site is protonated, forming a hydrogen bond with an adjacent carboxyl group and is detected, as shown in Figure 29. The ^{14}N shifts at NH_3^+, Nδ and Nε sites are 100ppm, 263ppm, and 280ppm, respectively.

The efficiency of the two way ^{14}N polarisation transfer was measured at each ^{14}N site, as a percentage of the ^{1}H spin echo signal, with echo periods equal to ^{14}N pulse lengths in the ^{14}N filtered spectrum. In His.HCl.H_2O, the transfer efficiencies were 27.0% at Nε, 24.4% at Nδ, and 7.5% at the NH_3^+ site. This experiment was optimised in terms of ^{14}N pulse amplitude and duration such that the signal intensity was highest at the two heterocyclic nitrogen sites over the NH_3^+ site, and could be optimised to achieve a more even spread of

signal intensity over the three sites. In the experiment on L-His, ^{14}N transfer efficiencies of 20.0% and 16.6% at Nε and NH_3^+ sites, respectively, were recorded. The efficiencies observed in this experiment are well in excess of those previously observed in the literature for indirect detection of ^{14}N via ^{1}H experiments. Using a *J*- and RDS based SQ-HMQC on a sample of partially deuterated glycine spinning at 30kHz with a B_0 field of 9.4T, Cavadini *et al*. reported experimental efficiencies of 2%[62]. Recently, in an experimental regime with much higher spinning rates, Nishiyama and coworkers performed a similar *J*- and RDS based transfer at a B_0 field of 14.1T and 90kHz MAS, and reported ^{14}N transfer efficiencies of 7.55% and 9.75% at the Nε and Nδ sites in His.HCl.H_2O[161]. Gan has reported transfer efficiencies of 5-6% on a sample of natural abundance glycine spinning at 20kHz at a B_0 field of 14.1T using a HMQC experiment with rotary resonance recoupling to recouple the ^{1}H/^{14}N dipolar interaction[64]. Our results compare very favourably with those in the literature without using deuteration or dipolar recoupling schemes, and the experiment presented here is likely to benefit from increased MAS rates to increase resolution.

The ^{14}N δ_Q^{iso} at each site was determined in the same way as was demonstrated for ^{13}C/^{14}N correlation spectra in the previous Chapter; the ^{14}N δ_{CS}^{iso} was subtracted from the ^{14}N shift (at the centre of gravity of the ^{14}N peak). The ^{14}N δ_{CS}^{iso} was assumed to be the same as the ^{15}N δ_{CS}^{iso} for each site and the ^{15}N δ_{CS}^{iso} was measured in a separate ^{15}N CPMAS experiment for each sample. The experimental values for the ^{14}N δ_Q^{iso} are given in Table 7, as well as expected values of the ^{14}N δ_Q^{iso} for each site at this field, calculated from literature values of C_Q and η, according to the expression for the ^{14}N δ_Q^{iso} in ppm given in Equations (1.49) and (1.50), repeated below:

$$^{14}\mathrm{N}_Q^{iso} = \frac{3}{40}\left(\frac{\chi_Q}{\nu_0}\right)^2 \times 10^6$$

$$\chi_Q = C_Q\sqrt{1+\frac{\eta^2}{3}} \qquad (4.2)$$

Table 7. Experimental and Calculated ^{14}N Second Order Quadrupolar Shifts

Values for C_Q and η used to calculate values of ^{14}N δ_Q^{iso} are those given in Table 6. Values are calculated according to Equation (4.2). The ^{15}N chemical shifts used to determine the experimental ^{14}N δ_Q^{iso} are also given in Table 6.

Sample	*Site*	*^{14}N $\boldsymbol{\delta_Q^{iso}}$ Experimental (ppm)*	*^{14}N $\boldsymbol{\delta_Q^{iso}}$ Calculated (ppm)*
His.Hcl.H_2O	NH_3^+	55.4	69.5
	Nε	85.8	91.1
	Nδ	90.2	94.9
L-His	NH_3^+	58.4	59.7
	Nε	105.4	105.9
NAV	N1	207.3	211.5

For ^{14}N sites in L-His, a good agreement is observed between the experimental values, and those calculated from literature ^{14}N EFG tensor parameters for the ^{14}N δ_Q^{iso} at both nitrogen sites in this molecule, to within ±1.5ppm. This discrepancy is acceptable considering the error introduced by reading the ^{14}N shift from the centre of gravity of peaks which are typically ~50ppm wide. The agreement between experiment and calculation is not as good for the three ^{14}N sites in His.HCl.H_2O, where the calculated ^{14}N δ_Q^{iso} is consistently larger than that observed experimentally. For both heterocyclic sites, the calculated ^{14}N δ_Q^{iso} differs from the experimental ^{14}N δ_Q^{iso} by ~5ppm. At the primary amine site the calculated ^{14}N δ_Q^{iso} is ~15ppm greater than that observed experimentally. These effects may be due to the motion in the sample scaling quadrupolar couplings, which are not accounted for in the CASTEP calculations, making the effect particularly pronounced at the amine site, which is likely to be undergoing rapid rotation. Indeed, our experimentally observed ^{14}N δ_Q^{iso} values for His.HCl.H_2O are in better agreement (within ±1.5ppm) with values determined by NMR on a single crystal at a similar temperature from the literature[24].

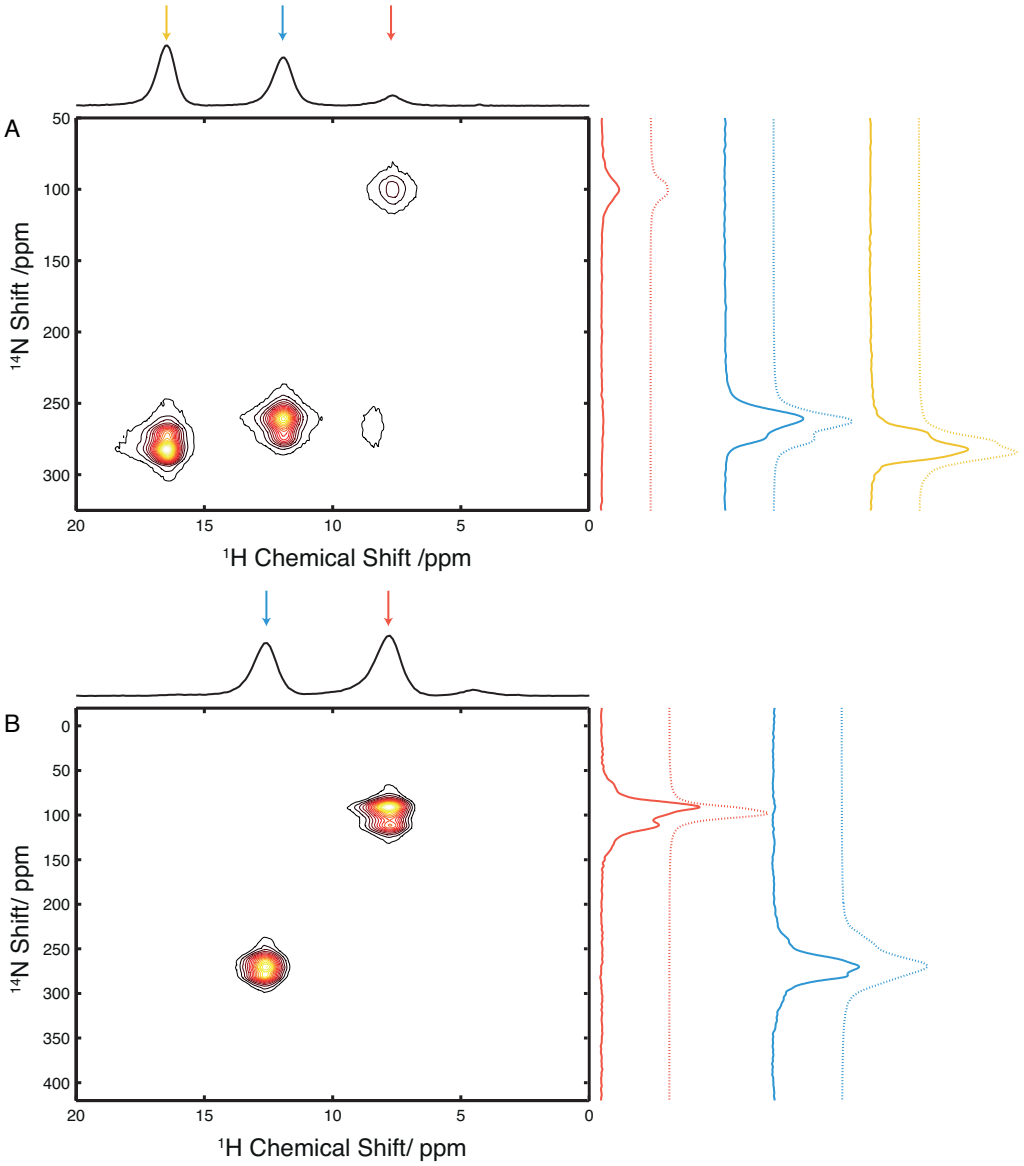

Figure 29. ^{1}H/^{14}N Correlation Spectra of Two Forms of Histidine at 14.1T

(A) Spectrum of His.HCl.H_2O recorded at 35.5kHz MAS frequency, (B) L-His spectrum Recorded at 40kHz MAS frequency. Experimental slices (solid lines) through the ^{14}N dimension at locations indicated by coloured arrows are plotted for each spectrum at the NH_3^+ sites (red), the Nε sites (blue) and the Nδ site (yellow). Best-fit simulations of each the sites (dashed lines) are plotted alongside each experimental trace. The slices are normalised to the maximum peak intensity in either spectrum.

4.4.1.1 Simulation Fitting of Histidine ^{14}N Lineshapes

From indirectly detected ^{14}N spectra, it is relatively simple to determine the ^{14}N δ_Q^{iso} from the ^{14}N shift, as described in the previous section. From this, the quadrupolar product, $\chi_Q = C_Q\sqrt{1+(\eta^2/3)}$, may be determined, which contains a contribution from both of the parameters required to define the ^{14}N EFG tensor, C_Q and η. However, their contribution to the ^{14}N δ_Q^{iso} is correlated, and therefore neither can be precisely determined from the ^{14}N shift. In order to determine the ^{14}N EFG tensor parameters more precisely, the asymmetry of the tensor needs to be determined from the lineshape. The ^{14}N lineshape also depends on both C_Q and η, and so these parameters can, theoretically, be found by a simulation fitting approach where these parameters are varied in separate simulations, and the fit to the experimental data evaluated. This has not previously been attempted for indirectly detected ^{14}N spectra. The most pressing reason for this has been the limited S/N in indirectly detected experiments has meant that subtle features in the indirectly detected lineshapes are not well defined. Here, having recorded indirect ^{14}N detected spectra with very high efficiencies, the quality of data was high enough to observe subtle spectral features in the ^{14}N lineshapes, for example, the shoulders on the sides of Nε and Nδ peaks in the His.HCl.H_2O spectrum. A simulation fitting approach was therefore adopted in order to assess the feasibility of determining C_Q and η from indirectly detected lineshapes.

Firstly, the accuracy with which C_Q can be determined simply from the ^{14}N δ_Q^{iso} was evaluated. Inspection of the equation for the ^{14}N δ_Q^{iso} demonstrates that a range of values of C_Q and η will be consistent with a given ^{14}N δ_Q^{iso}, which will depend on the field strength that the experiment is carried out at, and the magnitude of the C_Q. A simple way to visualise this is to plot the ^{14}N δ_Q^{iso} as a function of C_Q and η. From such a plot, upper and lower bounds can be placed on the C_Q values consistent with the observed ^{14}N δ_Q^{iso} for all (or particular) values of η. Such a plot is shown in Figure 30 for all sites in the histidine compounds investigated here.

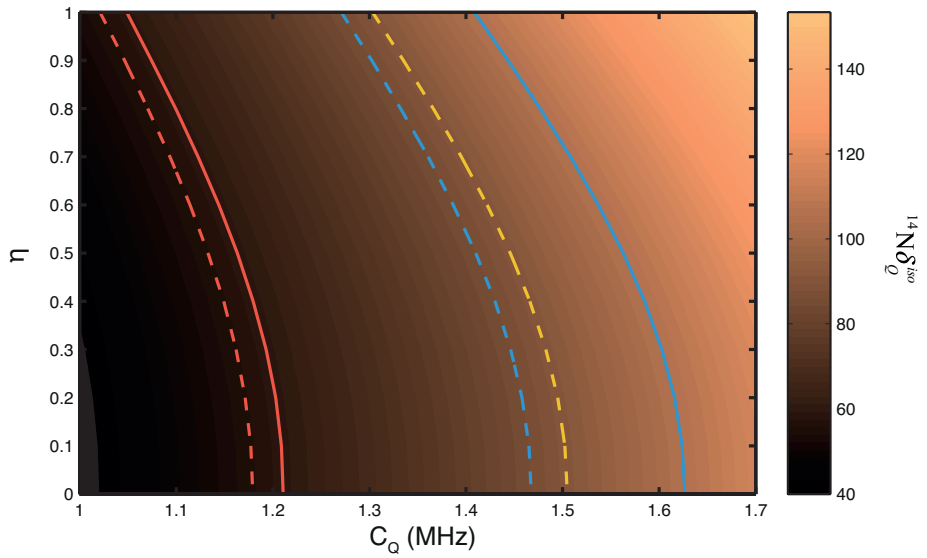

Figure 30. Plot of ^{14}N Second Order Isotropic Quadrupolar Shift at 14.1T

The $^{14}N\ \delta_Q^{iso}$ is plotted as a function of C_Q and η at a field strength of B_0=14.1T, according to Equation (4.2)*. Coloured contour lines are displayed for experimental values of $^{14}N\ \delta_Q^{iso}$(found in Table 7) for NH_3^+ (red), Nε (blue), and Nδ (yellow) sites in His.HCl.H_2O (dotted lines), and L-His (solid lines).*

Contour lines are displayed in Figure 30 which are equal to the value of the $^{14}N\ \delta_Q^{iso}$ (Table 7) for each nitrogen site in His.HCl.H_2O and L-His. One can easily read off the value of the ^{14}N C_Q where η=0 and η=1 for the contour line for each site, thereby determining the upper and lower bounds for C_Q for a given ^{14}N shift. For His.HCl.H_2O, for the sites NH_3^+, Nε and Nδ, respectively, ^{14}N C_Q =1.02MHz-1.19MHz, 1.27MHz-1.47MHz and 1.30MHz-1.50MHz when 0 > η >1. For L-His, at the NH_3^+ and Nε sites, respectively, C_Q=1.05MHz-1.21MHz and 1.41MHz-1.63MHz, when 0 > η >1. Therefore, at 14.1T, investigating ^{14}N EFG tensors with magnitudes in the range of 1MHz-1.5MHz, it is possible to constrain the C_Q to a range of ~200kHz for all values of η, simply by determining the $^{14}N\ \delta_Q^{iso}$. It is also apparent from these values that as the C_Q at a given site increases, the precision with which its C_Q can be determined from the $^{14}N\ \delta_Q^{iso}$ is reduced; the amine sites in each molecule have a $^{14}N\ \delta_Q^{iso}$

consistent with C_Qs within a range of 170kHz, while the Nε site in L-His, which has the largest C_Q of all those in the two histidine molecules, has its C_Q defined within a range of 220kHz.

In order to try and determine C_Q and η more precisely, the lineshapes were fitted to numerical simulations of the data. A 9x9 grid of values for C_Q and η close to the literature and calculated ^{14}N values (C_Q=±0.25MHz, η=±0.25) was constructed, and a 2D ^{1}H/^{14}N spectrum simulated for each point on the grid. The RMSD of the simulated ^{14}N lineshape in the indirect dimension with respect to the simulated lineshape for that site was calculated. A contour plot of the RMSD of simulated lineshapes with respect to the experimental ones is shown for each site in each histidine compound in Figure 31.

There is a strip of relatively good agreement (RMSD<~0.3) through the RMSD plot for each ^{14}N site where the simulated ^{14}N δ_Q^{iso} is consistent with that observed experimentally. A single black contour line through each plot marks the experimentally observed ^{14}N δ_Q^{iso}, which are also plotted on the contour plot in Figure 30, and tabulated in Table 7. This line generally coincides with the area of lowest RMSD in each plot. This shows that the largest contribution to the quality of the simulations fit is determined from the ^{14}N δ_Q^{iso}, which is to be expected since it determines the position of the peak. In the RMSD plot for the Nδ of His.HCl.H_2O, a single RMSD minimum where the RMSD is <0.15 (darkest blue contours) is observed. This covers a range of C_Q=1.49MHz-1.51MHz, and η=0.22-0.34, which is in close agreement with the CASTEP calculated values at this site, C_Q=1.53MHz, and η=0.23, as well as literature values of C_Q=1.47MHz, and η=0.27[24]. Additionally, visual comparison of the simulated lineshape with the experimental (Figure 29A, yellow traces) shows a good agreement. The slight deviation between the black contour line indicating the experimental ^{14}N δ_Q^{iso} and the literature values is quite likely due to the coarseness of the grid used for simulations; C_Q values were spaced at 50kHz intervals, which may be a limiting factor here in determining the C_Q. Nonetheless, this plot demonstrates that using a simulation fitting approach, one can define η to a range of ~±0.1, which allows the C_Q to be determined with a precision of tens of kHz, which is at

least an order of magnitude more precision than one can achieve considering only the ^{14}N δ_Q^{iso}.

In the RMSD plots for the Nε sites of both His.HCl.H_2O and L-His, the fit of the lineshape with respect to η is somewhat less robust. In the case of the Nε site in His.HCl.H_2O, two RMSD minima <0.2 are observed, at C_Q=1.29MHz-1.3MHz, η=0.93-0.9.7, and C_Q=1.39MHz-1.41MHz, η=0.60-0.68. The former of these minima, however, is consistent with the calculated values for this site, of C_Q=1.33MHz, η=0.94, as well as literature values of C_Q=1.29MHz, η=0.95[24]. For the Nε site in L-His, the same general pattern is observed, with two RMSD minima in similar positions with respect to η as those observed at the same site in His.HCl.H_2O (C_Q=1.35MHz, η=0.95-1.00, and C_Q=1.5MHz, η=0.60-0.65). However, the fit is not as good as that of the Nε site in the hydrochloride salt; in L-His the RMSD minima are broader, and the RMSD is not as low (minimum RMSD=0.24). This is probably due to there being less well defined features in the experimental lineshape at this site (as shown in Figure 29B) as compared to the Nε in His.HCl.H_2O. However, these sites are expected to have similar chemical environments and hydrogen bond geometries, and therefore similar ^{14}N spectral parameters in terms of their ^{14}N EFG and CSA tensors which give rise to the features observed in the ^{14}N lineshapes. The difference in the lineshapes may be due to the quite different length and amplitude of ^{14}N pulses used in the experiments on the different compounds giving rise to quite different lineshapes at these similar sites.

The NH_3^+ sites in both His.HCl.H_2O and L-His show a poor fit of the simulated lineshape with respect to η, with the minimum RMSD encompassing almost the whole range of η values simulated, and not falling below RMSD=0.3. This means it is not possible to use these plots to constrain the C_Q at either site more precisely than was achieved by considering simply the ^{14}N δ_Q^{iso}. There are a number of reasons that could account for the poor fit at these sites. In His.HCl.H_2O, the experimental ^{14}N lineshape is simply a featureless peak, without any discontinuities, shoulders or splitting. This makes it difficult to fit simulated spectra to. This may be due to fast ^1H T_2' decoherence or ^{14}N $T_{1\rho}$ at this site leading to broadening of the ^{14}N signal beyond detection of any sharp

features. Indeed, simulated spectra for this site displayed well defined shoulders and splittings in ^{14}N lineshapes, which were broadened out by the apodization function applied to match the decay of the FID of of the experiment, meaning a higher resolution lineshape at this site may be easier to fit. In this case spinning at a higher frequency may help to increase ^{1}H T_2' and resolve features in the ^{14}N lineshape. It would also be beneficial to perform another experiment specifically optimised for maximum signal at the NH_3^+ site, rather than the sidechain ^{14}N sites as in the experiment here. This should lead to greater S/N in the experimental lineshape, and a potentially higher quality spectrum. Alternatively, it has been previously demonstrated that SQ quadrupolar lineshapes are particularly sensitive to molecular motions[162-164]. This has also been demonstrated specifically for NH_3^+ groups in indirectly detected SQ ^{14}N lineshapes[77]. The authors demonstrated that the NH_3^+ group in a solid peptide underwent motions on a ns timescale that broadened ^{14}N SQ lineshapes from <1kHz to >5kHz. It is possible that the motion of the NH_3^+ in His.HCl.H_2O caused broadening of the ^{14}N resonance that was not accounted for in the simulations, resulting in a poor fit. At the NH_3^+ site in L-His, a well-defined splitting was observed in the ^{14}N lineshape. However, the fit of the simulated spectra to this lineshape was the least satisfactory of all those investigated here, with a minimum RMSD=0.35. As the experimental and simulated slices (Figure 29B, red traces) show, the experimental spectrum observed was significantly broader, with more features, than the best fit simulated spectrum. The reason for this may be that at the NH_3^+, the ^{14}N is coupled to three separate protons, which cannot be resolved in the ^{1}H spectrum, and so their contributions to the ^{14}N lineshape are superimposed in the ^{14}N dimension. The simulated lineshape only takes into account an approximation of the spin system with ^{14}N coupled to a single amine proton. This would mean that the simulation only accounted for part of the experimental lineshape, and that the remaining features were due to coupling to the other protons at the NH_3^+.

The quality of the fits of the simulated data with respect to the experimental data varies from site to site. In the best case, of the Nδ site in His.HCl.H2O, an RMSD of less than 1.5 is observed, however at the amine sites

in both compounds the fit is less robust, particularly in L-His where the minimum RMSD=0.35. The poor characterisation of the experimental lineshape by the simulations at the amine sites we attribute to the difficulty of simulating ^{14}N amine sites which are potentially highly mobile, and have three bound protons that are difficult to approximate accurately in a two spin systems for simulation. A potential solution to this would be to include effects of motions, or more spins in the simulations, though these simulations are already computationally intensive and close to the limit of what we can achieve in a reasonable timeframe (several weeks of supercomputer time) when fitting a grid of parameters to experimental data. A possible alternative, if the poor fit was due to motional broadening would be to investigate indirect detection of the ^{14}N DQ lineshape. This could be selected for experimentally by phase cycling, and can be easily selected for in simulations. While the DQ experiment is expected to be less efficient[2, 21, 62, 67, 74, 77], it is also expected to be less sensitive to molecular motions, meaning that spectral features arising from dynamics should be suppressed potentially improving the agreement with simulated spectra.

At the heterocyclic ^{14}N sites, where a better quality fit of the data is generally achieved, how well the ^{14}N EFG tensor is characterised by the simulations is variable. At the Nδ site of His.HCl.H2O, a single RMSD minimum is observed, which agrees well with calculated and literature values for this site, and the ^{14}N EFG tensor is well characterised by the simulated spectra. However, at the Nε sites of both histidines, while the minimum RMSD values are relatively low, there are multiple minima or broad minima that cover a wide range of C_Q and η values that are consistent with the experimental data; the ^{14}N EFG tensor is not well characterised. This may be due to the magnitude of the C_Q and η being correlated in their effects on both the ^{14}N lineshape and the position of the peak. In this case, the effects on the lineshape of η will be scaled by the C_Q meaning that when the C_Q is low, η has a limited effect on the ^{14}N lineshape, and its effects are observed in the lineshape more readily when the C_Q is higher. This would account for the EFG tensor parameters of the Nδ site being better characterised than at the Nε sites, where the C_Q is lower by 200kHz.

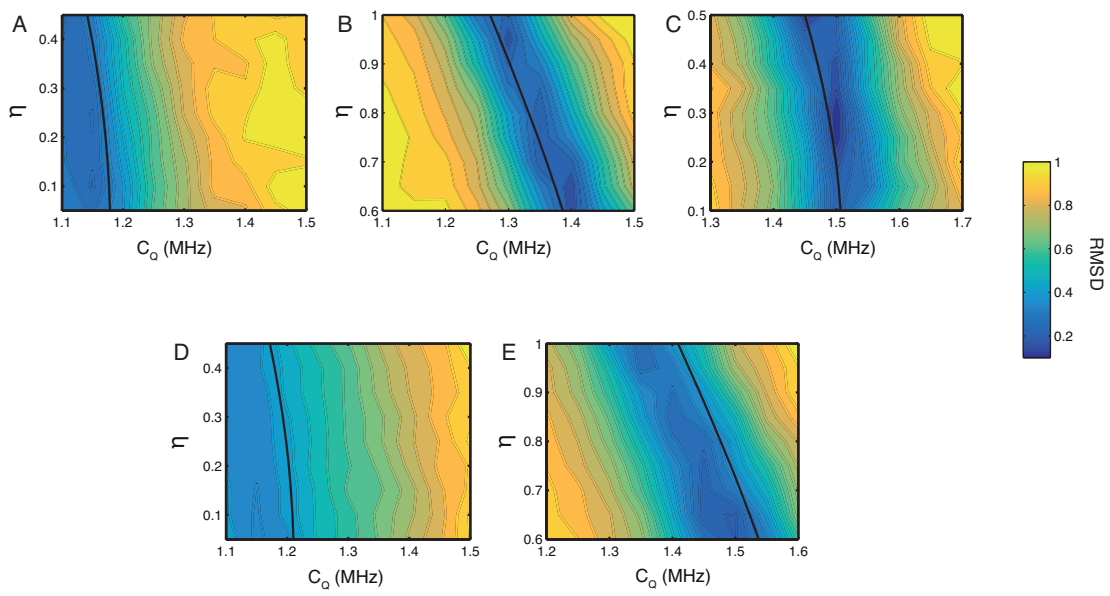

Figure 31. RMSD Plots of Simulated and Experimental ^{14}N Lineshapes of His.HCl.H_2O and L-His

The RMSD of simulated ^{14}N lineshapes with respect to experimental lineshapes is plotted as a function of the ^{14}N C_Q and η used in the simulation. (A) His.HCl.H_2O NH_3^+ (B) His.HCl.H_2O Nε (C) His.HCl.H_2O Nδ (D) L-His NH_3^+ (E) L-His Nε. Contour levels are spaced at 5% max RMSD from 5% max RMSD. A single black contour line in each plot marks the ^{14}N δ_Q^{iso} for each site, calculated from the ^{14}N shift-^{14}N δ_{cs}^{iso}.

4.4.2 2D $^{1}H/^{14}N$-Correlation Experiments: N-Acetyl-Valine

Further proton detected experiments were performed on NAV, which has an amide nitrogen site that differs from the primary amine site found in non-acetylated amino acids in that it exhibits an sp^2 hybridisation state, rather than an sp^3 state, with lower symmetry. This means that the quadrupolar coupling is increased compared to the amine sites of amino acids. The C_Q of ^{14}N in NAV is reported in the literature as 3.21MHz, η=0.32[22], this is similar to the C_Q expected at ^{14}N sites in the peptide bond in peptides and proteins, since the structure of these sites is similar. Experiments on NAV, therefore, should be a good mimic of experiments on peptides and proteins.

Figure 32 shows a 2D $^{1}H/^{14}N$ correlation spectrum recorded on a sample of NAV at 20.0T (850MHz ^{1}H Larmor frequency) and a MAS frequency of 78kHz. Experiments were performed at a higher field with ultra-fast MAS in order to optimise conditions for proton detection on this challenging, high-C_Q sample. An increased MAS rate leads to improved ^{1}H coherence lifetime, which is critical for the ^{14}N filtered experiment. Additionally, increasing the magnetic field strength increases the ^{1}H chemical shift range, meaning that proximal protons have an increased frequency separation, reducing dipolar flip-flop terms between them. This results in a reduction in homogenous linewidth that benefits the resolution of directly detected protons, as well as the coherence lifetime that benefits the sensitivity of ^{14}N filtered experiments. The ^{14}N shift of the amide ^{14}N in NAV at this field is 335ppm. From this shift, a ^{14}N δ_Q^{iso}=207.3ppm is calculated, by subtracting the δ_{cs}^{iso} of 127.7ppm (measured by ^{15}N CPMAS). This agrees well with the expected ^{14}N δ_Q^{iso} value of 211.5ppm calculated for this field from literature values[22]. The efficiency of the two-way ^{14}N transfer is 9% with respect to the ^{1}H spin echo signal under the same conditions. This is lower than the efficiencies of the proton detected experiment generally observed on the lower magnitude C_Q sites in the histidine compounds. This is a trend observed in all indirect ^{14}N detection experiments, and can be attributed to the increased magnitude of the ^{14}N C_Q in NAV, which is likely to mean that the pulses applied on the ^{14}N channel excite the far broader ^{14}N spectrum to a lesser extent. The

efficiency reported here cannot be compared to that of other ^{14}N indirect detection experiments on comparable compounds, since no such data has been reported in the literature for a proton detected experiment on a sample containing ^{14}N sites with large (~3MHz) quadrupolar couplings.

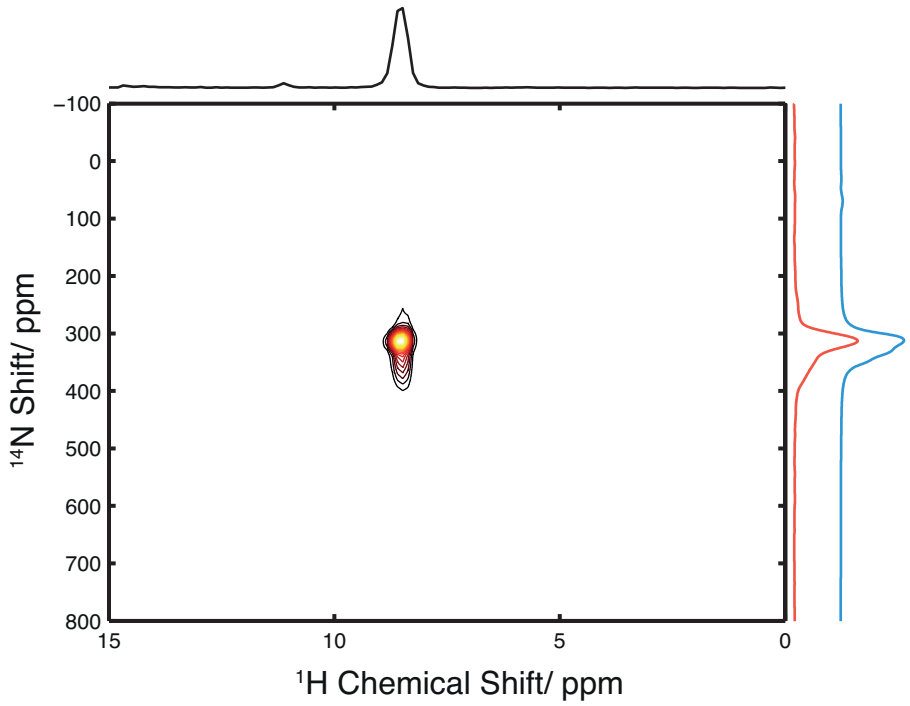

Figure 32. ^{1}H/^{14}N Correlation Spectra of NAV at 20.0T

Recorded at 78kHz MAS frequency. Experimental (red) and simulated (blue) projections of the ^{14}N dimension are displayed. Spectrum was simulated using ^{14}N EFG tensor parameters of C_Q=3.15, η=0.15, which were the lowest RMSD parameters from 81 simulations.

4.4.2.1 Simulation Fitting of NAV ^{14}N Lineshapes

In the previous section, a two-part strategy was presented to determine EFG tensor parameters at ^{14}N sites from indirectly detected ^{14}N spectra. Firstly, bounds were placed on the value of C_Q by plotting the ^{14}N δ_Q^{iso}, according to Equation (4.2), as a function of C_Q and η, marking the observed ^{14}N δ_Q^{iso} and reading off the highest and lowest C_Q values consistent with the observed ^{14}N δ_Q^{iso}. Subsequently, with the goal of determining η in order to further restrict the range of values of C_Q consistent with the ^{14}N δ_Q^{iso}, a series of simulated spectra were fit to the experimental ^{14}N lineshapes. In order to determine applicability of these techniques on a sample with a ^{14}N site with a large (>3MHz) C_Q, and to assess the effects on the spectra of working at high magnetic fields, the process was repeated with the data acquired on NAV.

A plot of the ^{14}N δ_Q^{iso}, calculated from Equation (4.2), as a function of C_Q and η at 20.0T, with the experimental ^{14}N δ_Q^{iso} marked, is shown in Figure 33. From this plot, one can read off the upper and lower limits placed on the C_Q at the ^{14}N site in NAV by the ^{14}N δ_Q^{iso} and determine the range of C_Q magnitudes consistent with the ^{14}N δ_Q^{iso} as 2.83MHz-3.26MHz, for all values of η. This range of 430kHz is far larger than the range of ~200kHz that could be obtained for the moderate C_Q (1.0MHz-1.5MHz) sites found in histidines at 14.1T. This is since when the C_Q is large, the effects of η on the ^{14}N δ_Q^{iso} will be larger, as evident from Equation (4.2), since η and C_Q are correlated, and the contribution of η is scaled by the C_Q magnitude. Inspection of the Equations for the ^{14}N δ_Q^{iso} and quadrupolar product, χ_Q, shows that the quadrupolar product is also scaled by the Larmor frequency in the ^{14}N δ_Q^{iso} expression. This indicates that working at high B_0 fields will increase the precision with which the C_Q can be determined in this manner, which will be particularly beneficial for samples with large C_Q.

Figure 34 shows a contour plot of the RSMD of simulated spectra with respect to the experimental ^{14}N lineshape in NAV as a function of the C_Q and η used in the simulations. A comparison of the best-fit ^{14}N simulated lineshape with the experimental lineshape is seen in the projections in Figure 32. The quality of the fit is high, with a minimum RMSD=1.5, however the RMSD

minimum is broad; covering a region of C_Q magnitude of 75kHz (C_Q=3.1MHz – C_Q=3.175MHz), and range of η values of 0.32 (0 – 0.32). The range of values covered by this RMSD minimum slightly underestimates the literature values of C_Q=3.21MHz, η=0.32[22]. However, the black contour line which indicates the ^{14}N δ_Q^{iso} calculated from the measured ^{14}N shift at the centre of gravity of the ^{14}N lineshape is also underestimated by the best RMSD fit, and in fact agrees very well with the literature values. The lack of accuracy with which the simulations fit the data suggest that there are features of the experimental ^{14}N lineshape that are not accounted for in the simulations. The reason for this is unclear. It may be that the ^{14}N spectrum of this site is broadened by high mobility. One possibility is that the signal is broadened or shifted by higher-order terms (>2) in the quadrupolar Hamiltonian that are not accounted for in the simulations and not removed by MAS. For sites with low (<1.5MHz) quadrupolar couplings, it was verified that these terms did not affect the simulated lineshapes. However, this was not verified for the case where the quadrupolar coupling was large (~3MHz) and high order terms are expected to be more significant, as running the simulations with the full Hamiltonian was found to be prohibitively time consuming.

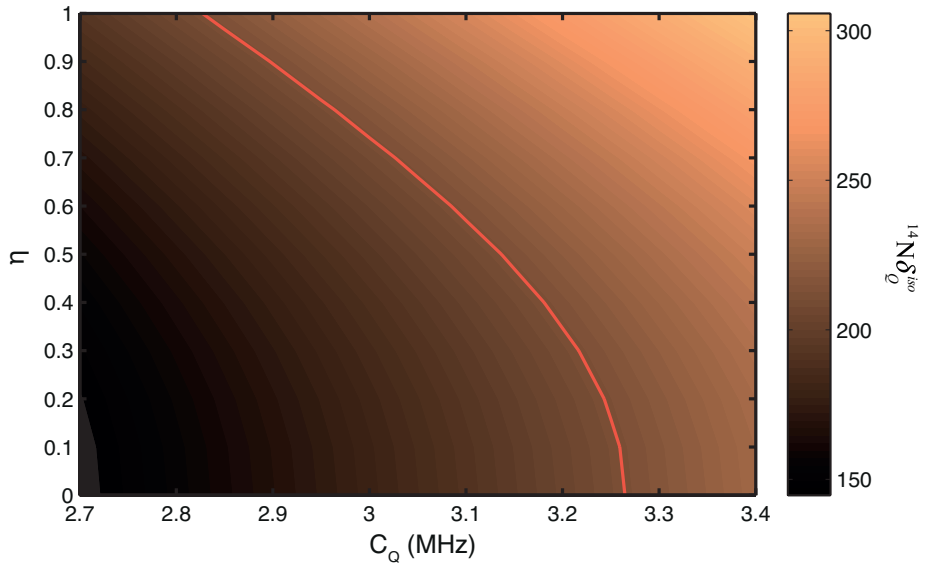

Figure 33. Plot of ^{14}N Second Order Isotropic Quadrupolar Shift at 20.0T

The ^{14}N δ_Q^{iso} is plotted as a function of C_Q and η at a field strength of B_0=14.1T, according to Equation (4.2). A red contour line is displayed for the experimental value of ^{14}N δ_Q^{iso} for the nitrogen site in NAV, of 207.3ppm

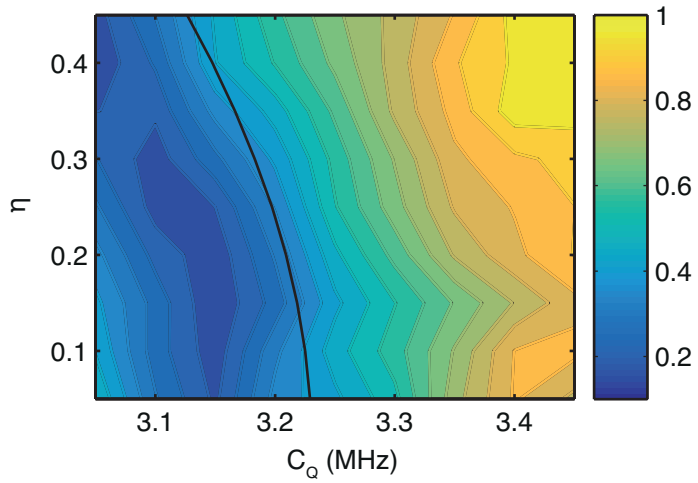

Figure 34. RMSD Plot of Simulated and Experimental ^{14}N Lineshapes of NAV

The RMSD of simulated ^{14}N lineshapes with respect to the experimental lineshape of the ^{14}N site in NAV is plotted as a function of the ^{14}N C_Q and η used in the simulation. A single black contour line marks the experimental ^{14}N δ_Q^{iso}, calculated from the ^{14}N shift-^{14}N δ_{cs}^{iso}.

4.5 Conclusion

In this Chapter, a number of developments in proton detected indirectly detected ^{14}N experiments have been demonstrated. Firstly, an RF-driven indirect ^{14}N detection pulse sequence has been presented. This method is the same as the method demonstrated in Chapter 3 for ^{13}C detection, but detected on protons to improve the sensitivity. The method employs long (100s μs) pulses of moderate RF amplitude (30-70kHz) pulses to ^{14}N to generate coherence between ^{1}H and ^{14}N, rather than using free evolution where the *J*- and RDS couplings evolve. It has been shown in this Chapter that this method is more efficient than many other HMQC-type methods presented in the literature. We observe transfer efficiencies of up to 27% on samples with moderate ^{14}N C_Q magnitude (<1.6MHz), compared to the highest reported efficiency in the literature of <10%, measured on the same compound with ultrafast (90kHz) MAS, with more typical efficiencies of pulse sequences using dipolar recoupling sequences of <5%. The type of efficiencies we report allow the acquisition of spectra with well resolved ^{14}N lineshapes on unlabelled small molecules with ^{14}N quadrupolar couplings of 1MHz – >3MHz in a matter of hours. Furthermore, we use solely fast MAS (35kHz – 78kHz) to resolve proton detected data, without sample deuteration or homonuclear decoupling.

Secondly, having acquired data of a quality sufficient for a numerical analysis for the first time, we have demonstrated an approach for determining the magnitude and asymmetry of the EFG tensor at ^{14}N sites by simulation fitting of ^{14}N SQ lineshapes. Previously, the EFG has been characterised by indirect detection only in terms of the quadrupolar product, χ_Q, which depends on both the C_Q and η. From this value, it is not possible to determine the C_Q more precisely than ±100kHz for sites with moderate C_Q values (~1MHz) at moderate B_0 fields, with the precision reducing as the C_Q increases. In this Chapter, it was shown that fitting simulations to the experimental lineshapes can, in some cases, define η to a small range of values, allowing the C_Q, and therefore the ^{14}N EFG tensor to be characterised with increased precision. In other cases, however, the simulations did not fit well to the experimental lineshapes, or the fit

was consistent with a broad range of ^{14}N EFG tensor values, and the fitting approach did not provide increased precision. There are a number of potential reasons for the shortcomings of the method. In some cases, molecular motions may have contributed to distortions in experimental lineshapes that were not accounted for in simulations, and poor resolution of sites due to fast ^{1}H T_2' or ^{14}N $T_{1\rho}$ relaxation could have broadened spectral features to the point they were difficult to fit. Other ^{14}N lineshapes may have been difficult to fit due to approximations of spin systems in the simulations not accurately reflecting the properties of the site, and there remains a question as to whether high order (>2) effects in the quadrupolar Hamiltonian, not accounted for in simulations, can affect the lineshapes of sites with large quadrupolar couplings (≥3MHz) in particular. It is envisioned that a number of approaches could help alleviate issues with the fitting of simulations. Firstly, a number of methods to increase spectral resolution could be employed. Faster MAS, partial sample deuteration and use of homonuclear decoupling sequences could be employed to increase resolution in the detected proton dimension, as well as increasing the proton T_2', thereby removing homogeneous broadening effects from strong proton-proton couplings from indirectly detected ^{14}N lineshapes. Secondly, DQ ^{14}N spectra could be acquired instead of SQ spectra, and simulated. DQ spectra are not affected by the FOQI, making them less sensitive to dynamics[77, 163], and thus may fit better to simulations where dynamics are not accounted for. While the DQ transfer is typically less efficient than that of the SQ[2, 21, 62, 67, 74, 77], with the experimental efficiencies observed for the proton detected SQ transfer, sensitivity is expected to be more than sufficient to acquire DQ ^{14}N lineshapes of adequate quality to fit simulations.

CHAPTER 5

INDIRECT DETECTION OF ^{14}N IN COMPLEX BIOMOLECULES: INDIRECT DETECTION OF ^{14}N IN GB3

5.1 Introduction

Recent decades have seen the development of biological ssNMR into a staple technique in structural biology, and ssNMR techniques for investigating the structure and dynamics of proteins in the solid-state are well established. It is an invaluable technique for studying protein samples which cannot be studied by solution-state NMR or X-ray crystallography in their native state, such as membrane proteins and amyloid fibrils[11, 165]. Typically, NMR studies on proteins rely on isotopically labelling samples with the spin-½ isotopes ^{13}C and ^{15}N. This is frequently a practical necessity in the case of ^{13}C, where the highly naturally abundant (98.9%) isotope, ^{12}C, has a spin=0 and cannot be used for NMR. The naturally abundant (99.6%) isotope of nitrogen, ^{14}N, on the other hand, has a nuclear spin=1, and can be used for NMR. The moderate quadrupole moment of Q = 20.44mbarn for ^{14}N leads to typical quadrupole couplings on the order of 1-7MHz, making it difficult to study directly by NMR[5]. In the last decade, methods for studying ^{14}N have been proposed and developed which enhance resolution and sensitivity by indirectly detecting ^{14}N via a spin-½ nuclei potentially allowing their application to complex molecules with many ^{14}N sites, as reviewed in Chapter 1. Furthermore, the quadrupolar coupling is highly sensitive to small changes in the local electrostatic environment. This makes it potentially an excellent probe of structure of nitrogen sites in a variety of complex materials such as proteins and peptides, DNA, and nanostructures. For example, at the Nδ site in the imidazole sidechain of histidine, the ^{15}N chemical shift changes by ~80ppm (~5kHz at B_0=14.1T) between protonated and protonated isomers[166]. However, the ^{14}N C_Q changes by 1.9MHz at the same nitrogen site between the two isomers[134]. Moreover, the protonation state and hydrogen bonds formed in the sidechain of histidine compounds can affect the ^{15}N chemical shift of the Nε by only~10ppm

(~600Hz at B_0=14.1T)[166, 167]. The ^{14}N C_Q, however, can differ by over 100kHz depending on the hydrogen bonding in the imidazole ring[160], which is manifested by changes in ^{14}N δ_Q^{iso} of 20ppm, or 1.2kHz, at B_0=14.1T, as shown in the previous Chapter. It has also been shown, in model polypeptides, that the modest difference in hydrogen bond length between α–helices and β–sheets is sufficient to change the ^{14}N C_Q by ~200kHz for a in a given residue[44]. Hence, the ^{14}N C_Q has the potential to act as an extremely sensitive reporter on the local structure in proteins.

However, despite its potential for investigating complex structures, the application of indirect detection ^{14}N NMR has so far not progressed further than relatively small organic molecules with <7 ^{14}N sites with moderate C_Q magnitudes (<1.5MHz)[66, 79, 168], or molecules with <3 high C_Q (~3MHz) sites[1]. A current limitation of indirectly detected ^{14}N ssNMR experiments is their limited ^{14}N/spy nucleus transfer efficiency and sensitivity which has so far prevented their application to more complex materials and large biomolecules such as proteins which contain many high C_Q amide ^{14}N sites. In Chapters 3 and 4 of this thesis, an RF-driven indirect ^{14}N detection method was introduced[4], and its increased efficiency with respect to previous HMQC-type indirect detection methods was demonstrated, as well as an analysis of the precision with which ^{14}N EFG tensor parameters, C_Q and η can be determined from the resulting spectra. In this Chapter, exploiting the efficiency enhancement of the RF-driven indirect detection method introduced in this thesis, the application of this technique is demonstrated on a model protein sample of protein GB3, the expression, purification and characterisation of which was detailed in Chapter 2. We combine this method with a number of general techniques for enhancing sensitivity in ssNMR experiments. Presented are $^{1}H/^{14}N$ correlation spectra of microcrystalline GB3 recorded with ^{1}H detection, at high field (B_0=20.0T) and fast MAS (75kHz), using paramagnetic dopants to increase sensitivity by paramagnetic relaxation enhancement (PRE) of the ^{1}H T_1 to allow fast recycling. Also presented are ^{13}C detected spectra, performed under MAS at room temperature, as well as with sensitivity enhancement afforded by dynamic nuclear polarisation (DNP) at cryogenic temperatures. The spectra recorded

allow us to perform the first analysis of observed ^{14}N shifts in a full-length protein, and characterise the ^{14}N EFG tensor magnitudes found at amide sites in GB3. It is shown that the magnitude of the quadrupolar coupling is sensitive to the H-bonding associated with secondary structural motifs in the protein, which is manifested by large, measurable changes in the second order isotropic quadrupolar shift in indirectly detected spectra.

5.2 Sensitivity Enhancement in ssNMR

In this Chapter, two general methods for increasing the sensitivity of ssNMR experiments are employed; DNP and PRE of ^{1}H T_1 to enable fast recycling. In this thesis, the focus in not on developing theoretical or methodological aspects of either. However, some aspects of the theory and application of each technique pertinent to their use in this work are iterated here for clarity. For more detailed accounts of the theory and practice of each technique, interested readers are referred to more thorough references on the subjects of DNP[169-171], and paramagnetic relaxation in NMR[172-175].

5.2.1 Dynamic Nuclear Polarisation: Theory and Background

DNP is one of a class of hyperpolarisation NMR experiments where the sensitivity of nuclear spins is increased by the transfer of magnetisation from highly polarised unpaired electrons. In the high temperature limit, the difference in population of energy levels of a spin-½ particle at thermal equilibrium, which is proportional to the available signal in an NMR experiment, can be given by the Boltzmann distribution:

$$\frac{N_{upper}}{N_{lower}} = e^{\frac{-\Delta E}{kT}} = e^{\frac{-\gamma B_0}{kT}} \cong 1 - \frac{\gamma B_0}{kT} \tag{5.1}$$

where N_{upper} and N_{lower} represent the population of nuclei in the upper and lower energy states, k is Boltzmann's constant, and T is the temperature in Kelvin. It is clear, then, that the equilibrium polarisation, and thereby the sensitivity of an NMR experiment, can be increased by reducing the temperature, increasing the magnetic field, or both. An alternative method involves polarising some species with a higher gyromagnetic ratio than atomic nuclei, such as electrons, and then transferring polarisation to the nuclear spins. This process is known as hyperpolarisation. With DNP, the polarisation of electrons is transferred to nuclear spins by microwave irradiation at or near the electron paramagnetic resonance (EPR) transition frequency. The theoretical enhancement achievable with DNP is governed by the ratio of electron and nuclear gyromagnetic ratios (γ_e / γ_n), which is ~660 for electrons and protons.

5.2.1.1 The Cross Effect

There are a number of mechanisms by which polarisation can be transferred from electrons to nuclear spins. That which has emerged as the method of choice for high field ssNMR experiments in recent years is the cross effect (CE)[176, 177]. The CE is most efficient with a biradical polarising agent where the two electrons are dipolar coupled, and have EPR transition frequencies that are separated by the nuclear Larmor frequency: $\omega_{0S_1} - \omega_{0S_2} = \omega_{0I}$. One can then view the experiment in terms of a three spin-½ system (two electrons and one nucleus) with eight energy levels, two of which are degenerate. In this case, continuous wave irradiation at either of the electron Larmor frequencies will saturate the degenerate energy levels, and the population difference between energy levels corresponding to NMR transitions is increased[170, 171, 178].

Optimal radicals for CE DNP are biradical polarising agents, where the EPR spectrum of each radical site has a small homogeneous linewidth, but experiences large inhomogeneous broadening due to *g*-anisotropy. In this situation, there is an increased chance that a given two spin packets in the inhomogeneously broadened EPR spectrum will match the condition, $\omega_{0S_1} - \omega_{0S_2} = \omega_{0I}$ [178, 179]. However, it also means that the efficiency of the transfer process decreases by ω_0^{-1} as the size of the B_0 field and the electron Larmor frequency increases, since the probability of matching this condition falls. Another consideration for a useful polarising agent is the rate of electron T_1 relaxation, and the concentration at which the radical can give a useful signal enhancement, since these variables can give rise to detrimental effects in NMR spectra due to nuclear T_2 broadening. The solubility of the radical in water is another factor to consider in the case of experiments on biological samples. In this study we have used AMUPol, a recently developed bTUrea derivative which is ideal for our study for a number of reasons[180]. It is highly water soluble, making it compatible with the buffers used for GB3 and it provides large DNP enhancements compared to other commonly used soluble CE biradicals such as TOTAPOL. The DNP enhancement from AMUPol is relatively insensitive to

MAS frequency below 15kHz, and enhancements has have been shown to be less attenuated at higher B_0 fields than for other biradical polarising agents[180].

Another experimental variable that should be considered is the temperature at which experiments are conducted. The efficiency of the CE mechanism is severely attenuated at temperatures close to ambient temperatures. This is thought to be principally due to fast nuclear T_1 relaxation at high temperatures. Indeed, it has been demonstrated that increased deuteration at exchangeable amide sites in proteins can increase DNP enhancements in a manner that correlates with decreased ^1H T_1 relaxation rates, and that DNP enhancements are larger at higher temperatures in deuterated samples with decreased relaxation rates[181]. For this reason, MAS ssNMR DNP measurements are typically carried out at cryogenic temperatures (90K – 100K) as a compromise between maximal sensitivity and the difficulty of achieving stable low temperature MAS (LT-MAS) for extended periods of time.

5.2.2 Paramagnetic Relaxation Enhancement: Theory and Background

Another obstacle to higher sensitivity in NMR experiments is the rate of nuclear T_1 relaxation that governs the optimal recycle delay in experiments. Optimal sensitivity per unit time is achieved by using a recycle delay of at least $3xT_1$ of the excited nucleus. In biological samples, typically protons are excited, even if other nuclei are detected after polarisation transfer, and typical T_1 times can be ~500ms. One can endeavour to reduce this time by increasing relaxation rates by introducing unpaired electrons into the sample.

Introducing unpaired electrons into NMR samples can have a number of effects on the NMR spectrum for the sample. As stated above, unpaired electrons have gyromagnetic ratios orders of magnitude greater than nuclei, and the electrons couple to nuclei in the sample by the hyperfine coupling[172]. These effects manifest themselves in NMR spectra in the form of paramagnetic shifts, including pseudocontact shifts (PCS), or relaxation enhancements. Paramagnetic centres with anisotropic magnetic susceptibility tensors, e.g. Co^{2+} give rise to modulations of the frequency of NMR shifts, while those with close

to isotropic magnetic susceptibility tensors, e.g. Cu^{2+}, give rise to larger PRE effects[174, 182].

Approximated expressions for the longitudinal (Γ_1) and transverse (Γ_2) nuclear PREs in the solid-state are given by[172, 174]:

$$\Gamma_1 \approx \frac{2C}{r^6}\left(\frac{3T_{1e}}{1+\omega_I{}^2T_{1e}{}^2}+\frac{7T_{1e}}{1+\omega_e{}^2T_{1e}{}^2}\right)$$
$$\Gamma_1 \approx \frac{C}{r^6}\left(4T_{1e}+\frac{3T_{1e}}{1+\omega_I{}^2T_{1e}{}^2}+\frac{13T_{1e}}{1+\omega_e{}^2T_{1e}{}^2}\right) \quad (5.2)$$

Where T_{1e} is the longitudinal relaxation at the paramagnetic centre, r is the distance between the nucleus and the paramagnetic centre, the Larmor frequencies of the electron and the nucleus are given by ω_I and ω_e, respectively, and C is a constant that depends on a number of fundamental constants. It is apparent, then, that one may choose PRE agents that will affect the nuclear T_1 and T_2 rates differently, according to the T_{1e} of the paramagnetic ion. These will be selected depending on the purpose of the PRE in the experiment. In the case of enhancing sensitivity by using paramagnetic ions, those which enhance nuclear T_1 with minimal effects on the nuclear T_2 are desired in order to maximise the potential speed of recycling, while minimising linebroadening from decreased T_2. The effects of the paramagnetic dopants introduced will also depend on their concentration in the sample. A number of different paramagnetic dopants have been utilised in order to increase the rate of nuclear T_1 relaxation with minimal effects on nuclear T_2 relaxation, to enable fast recycling with minimal linebroadening. These include Cu^{2+} containing compounds such as Cu^{2+}-EDTA[175, 183-185], and Gd^{3+} containing compounds such as Gd^{3+}-DOTA and Gd^{3+}-DTPA, which can be used in protein samples at lower concentrations than Cu^{2+}-EDTA to achieve equivalent or greater T_1 relaxation enhancements [186, 187]. In experiments in this Chapter we use Gd^{3+}-DTPA which can be used to increase nuclear T_1 rates at relatively

low concentrations (1mM) to reduce nuclear T_1 rates tenfold with minimal nuclear T_2 effects[187].

5.3 Materials and Methods

5.3.1 Solid-State NMR

All Experiments described in this Chapter were performed on samples of GB3. GB3 samples were expressed and purified identically to all GB3 samples prepared and characterised in this thesis, as described in Chapter 2.

5.3.1.1 ^{13}C Detected Experiments at Room Temperature

For ^{13}C detected experiments described in this Chapter, a U-^{13}C labelled sample was used. This was not crystallised in the manner described in Chapter 2, but prepared by lyophilisation from solution overnight. The sample was packed into a 2.5mm Bruker MAS-NMR rotor.

NMR experiments were performed on a Bruker Avance II spectrometer operating at 14.1T (600MHz ^{1}H Frequency), equipped with a triple resonance 2.5mm MAS probe modified in house to tune to ^{14}N on the third channel, tuned in triple resonance mode to ^{1}H/^{13}C/^{14}N frequencies. The sample was spun at 25kHz. The ^{13}C detected experiment shown in Figure 17 was used for indirect ^{14}N detection. ^{1}H to ^{13}C cross-polarisation was linearly ramped on the proton channel to match a 50kHz spin lock field on ^{13}C. Excitation and reconversion pulses on ^{14}N were applied for 1.75ms, with an amplitude of 50kHz. SPINAL-64[114] decoupling was employed on the proton channel at an RF field of 115kHz during ^{14}N pulses, and at 100kHz during acquisition. The recycle delay was 2.5s and 160,000 scans were acquired. The data were processed in MatNMR[126] where 30Hz linebroadening was applied and data was zerofilled to 2048 points before Fourier transform.

5.3.1.2 ^{1}H Detected Experiments at Room Temperature

Proton detected experiments were performed on a sample of unlabelled GB3. 1mM Gd^{3+}-DTPA was added to the sample before crystallisation, which was performed as described in Chapter 2, and packed into a 1.0mm JEOL MAS-NMR rotor.

NMR experiments were performed on a Bruker Avance II magnet operating at 20.0T (850MHz ^{1}H frequency), equipped with a 1mm double resonance JEOL MAS probe tuned to ^{1}H and ^{14}N Larmor frequencies. The sample was spun at 75kHz, and the VT air temperature was regulated at -50°C. The ^{1}H detected experiment shown in Figure 27 was used for ^{14}N indirect detection. Proton pulse widths were 0.8μs and 1.6μs, respectively, for π and π/2 pulses. Excitation and reconversion pulses on ^{14}N were applied for 387μs at 70kHz. A recycle delay of 0.25s was used to acquire a spectrum with 10 t_1 increments and 40,000 scans per increment, acquired using the States method[125].

Data were processed in NMRPipe[116] where 150Hz exponential linebroadening was added to the F_1 dimension and both dimensions zerofilled to 2048 points before 2D Fourier transform.

5.3.1.3 ^{13}C Detected Experiments at Cryogenic Temperatures with DNP

^{13}C detected, DNP enhanced experiments were performed on samples of U-^{13}C-GB3. In order to increase the level of deuteration at exchangeable amides in the protein to ~50%, and at least partially dueterate the solvent environment to attempt to reduce the proton T_1 relaxation rates[181], The crystallisation method was modified. GB3 was concentrated to 55mg/ml in >95% D_2O before protonated hexylene glycol was added to a 400μl aliquot of concentrated GB3 at a ratio of 1:1.4 protein:hexylene glycol to induce crystallisation. The crystals were spun down, 80% of the supernatant was removed, and a stock of 20mM AMUPol[180] in 70% d-glycerol, 30% D_2O was added. Samples were prepared which contained 2mM, and 12.5mM final concentrations of AMUPol. The final ratios of d-glycerol:D_2O:hexylene glycol in the samples were 1:10.7:4.2 and 1:1.85:0.62 in the 2mM and 12.5mM AMUPol samples, respectively. The samples were centrifuged into Bruker 3.2mm sapphire LT-MAS NMR rotors in cycles of pelleting solid protein into the rotor and removing supernatant.

All MAS-DNP experiments were performed using a Bruker Avance III spectrometer operating at 18.8T (800MHz ^{1}H Larmor frequency), with a gyrotron oscillator generating continuous wave microwaves at a frequency of 527GHz,

and a 3.2mm triple resonance MAS-DNP probe tuned to $^{1}H/^{13}C/^{14}N$ frequencies. Bearing, drive and VT gas were supplied with cooled nitrogen, to achieve sample temperatures of ~100K. The pulse sequence shown in Figure 17 was used to record $^{13}C/^{14}N$ correlation spectra of GB3, and $^{1}H/^{13}C$ CP and spin echo spectra were also recorded to assess enhancements due to DNP, and the ^{14}N transfer efficiency of the pulse sequence. Spectra of GB3 were recorded at 13.5kHz MAS. In all experiments on GB3, the $^{1}H/^{13}C$ CP contact time used was 1.75ms, with ramped ^{1}H pulse matched to a ^{13}C spin lock field of 40kHz amplitude. Excitation and reconversion pulses on ^{14}N were optimised to 1.7ms (24 rotor periods) at 30kHz, and decoupling at 100kHz using TPPM[188] was used in all cases during acquisition and t_1 periods of 2D experiments. Two-dimensional spectra were acquired in a phase sensitive manner using States-TPPI[115], with 76 increments acquired with 640 scans per increment. The recycle delay was 6s, giving a total acquisition time of 3.3 days. Data were processed in NMRpipe, where 20Hz linebroadening was applied in either dimension, before both dimensions were zerofilled to 1024 points and Fourier transformed. Spectra are referenced to solid NH_4Cl at 39.3ppm in the ^{14}N dimension, corresponding to liquid ammonia at 0ppm[127]. Carbon dimensions are referenced to DSS at 0ppm, using the downfield peak of adamantane at 40.48ppm as a secondary reference[117].

5.3.3 Calculated $^{13}C/^{14}N$ Spectra

To aid in interpretation of 2D $^{13}C/^{14}N$ spectra recorded on GB3 with DNP, 'calculated' $^{13}C/^{14}N$ spectra of GB3 were constructed using the $^{13}C/^{15}N$ NCA and NCO spectra shown in Figure 15. The ^{13}C and ^{15}N chemical shifts of each assigned peak in the NCO and NCA spectra were recorded. To the ^{15}N chemical shift of each nitrogen was added an 'expected' ^{14}N δ_Q^{iso}, according to the secondary structure at that site, which was evaluated from the CSI. The expected shifts were estimated from the known ^{14}N EFG tensor parameters of NAV, C_Q=3.21MHz, η=0.32[22], whose ^{14}N site is structurally similar to those found in the peptide backbone of proteins. This was assumed to be the magnitude of the ^{14}N C_Q if the site was expected to be a β-sheet. According to

the study of Fukazawa *et al*.[44], 200kHz was added to the β-sheet C_Q of 3.21MHz if the site was expected to be in an α-helix. Those sites predicted to be a random coil were assigned a C_Q of 130kHz less than that of the β-sheet. It was assumed that the lack of hydrogen bonding at amide sites in a random coil would reduce the asymmetry of the charge distribution at the ^{14}N nucleus, leading to a smaller value of the C_Q magnitude. We also observed a number of resonances on the high field edge of the experimental ^{14}N spectrum which appeared to be too low to be consistent with expected values for α-helix or β-sheet signals, that we thus fit the random coil value to match. At all sites, η was kept constant and assumed to be η=0.3, which is within η=±0.01 of amide sites found in a number of peptides and acetylated amino acids, such as NAV[22], triglycine[132], AGG and AAG[138]. From the equation for the ^{14}N δ_Q^{iso} given in Equation (1.49), the expected ^{14}N δ_Q^{iso} at B_0=18.8T for each type of secondary structure was calculated to be ^{14}N δ_Q^{iso}=261 ppm for an α-helix (C_Q=3.41MHz, η=0.3), ^{14}N δ_Q^{iso}=237 ppm for a β-sheet (C_Q=3.21MHz, η=0.3), and ^{14}N δ_Q^{iso}=219ppm for an random coil (C_Q=3.08MHz, η=0.3). Spectra were calculated in the time domain using ^{13}C shifts and ^{15}N shifts with secondary structure dependent 'expected' ^{14}N shifts added. A sinebell squared window function was applied in the time domain to the generated signals to add 150Hz linebroadening in either dimension to simulate the effect of cryogenic temperatures on ^{13}C linewidths, and cryogenic temperatures and second order broadening effects on ^{14}N linewidths. Data were zerofilled to 1024 points and Fourier transformed in both dimensions.

5.6 Results

Indirect ^{14}N detection studies have been carried out on GB3 under a range of conditions in order to measure ^{14}N quadrupolar couplings in the protein.

5.6.1 ^{13}C Detected Experiments at Ambient Temperature

A 1D ^{14}N filtered ^{13}C spectrum of lyophilised ^{13}C-GB3 is shown in Figure 35. This spectrum required 160,000 scans, almost 4 days of signal averaging, to acquire. The efficiency at the most intense peak in the spectrum with respect to a ^{13}C spin echo on the same sample under the same conditions was <0.00001%, which is over five orders of magnitude less than the 4% efficiency observed with ^{13}C detection in powdered peptides such as triglycine, demonstrated in Chapter 3. A sample of lyophilised GB3 was used instead of a microcrystalline sample in this case since a comparable sample of microcrystalline protein was found not to yield any observable signal in the same timeframe under the same experimental conditions. The increased ^{14}N transfer efficiency of the lyophilised sample we attribute to the reduced dynamics of the sample compared to the microcrystalline sample. Protein dynamics on the ms timescale may cause interference with ^{14}N RF fields of kHz amplitude since they occur on a similar timescale. Interestingly, the most intense peaks in the ^{14}N filtered spectrum appear at ^{13}C chemical shifts between 42.0ppm and 43.7ppm, and are thus assigned to the Cε sites of lysine residues, adjacent to the amine ^{14}N in the sidechain. A far smaller peak is visible in at ~51.2ppm, which is attributed to $^{13}C\alpha$ sites adjacent to ^{14}N amide sites in the protein backbone. This lack of sensitivity at the Cα sites may be due to the ^{14}N filtering experiment being more efficient at amine sites over amide sites, as was generally found in the experiments on model compounds in Chapters 3 and 4. Alternatively, it could be due to the high abundance of lysine in this protein (7/56 residues) and chemical shift degeneracy of the lysine $^{13}C\varepsilon$ resonances, which is more pronounced in this inhomogeneously broadened lyophilised sample over microcrystalline samples, leading to all the lysine signals falling on top of one another adding to the intensity in this region.

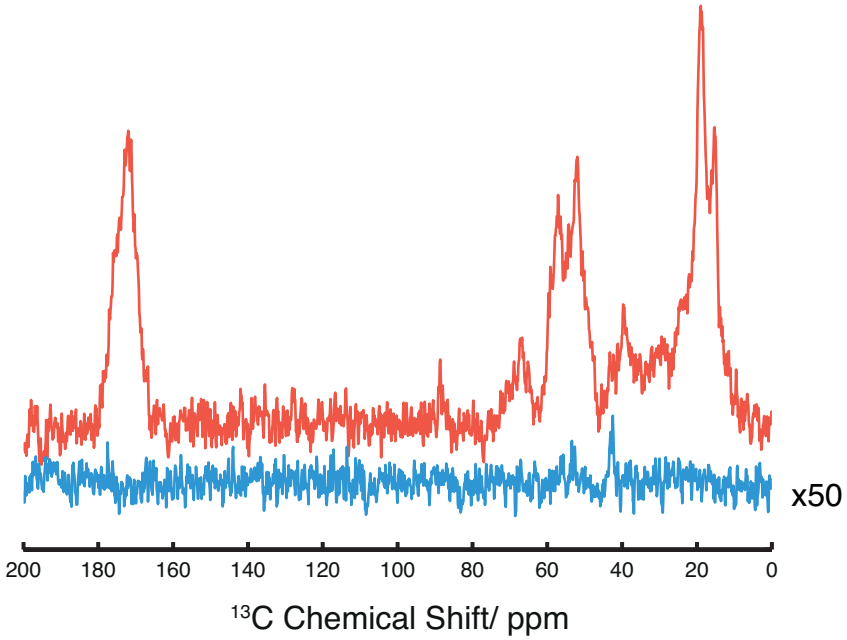

Figure 35. 1D ^{14}N Filtered ^{13}C Spectrum and Spin Echo Spectrum of Lyophilised U-^{13}C GB3

Spectra Acquired on Lyophilised ^{13}C-GB3 at room temperature at 14.1T with 25kHz MAS. Red spectrum is a ^{13}C spin echo spectrum, acquired with 16 scans, where echo periods are the same length as the ^{14}N pulses in the ^{14}N-filtered spectrum. Blue spectrum is the ^{14}N filtered ^{13}C spectrum (i.e. a single slice of the indirect detection sequence, where t_1=0), with 160,000 scans acquired.

5.6.2 ^{1}H Detected Experiments at Ambient Temperature

It was found that the ^{13}C detected experiment was not efficient enough to record 2D ^{13}C/^{14}N spectra on protein samples within a reasonable timeframe, months would be required. The ^{1}H detected version of the same experiment is intrinsically more sensitive than the ^{13}C detected experiment, owing to the gyromagnetic ratio of ^{1}H being ~4x greater than that of ^{13}C. In addition to the sensitivity enhancement from proton detection, in Chapter 4 it was demonstrated that transfer efficiencies of 9% were achievable on a sample (NAV) with a ^{14}N quadrupolar coupling of ~3MHz using ^{1}H detection, which is over double the 4% achieved on a sample (triglycine) with a similar C_Q with the ^{13}C detected experiment in Chapter 3. The C_Qs of ~3MHz at ^{14}N sites in samples of NAV and triglycine are expected to be similar to those found in proteins due to the structural similarity of the ^{14}N sites in these small molecules to the amide ^{14}N sites in the protein backbone. Hence, the ^{1}H detected experiment was employed to try and increase the sensitivity and efficiency of the experiment in order to record a 2D ^{1}H/^{14}N correlation spectrum and characterise ^{14}N sites in GB3. In addition, a higher field instrument (850MHz) was used to further increase sensitivity, fast MAS (75kHz) was employed to increase ^{1}H coherence lifetimes providing additional resolution. The sample was also doped with a paramagnetic relaxation agent (1mM Gd^{3+}-DTPA) in order to increase the rate of T_1 relaxation in the sample, enabling faster recycling, and a large increase in terms of sensitivity per unit time.

Under these conditions, it was possible to record the 2D spectrum shown in Figure 36 in 28 hours, with 10 t_1 increments with 40,000 scans per increment. The spectrum is dominated by intense signals that are assignable to the protonated solvents used to prepare the protein samples. Those resonances at ~1.0ppm and 3.9ppm are from hexylene glycol, and those at 4.7ppm and 5.2ppm are assigned to bulk and crystalline water, respectively. These artefacts can be removed in future work by preparing samples with deuterated samples or using pulsed solvent suppression techniques[189]. A block of resonances are observed, at ^{1}H chemical shifts between 6ppm-12ppm, that originate from amide protons correlated with amide ^{14}N spins in the protein backbone. These

are centred at a ^{14}N shift of ~340ppm. As observed in ^{15}N spectra of GB3 (Figure 15), the average ^{14}N δ_{cs}^{iso} of amide sites in GB3 is ~120ppm. One can calculate a rough average of the ^{14}N δ_{Q}^{iso} at amide sites as 340ppm-120ppm=220ppm. According to Equation (1.49), this ^{14}N δ_{Q}^{iso} value is consistent with ^{14}N C_Q magnitudes of between 2.89MHz – 3.35MHz for all values of η, which is in agreement with the values that are observed in NAV[22] and tripeptides such as triglycine[132], AAG and AGG[138] of 3.0MHz – 3.21MHz. Unfortunately it is not possible to be more specific about the range of ^{14}N C_Q magnitudes found in the protein from this spectrum for the primary reason that the FID in the indirect dimension is truncated, hence, ^{14}N resonances are broadened and one cannot precisely measure the boundaries of the observed ^{14}N shifts. Repeating this experiment with more points in the indirect dimension would allow a more precise determination of the range of ^{14}N δ_{Q}^{iso} values, and therefore, the range of ^{14}N C_Q magnitudes. Furthermore, the ^{1}H dimension is not completely resolved, and there is significant overlap of resonances. This precludes any potential for the assignment of any sites specifically, which would allow more accurate determination of the ^{14}N δ_{Q}^{iso} from a known ^{14}N δ_{cs}^{iso}, rather than an estimate based on the average observed shift in the protein.

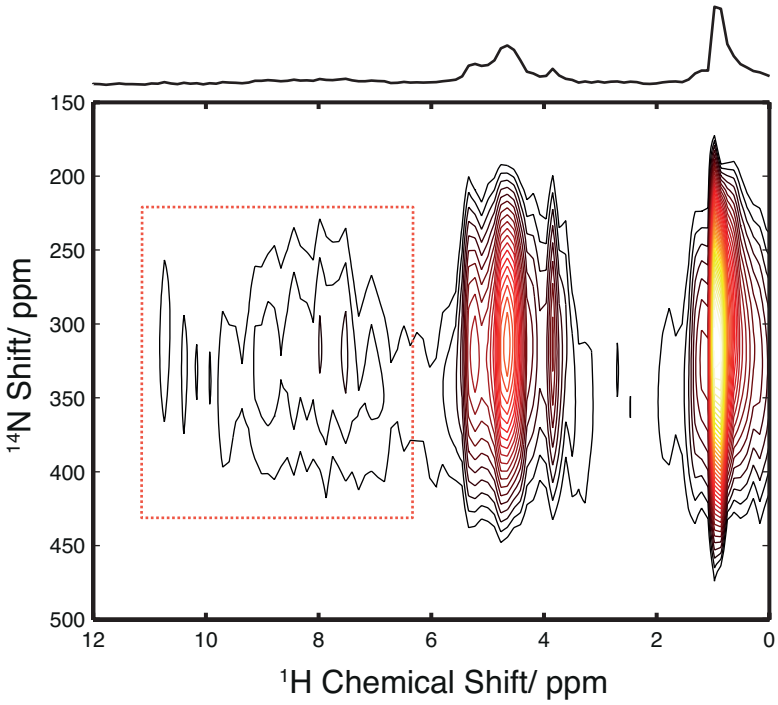

Figure 36. ^{1}H/^{14}N Correlation Spectrum of Microcrystalline GB3

2D ^{1}H/^{14}N spectrum recorded at B_0=20.0T, 75kHz MAS on a sample of unlabelled GB3 containing 1mM Gd^{3+}-DTPA. Spectrum acquired with 10 t_1 increments with 40,000 scans per increment and a recycle delay of 0.25s. The red box indicates the expected location of amide resonances. The intense peaks at ~1ppm, 3.9ppm, 4.7ppm, 5.2ppm arise from the solvents hexylene glycol and water.

5.6.3 ^{13}C Detected Experiments at Cryogenic Temperatures with DNP

5.6.3.1 Sensitivity of Cryo-DNP Experiments

Further sensitivity enhanced spectra were recorded at cryogenic temperatures (100K), with sensitivity enhancement from DNP. DNP experiments were ^{13}C detected rather ^{1}H detected. This was due to the limitation of the sample spinning speed of DNP probes (currently only commercially available with a 3.2mm rotor size, MAS limited to ~15kHz), which is insufficient to record high resolution spectra of protons. Additionally, ssNMR experiments at cryogenic temperatures suffer from inhomogeneous broadening due to a lack of mobility[92], which combined with the homogeneous broadening effects of the low chemical shift dispersion and strong homonuclear dipolar coupling network of protons, makes acquiring resolved ^{1}H spectra at cryogenic temperatures extremely challenging.

Figure 37 shows the ^{1}H/^{13}C CP spectra of GB3 samples containing 12.5mM and 2mM AMUPol, recorded at 100K, B_0=18.8T with 13.5kHz MAS, with and without CW microwaves to demonstrate the signal enhancement from DNP at each radical concentration. The signal enhancement factors, ε, observed with DNP for the samples containing 12.5mM AMUPol and 2mM AMUPol were ε=18x and ε =3x, respectively, calculated according to:

$$\varepsilon = \frac{S_{on}}{S_{off}} \tag{5.3}$$

Where S_{on} and S_{off} are the intensities of the signal acquired respectively with and without microwave irradiation at the same temperature (100K). Despite the larger signal enhancement due to DNP from the sample containing 12.5mM AMUPol over that of the sample containing 2mM AMUPol, it was found that the efficiency of the ^{13}C/^{14}N transfer was lower in the sample containing a higher concentration of radical. This is demonstrated in Figure 38, where the intensity of the ^{14}N filtered ^{13}C spectrum is compared with a ^{13}C spin echo under identical conditions for either sample. The efficiency of the ^{14}N filtered experiment,

expressed as a percentage of the ^{13}C spin echo spectrum, is calculated as 0.0006% for the sample with 12.5mM AMUPol and 0.007% for the sample containing 2mM AMUPol. Therefore, absolute sensitivity of the ^{14}N detection experiment was actually greater for the sample with a lower concentration of radical, despite a lower ε. The experimental ^{14}N transfer efficiencies observed here allowed for the acquisition of the 1D ^{14}N filtered spectra shown in Figure 38 in 1.5 hours for the 12.5mM AMUPol sample, with 896 scans, and 13 minutes for the 2mM AMUPol sample with 128 scans. The 2mM AMUPol sample was used the record subsequent 2D spectra. The reduction in intensity of the ^{14}N filtered spectrum of the sample containing higher radical concentration cannot be simply attributed to the increased radical concentration causing faster T_2 relaxation of ^{13}C. This is because the efficiency of the ^{14}N transfer is expressed as a percentage of the ^{13}C spin echo signal, meaning that radical quenching of ^{13}C T_2 is present in both ^{14}N filtered and spin echo spectra and is therefore accounted for in the calculation of transfer efficiency. Furthermore, the ^{13}C coherence lifetime of the high radical concentration sample is not reduced to a point where a spin echo long enough to insert ^{14}N pulses long enough to generate $^{13}C/^{14}N$ coherence cannot be performed. The 1.7ms echo time in Figure 38 is sufficient time for $^{13}C/^{14}N$ coherence to build up under ^{14}N pulses, and in both samples a spin echo with 1.7ms delay periods could be acquired. Unfortunately, the ^{13}C coherence lifetime of each sample was not measured, so this observation regarding ^{13}C coherence lifetime cannot be quantified, however, it seems that the AMUPol radical negatively affected ^{14}N $T_{1\rho}$ relaxation directly at higher concentrations, necessitating the use of moderate radical concentrations with reduced ε factors.

Another point of note from the calculated ^{14}N transfer efficiencies, which are independent of sensitivity enhancements from DNP, which affect both the spin echo and ^{14}N filtered experiments, is that the ^{13}C detected ^{14}N filtered experiment performs far more efficiently at cryogenic temperatures than at room temperature, where the efficiency was less than 0.00001% (Section 5.6.1), and 4 days of signal averaging was required to observe signals. The efficiency is also independent of the increased sensitivity afforded by cryogenic temperatures,

which increase the Boltzmann equilibrium polarisation, according to Equation (5.1), by a factor of 3 between 300K and 100K, and affects both the CP and spin echo spectra. The gain in transfer efficiency at low temperature, rather than sensitivity, we attribute to a lack of dynamics in the GB3 sample at low temperature. It has been shown by NMR relaxation studies on hydrated microcrystals of the homologous protein GB1, that at temperatures below ~200K, motions of the protein backbone on the ms timescale are arrested[92]. It is possible that the increased ^{14}N transfer efficiencies that we observe in this cryogenic temperature regime are due to reduced protein backbone dynamics, that at higher temperatures are either occurring on the on the timescale of the amplitude of the applied ^{14}N RF field (kHz), and interfering with it, or are broadening resonances beyond detection by dynamics induced quadrupolar relaxation[163, 164]. This observation is in accordance with our earlier findings on a lyophilised sample of GB3 at room temperature (Section 5.6.1). The lyophilised sample of GB3 was employed once it was established that in a microcrystalline sample of GB3, no signal was observed under the same conditions as the experiment of Figure 35. In this dehydrated sample, the backbone and sidechain dynamics are expected to be reduced compared to the hydrated microcrystalline form.

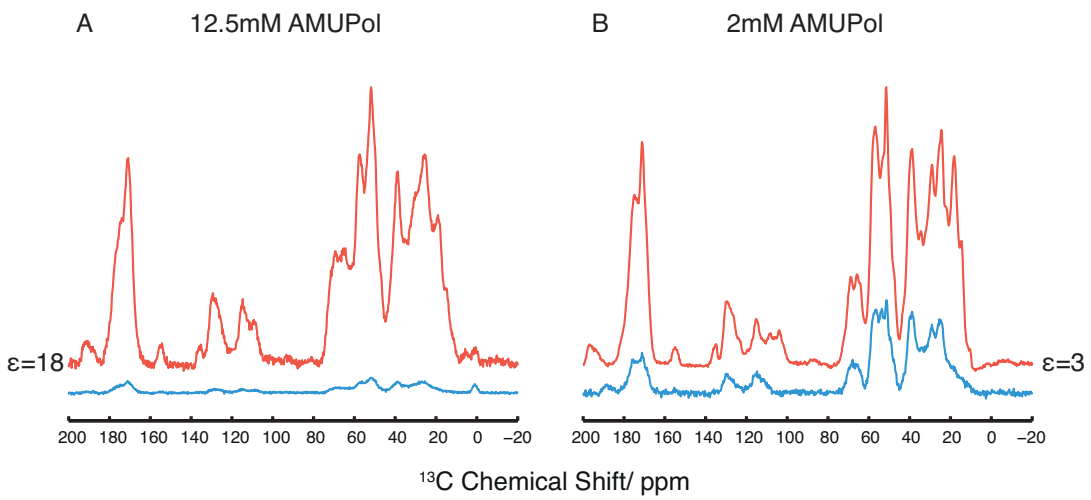

Figure 37. DNP Enhancements on Microcrystalline GB3

$^{1}H/^{13}C$ CP spectra of samples GB3 containing (A) 12.5mM and (B) 2mM AMUPol. Blue traces are without microwave irradiation, and red traces with microwave irradiation. Enhancement factors of ε=18 and ε=3 are observed in the 12.5mM and 2mM AMUPol samples, respectively. Note that in (A) and (B) the spectra are normalised to the maximum intensity in the respective DNP enhanced spectra; intensities are not comparable across the two samples.

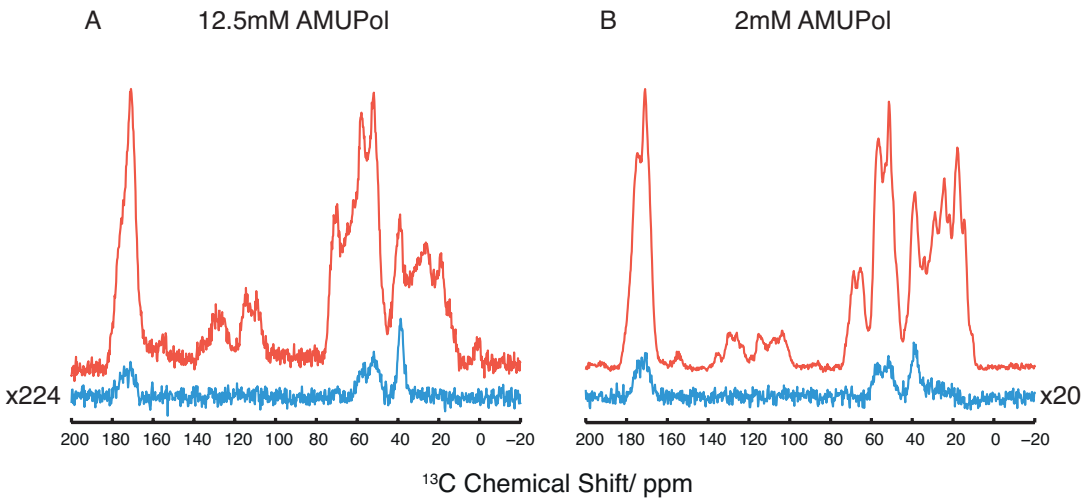

Figure 38. ^{14}N Transfer Efficiencies on GB3 with DNP

^{13}C spin echo (red) and ^{14}N filtered (blue) spectra of samples GB3 containing (A) 12.5mM and (B) 2mM AMUPol. Spin echo delays and ^{14}N pulse widths in all spectra are 1.7ms. Note that in (A) and (B) the spectra are normalised to the maximum intensity in the respective spin echo spectra; intensities are not comparable across the two samples.

5.6.3.2 $^{13}C/^{14}N$ Correlation Spectra of U-^{13}C Microcrystalline GB3

Shown in Figure 39 are the carbonyl and aliphatic regions of the $^{13}C/^{14}N$ correlation spectrum recorded on a sample microcrystalline GB3 containing 2mM AMUPol. In the aliphatic region, a region of high intensity corresponding to a number of unresolved resonances is observed at 38.5ppm-46.0ppm in the ^{13}C dimension and between 280ppm-316ppm in the ^{14}N dimension. These are assignable to the amine groups in lysine residues in GB3 on the basis of their ^{13}C chemical shifts and their reduced ^{14}N shifts compared to the amide sites in the protein backbone. The high relative peak intensity of these sites can be attributed to the fact that there are 7 lysine residues in the protein, whose $^{13}C\varepsilon$ chemical shifts are not resolved due to the inhomogeneous broadening due to structural inhomogeneity in the sample at this temperature, and whose ^{14}N shifts are also not resolvable.

All other resonances are observed at ^{13}C shifts of between 53.0ppm-64.5ppm and 170.5ppm-181.5ppm, meaning that they are attributable to Cα and CO nuclei, respectively, in the protein backbone. In the ^{14}N dimension, these resonances are correlated with ^{14}N nuclei with ^{14}N shifts of between 320ppm-420ppm. It is immediately apparent, then, that the distribution of resonances in the ^{14}N dimension is greatly increased over that of the ^{15}N dimension in the same molecule. In the $^{13}C/^{15}N$ spectra of microcrystalline GB3, shown in Figure 15, amide ^{15}N shifts, which are solely a function of the nitrogen δ_{cs}^{iso}, are distributed over <30ppm, from 102.2ppm –131.8ppm. The shifts in the ^{14}N spectrum are distributed over ~100ppm. Importantly, this additional spectral range in the ^{14}N dimension of GB3 as compared to the ^{15}N dimension is a function of the ^{14}N δ_Q^{iso}, which in turn depends on the local electronic environment at the ^{14}N site, and can inform on the magnitude of the quadrupolar couplings at ^{14}N amide sites in the protein. One can roughly estimate the magnitude and range of quadrupolar couplings at amide sites in GB3 by subtracting the centre ^{15}N shift of the ^{15}N spectrum of GB3 (120ppm) from the centre of the shift in the shift range in the ^{14}N dimension (370ppm), to give a rough "average" ^{14}N δ_Q^{iso} of 250ppm. Using Equation (1.49), without any

knowledge of η, one can determine that at a field of 18.8T, this corresponds to a ^{14}N C_Q of 2.89MHz – 3.34MHz. This value is consistent with the values for ^{14}N C_Q magnitudes expected in proteins from experiments on amide containing model samples and peptides carried out in this work (triglycine and NAV, C_Q=3.0MHz and 3.2MHz, respectively) and others[22]. It is also in very good agreement with the estimate of ^{14}N C_Q magnitudes made from the ^{14}N shifts of GB3 observed in the ^{1}H detected spectrum. Similarly, a rough estimate of the variation in the magnitude of the C_Q at ^{14}N amide sites in the protein can be made from the variation in ^{14}N δ_Q^{iso} shifts of 100ppm at this field, which could account for a variation in ^{14}N C_Q magnitude of ~500kHz at amide sites in the protein if all values of η are considered. This finding would easily allow for the ~200kHz variation in ^{14}N C_Q magnitude observed between identical amino acid residues in different secondary structures in model peptides[44].

To assess whether the distribution of observed ^{14}N shifts in GB3 were consistent with the C_Q values of model peptides, and the differences expected for residues in different secondary structures, the experimental spectrum was compared with a 'calculated' $^{13}C/^{14}N$ spectrum. The carbonyl and aliphatic regions of the calculated and experimental spectra are shown in Figure 39. The calculated spectrum was made by taking the $^{13}C/^{15}N$ NCO and NCA spectra shown in Figure 15, and adding a calculated ^{14}N δ_Q^{iso} to each amide ^{15}N resonance according to the secondary structure of the reside. As demonstrated in Figure 39, the distribution in ^{14}N shifts observed experimentally may be accurately mimicked in the calculated spectrum, using a range of ^{14}N EFG tensor magnitudes of 320kHz. The C_Q magnitudes in the calculated spectrum were defined based on the observed ^{14}N C_Q magnitudes in model peptide samples, and distributed according to the magnitude of expected differences between ^{14}N C_Q magnitudes as a function of the secondary structure of the protein, according to [44]. This provides the first estimate of the magnitude and variation of ^{14}N quadrupolar couplings that are observed at amide ^{14}N sites in proteins. Furthermore, comparison of experimental and 'calculated' spectra suggests that the magnitude and distribution of ^{14}N shifts observed is consistent

with that predicted on the basis of those found in model peptide samples, and the expected variation due to the protein secondary structure.

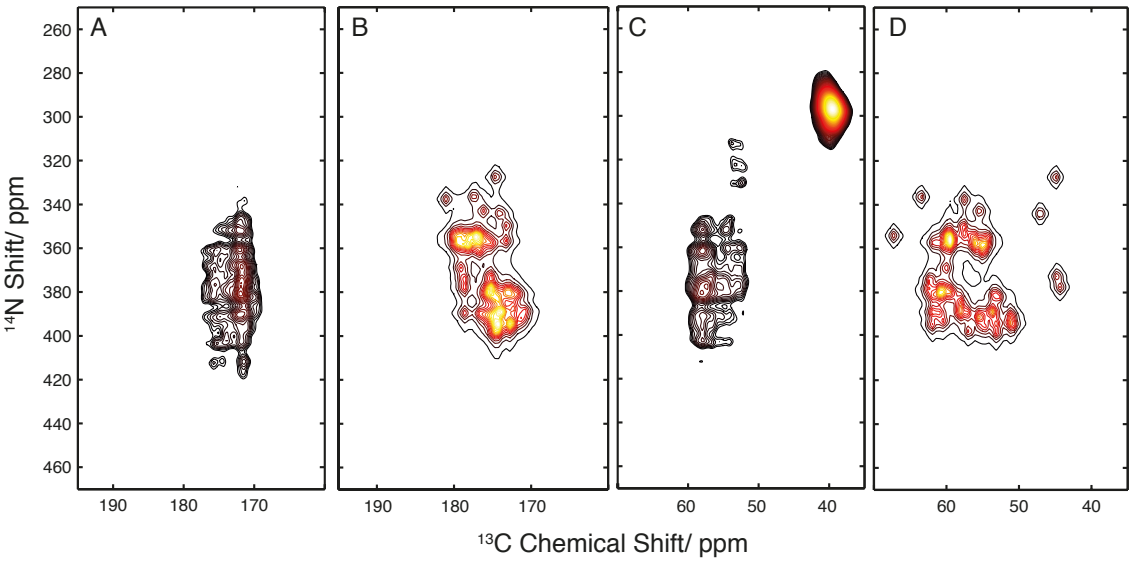

Figure 39. Experimental and 'Calculated' 2D ^{13}C/^{14}N Correlation Spectra of Microcrystalline GB3 With DNP

Experimental 2D ^{13}C/^{14}N spectrum of microcrystalline GB3 containing 2mM AMUPol is shown in panels (A) and (C), which are carbonyl and aliphatic regions, respectively. 'Calculated' spectra for the same regions are shown in panels (B), carbonyl, and (D), aliphatic. Experimental spectrum acquired at 18.8T, 100K, 13.5kHz MAS with DNP enhancement.

5.7 Discussion

In this Chapter, results of indirect detection of ^{14}N detection experiments in GB3 via both ^{13}C and ^{1}H were presented, which allowed an analysis of the ^{14}N EFG tensor parameters in this protein. We find that in GB3, there is a distribution of ^{14}N EFG tensor C_Q magnitudes of over 300kHz. This distribution is consistent with expected changes in the C_Q magnitude as a function of the secondary structure, and gives rise to well dispersed ^{14}N spectra spread over 100ppm. This is the first observation and analysis of ^{14}N ssNMR spectra in a protein, and a first insight into the properties of the ^{14}N C_Q tensors at amide sites in the protein backbone. It provides a first indication that it is possible to determine ^{14}N C_Q tensor parameters from indirect detection in proteins; and that their determination has the potential to provide valuable structural information at nitrogen sites in proteins. Unfortunately, the lack of resolution in the directly detected dimensions of both ^{1}H and $^{13}C/^{14}N$ correlation spectra precluded the assignment of any individual residues in GB3 with these spectra. Assignment of sites would allow a site-specific analysis of the ^{14}N EFG tensor parameters found in this protein. In the proton detected spectrum, proton resolution was low due to strong homonuclear $^{1}H/^{1}H$ couplings in the fully protonated protein, and the use of protonated solvents giving rise to high intensity artefacts that overlapped with protein signals. In ^{13}C detected 2D spectra with DNP, resolution was low due to the inhomogeneous broadening of protein resonances at cryogenic temperatures due to low mobility of the protein. These resolution problems are present generally in proton detected experiments of large biomolecules, and experiments at cryogenic temperatures, and a number of strategies could be employed to improve resolution.

Proton detected experiments on ^{14}N at room temperature have the advantage over cryo-DNP experiments that they do not suffer from inhomogeneous broadening due to structural inhomogeneity. However, due to the low chemical shift dispersion of protons, and relatively large homogeneous linewidth due to strong homonuclear couplings, the resolution afforded by the protonated GB3 sample used here was low. There are a number of experimental strategies that can be employed to increase ^{1}H resolution in these

experiments in future. Perdeuteration of proteins and back exchange with protons at amide sites can significantly enhance proton resolution in proteins[144, 154, 156]. Spinning at 60kHz, perdueterated and 100% back exchanged proteins can have tenfold lower ^{1}H linewidths than protonated samples, and longer coherence lifetimes[190]. Alternatively, faster MAS could be employed to reduce the homogeneous proton linewidth and increase the coherence lifetime[154, 156]. Performing experiments at higher magnetic fields can also increase resolution by improving chemical shift dispersion. These three approaches could be used individually or in combination to increase resolution in ^{1}H detected experiments on ^{14}N. Preparation of samples with deuterated solvents would also be advantageous; removing intense artefacts from the ^{1}H detected spectra.

Tackling resolution in ^{13}C detected cryo-DNP spectra is more challenging, owing to the inhomogeneous nature of the temperature dependent linebroadening that gives rise to poor resolution. One viable solution to this problem would be to selectively label protein samples, in order to selectively detect few ^{14}N sites and analyse them specifically. One could express GB3 with uniform ^{13}C labelling, and near uniform ^{15}N labelling with certain amino acids not labelled with ^{15}N, or one could simply sparsely label with ^{13}C with natural abundance nitrogen in order to observe only ^{14}N sites adjacent to ^{13}C. Alternatively, higher dimensional experiments (3D, 4D) can be used to resolve specific sites in cryo-DNP spectra[191]. Three dimensional experiments involving ^{14}N have been applied before on small molecules, to record ^{1}H/^{13}C/^{14}N correlation spectra[72], and recently to record ^{1}H/^{14}N/^{14}N experiments[161]. For the cryogenic DNP experiments discussed here, a ^{13}C/^{13}C/^{14}N correlation experiment in the manner of a ^{15}N NCOCX with ^{13}C/^{13}C PDSD or DARR mixing would seem the optimal method for increasing resolution. Recording the relatively poorly dispersed (~20ppm) carbonyl region in the indirect dimension would allow one to keep that indirect dimension narrow, and since both ^{13}C and ^{14}N linewidths are inhomogeneously broadened at cryogenic temperatures, both indirect ^{14}N and ^{13}C dimensions could be evolved for modest periods of 3-4ms. With the S/N observed with the DNP experiments described in this Chapter, it

would be possible to record such a spectrum on a sample of microcrystalline GB3 in less than a week. Higher dimensional experiments could also be considered for ^{1}H detected spectra at room temperature.

Another issue with the ^{14}N indirect detection experiments at room temperature presented in this Chapter was one of sensitivity, caused by the apparent detrimental effects on ^{14}N transfer efficiency caused by protein dynamics. Detecting ^{13}C a sample of of U-^{13}C microcrystalline GB3 at ambient temperature, we were unable to record ^{14}N filtered spectra after 4 days of signal averaging. Only after moving to a lyophilised sample of GB3, expected to be less dynamic, were we able to acquire 1D filtered ^{14}N spectra with poor S/N in the same timeframe. Furthermore, performing experiments at a cryogenic temperature where slow (ms) backbone dynamics are expected to be entirely frozen out[92], ^{14}N transfer efficiencies increased orders of magnitude. This effect is thought to be due to protein dynamics on the ms timescale interfering with applied ^{14}N RF fields with amplitudes on the same timescale. This effect ought to be investigated further, with a more rigorous investigation of ^{14}N SQ transfer efficiency as a function of temperature. It may also mean that greater transfer efficiencies could be observed on protein samples that are intrinsically more rigid, such as amyloid fibrils, and it seems other samples should also be investigated. Alternatively, the DQ ^{14}N spectrum, which should be less sensitive to dynamics[77, 163, 164], could be acquired in order to circumvent this issue. This may provide a route to increasing the ^{14}N transfer efficiency at ambient temperatures, which would be highly desirable, since inhomogeneous linebroadening should not be an issue.

A way that the transfer efficiency of ^{14}N experiments in this Chapter could be increased is by increased RF amplitudes on ^{14}N. In particular, ^{13}C detected DNP experiments were performed using sub-optimal power levels, according to the simulations of Figure 25, of 30kHz. However, this was the highest power at which we could reasonably apply 1.7ms pulses on the ^{14}N channel with. If probe developments allowed application of pulses on the ^{14}N channel of 50kHz – 60kHz, the transfer efficiency could potentially be doubled. While these simulations were performed for ^{13}C detection, not ^{1}H detection, they do also

suggest that there may be a regime where very high RF amplitudes on ^{14}N (≥100kHz) could increase transfer efficiencies. Sub-1mm rotor size probes may be able to deliver this type of power and efficiencies in this experimental regime could be investigated.

5.7 Conclusion

In this Chapter, the application of the novel pulse sequences for indirect detection of ^{14}N via ^{13}C and ^{1}H, introduced in Chapters 3 and 4, have been applied to samples of the 56 amino acid residue protein GB3. Firstly, the ^{13}C detected pulse sequence was applied to a sample of lyophilised protein to record a 1D ^{14}N filtered ^{13}C spectrum. The efficiency was found to be low, yet it was determined that this was likely to be due to dynamics in the protein, since the experiment was more efficient as applied to a sample of lyophilised protein, rather than microcrystalline material. Further evidence for this observation is provided by increased ^{14}N transfer efficiencies at cryogenic temperatures. Exploiting the increased sensitivity afforded by ^{1}H detection, high magnetic fields and fast-recycling enabled by paramagnetic doping, a 2D $^{1}H/^{14}N$ spectrum was recorded on GB3 which indicated that the ^{14}N C_Q at amide sites in the protein are similar in magnitude to those found in short peptides and acetylated amino acids. Exploiting the sensitivity enhancements afforded by cryogenic temperatures, which increases equilibrium polarisation and reduces dynamics in the sample, and DNP, 2D $^{13}C/^{14}N$ spectra have been recorded on U-^{13}C microcrystalline GB3. These are the first examples of ^{14}N NMR spectra recorded from a full length protein, and permitted an initial analysis of the nature of ^{14}N EFG tensors at the ubiquitous amide sites in the protein backbone. It was found that at amide ^{14}N sites, typical ^{14}N C_Q magnitudes of ~3.2MHz are observed, consistent with studies on model samples. The range of ^{14}N C_Q magnitudes at ^{14}N sites in GB3 can be estimated to be ~300-500kHz, depending on the asymmetry of the ^{14}N EFG tensors, and is consistent with previous studies that indicate changes in C_Q magnitude of ~200kHz between residues in α-helices and β-sheets. This span of ^{14}N C_Q magnitudes gives rise to a dispersion of ^{14}N peaks over ~100ppm at 18.8T, meaning that there is excellent potential for site specific resolution of ^{14}N sites in proteins. While the experimental methods employed here (proton detection, cryo-DNP) suffered from poor resolution in the direct dimension, we envision that strategies to improve resolution will pave the way to site specific resolution, and a sensitive, novel method to investigate structure and dynamics in proteins.

CHAPTER 6

CONCLUSIONS AND PERSPECTIVE

In this thesis, a number of advancements have been presented in the development of ssNMR methodology for recording indirect ^{14}N spectra in organic and biological molecules. These may be split into three categories. Firstly, a novel method for recording 2D correlation spectra between spin-½ nuclei and ^{14}N has been introduced, with a number of benefits over previous HMQC type experiments. Secondly, we have demonstrated advances in the analysis of indirectly detected ^{14}N lineshapes for determining ^{14}N EFG tensors. Thirdly, we present, to our knowledge, the first ^{14}N ssNMR spectra of a full-length protein, which allows us to analyse the characteristics of the EFG tensor ^{14}N sites in proteins for the first time. In the following, the conclusions that are drawn from these three areas of research are presented, before future directions for the research project are discussed.

6.1 A Novel Method for Indirect Detection of ^{14}N

The primary goal of this work was to apply ^{14}N indirect detection ssNMR to a protein in order to characterise ^{14}N EFG tensors at the abundant ^{14}N sites in proteins. In order to achieve this, issues regarding the sensitivity and robustness of existing ^{14}N indirect detection methods were addressed. A novel method for indirect detection of ^{14}N was developed that unlike previous techniques, does not rely on delays for free evolution[1, 2, 62, 67], or the application of traditional dipolar recoupling sequences such as rotary resonance recoupling[63, 64] or symmetry based recoupling sequences[66, 78, 79] for generating coherence between ^{14}N and spin-½ nuclei. Instead, delays are omitted, and coherence is generated directly under the action of long (100s μs - ms) pulses of moderate amplitude (≥30kHz) RF irradiation on the ^{14}N channel only[4]. This method was used to detect ^{14}N correlation spectra with ^{13}C and ^{1}H nuclei on model

compounds of amino acids, acetylated amino acids, and peptides. It was shown that this method had a number of benefits over previous methods.

It was demonstrated that the RF-driven method has increased efficiency over previous ^{14}N HMQC methods relying on either *J*- and RDS based coherence transfer, or coherence transfer via dipolar recoupling. As detected on ^{13}C, efficiencies, expressed as a percentage of the signal intensity of a spin echo, of 16% at amine sites and 9% at amide sites in ^{13}C labelled materials were demonstrated. These compare well with previous experimental efficiencies in the literature. At amine sites in unlabelled materials, efficiencies of ~22% were achieved, which are in excess of those previously observed in unlabelled materials. Furthermore, we demonstrated the extension of the technique using ^{1}H as the detection nucleus instead of ^{13}C. Protons are the nuclei of choice for indirect ^{14}N detection in terms of sensitivity due to their increased natural abundance and gyromagnetic ratio compared to ^{13}C. Detecting with protons, we were able to achieve efficiencies of up to 27% at some ^{14}N sites, which is a more than twofold increase on those efficiencies previously reported in the literature[161].

Additionally, The RF-driven correlation experiment described in this work is not highly reliant on high RF amplitudes, as typical HMQC experiments on ^{14}N are. Moderate ^{14}N RF fields of ~30kHz can be used to generate coherence between ^{14}N and detected nuclei with good efficiency, with numerical simulations indicating that experimentally accessible optimal efficiencies should be observed at ~50-60kHz depending on the size of the quadrupolar interaction. This makes the experiment possible on a wide variety of probes with various coil and, therefore, sample sizes, and is not limited to small diameter coil probes. The numerical simulations performed also indicate that the experiment is robust with respect to RF field strength. Fine optimisation of the RF field amplitude and length to achieve signal at sites with different ^{14}N EFG tensor parameters is not necessary, meaning that ^{14}N sites with different EFG tensor characteristics can be acquired in a single experiment, if investigating complex samples.

6.2 Lineshape Analysis of Indirectly Detected ^{14}N Spectra

The combination of high sensitivity afforded by proton detection combined with the increased efficiency of the RF-driven indirect detection pulse sequence allowed acquisition of indirect ^{14}N lineshapes with a high S/N ratios and well defined lineshapes on a number of model samples. This allowed us to address a current limitation in the analysis of indirectly detected ^{14}N lineshapes. It is simple to extract, from the centre of gravity of indirect ^{14}N lineshapes, the quadrupolar product, χ_Q, since the ^{14}N shift is the sum of the isotropic chemical shift and second order quadrupolar shift. However, this only provides a range of C_Q and η values for the ^{14}N EFG tensor at a given site. In order to determine the C_Q more precisely, a lineshape fitting approach is required in order to determine η. Previously, the S/N of indirect ^{14}N lineshapes has been one of the limiting factors that has prevented this being attempted. We used SQ ^{14}N spectra of a number of ^{14}N sites in a number of compounds with different quadrupolar couplings and fit simulations of ^{14}N indirectly detected lineshapes accounting for interactions up to second order to the spectra. It was found that the precision with which ^{14}N EFG tensor parameters could be determined depended on the characteristics of the ^{14}N EFG tensors at individual sites. At a site with moderate C_Q coupling magnitude (1.53MHz), a robust fit was achieved, with η defined to a range of ±0.05, and the C_Q could be constrained to within 200kHz. At sites with lower values of C_Q, numerous values of η became consistent with experimental lineshapes, and the C_Q could be determined less precisely. This was attributed to C_Q and η being correlated, and η therefore having a less pronounced effect on the ^{14}N lineshape when the C_Q is low. Experimental SQ ^{14}N spectra of amine sites were difficult to fit with simulations, though it is well known these sites are mobile and that SQ ^{14}N spectra can suffer from considerable motional broadening[77, 163, 164]. The EFG parameters for ^{14}N sites with larger C_Q magnitudes (~3MHz) were somewhat underestimated by simulations, and the reason remains unclear. It may potentially be attributable to higher order terms in the quadrupolar Hamiltonian not being included in simulations, or it could be scaling due to motions. The fitting of simulations to experimental spectra

presented here constitutes a more robust and rigorous analysis of the accuracy with which ^{14}N EFG tensor parameters can be determined from indirect detection spectra than is found elsewhere in the literature.

6.3 ^{14}N Spectra and EFG Tensors in a Protein

Using variants of the novel indirect ^{14}N detection pulse sequence we have developed, we have acquired a series of indirectly detected ^{14}N spectra from the protein GB3. Detected via protons, at high field with fast MAS, and using paramagnetic agents to enhance T_1 relaxation rates and recycle quickly, a 2D ^{1}H/^{14}N correlation spectrum of microcrystalline GB3 was recorded in 28 hours. This spectrum provided the first insight into the ^{14}N EFG tensors found at amide sites in proteins. As expected from amide containing model samples, these were found to be consistent with ^{14}N EFG tensors with a magnitude of ~3MHz.

Using cryo-DNP to enhance the sensitivity of indirect detection experiments, we were able to record 2D ^{13}C/^{14}N correlation spectra on a sample of microcrystalline GB3. These experiments confirmed that ^{14}N shifts in GB3 were consistent with quadrupolar couplings on the order of 3MHz. Furthermore, the FID was allowed to evolve for longer in this experiment than in the proton detected spectrum, meaning the distribution of ^{14}N resonances was better defined. This allowed us to estimate the range of quadrupolar couplings present at the amide sites in GB3. With the aid of a 'calculated' spectrum, constructed from a ^{13}C/^{15}N spectrum of GB3 with estimated second order quadrupolar shifts added to the ^{15}N dimension according to the secondary structure at the site, it was determined that the variation in ^{14}N shifts could be explained by ^{14}N EFG tensors varying in magnitude by 0.3MHz - 0.5MHz in the protein backbone. This was manifested in a range of ^{14}N shifts in the protein of >100ppm, which compares favourably to the chemical shift range of ^{15}N in proteins of ~30ppm. This provides an illustration of how minutely sensitive the ^{14}N EFG tensor is to local electrostatic structure, and its potential to serve as a useful reported on the local structure in proteins.

6.4 Future Directions

We have demonstrated a number of advances in the development of indirect ^{14}N detection methodology, and its application to complex molecules such as proteins, however there is scope for further developments in all aspects of the research project.

The pulse sequence we have developed for the indirect detection of ^{14}N has potential for improvement in a number of areas. A number of methods that have been applied in the literature for similar experiments could be used to augment the efficiency or accuracy of second order quadrupolar lineshapes in these experiments. For instance, in the proton detected version, homonuclear decoupling sequences applied to ^{1}H during the t_1 evolution could prolong the ^{14}N FID, alleviating ^{14}N linebroadening due to strong homogeneous $^{1}H/^{1}H$ couplings and yielding more accurate second order quadrupolar lineshapes. Ultrafast MAS could be used to achieve the same effect. This would also increase the resolution of experiments by increasing the resolution in the direct dimension by more significantly attenuating or removing ^{1}H homonuclear dipolar couplings. The new generation of ultrafast MAS probes, capable of spinning frequencies up to 111kHz, also have the potential to provide very high power RF irradiation (>100kHz) on the ^{14}N channel. This regime was not investigated in this thesis, even in simulations, and it would be interesting to examine the performance of the experiment using extremely high RF amplitudes. The methods for increasing ^{14}N resolution mentioned here may also be beneficial for fitting of simulated lineshapes to experimental spectra. Better resolved ^{14}N lineshapes could give rise to better defined features in ^{14}N signals easier to fit with simulations.

In both ^{13}C and proton detected versions of the experiment, one could also investigate other types of pulses applied on the ^{14}N channel over square pulses. Adiabatic pulses such as WURST pulses, and pulse trains such as DANTE have both been shown to provide wide excitation bandwidths up to the MHz range[70, 71, 192], which may increase the efficiency of the method.

There a number of interesting lines of investigation that could be pursued in the analysis of indirectly detected ^{14}N lineshapes by simulation fitting. For a

number of reasons, it would be informative to repeat the procedure using DQ indirect ^{14}N spectra. While the DQ transfer efficiency is lower than that of the SQ transfer, the DQ experiment has two particular advantages over the SQ for this type of lineshape analysis that should help identify sources of error in the procedure, and potentially improve the robustness of the fit. Firstly, unlike SQ ^{14}N spectra, DQ ^{14}N spectra are not affected by the FOQI. This makes them far less susceptible to lineshape distortions due to mis-setting of the magic angle, and/or unstable MAS. Secondly, DQ spectra of quadrupolar nuclei with moderate or large quadrupolar couplings (>1MHz) are expected to be less sensitive to broadening due to dynamics[77, 163, 164]. This may be of particular benefit when investigating sites known to have high mobility, such as amines. Investigating the effects of higher (third and possibly fourth) order terms in the quadrupolar Hamiltonian on indirectly detected ^{14}N lineshapes by calculation and simulation could be of benefit in determining sources of error in the fitting of ^{14}N lineshapes. This would be of particular use when investigating compounds with larger quadrupolar couplings, such as NAV, where these terms would be expected to be larger, and potentially have a more significant affect on ^{14}N peak position and lineshape.

Considering the application of indirect ^{14}N detection experiments to proteins, one can envision a number of beneficial future experiments to perform. In terms of increasing the resolution of the experiments described in this thesis on GB3, some suggestions are put forward. The proton detected 2D ^{1}H/^{14}N correlation experiment presented in this thesis was limited in resolution due to time constraints performing experiments at a national facility. Simply acquiring for longer would increase the resolution of this spectrum and allow a better analysis of ^{14}N signals in GB3. This experiment could benefit significantly from MAS at higher frequencies (>100kHz), which would increase resolution in the ^{1}H dimension. Additionally, use of a perdeuterated sample back exchanged with protons at amide sites would be beneficial for increasing resolution. Use of deuterated solvents in sample preparation, or pulsed solvent suppression techniques would be of benefit in reducing artefacts in the spectra. It would also be interesting to indirectly acquire the DQ ^{14}N spectrum of GB3 at fast MAS. The

attenuated ^{14}N transfer efficiencies at room temperature over those at cryogenic temperatures lead us to believe that protein dynamics on the timescale of applied ^{14}N RF fields interfere with the transfer and attenuate the signal. Since DQ spectra should be less sensitive than SQ spectra, they may in fact work well in this situation. Furthermore, the ^{14}N DQ spectrum should be better resolved, as signals are expected not to be broader, but to be distributed over twice the spectral width as the SQ spectrum, as resonances appear at $2\times\delta_{CS}^{iso} + 2\times\delta_{Q}^{iso}$.

For ^{13}C detected DNP experiments, use of smaller coil diameter probes with faster MAS rates, when available, could also be of benefit. Faster MAS would facilitate the use of larger indirect spectral widths, potentially allowing DQ spectra of inherently higher resolution to be acquired. A smaller sapphire rotor surface area may also increase sensitivity enhancements from DNP for a given radical concentration as less of the microwaves would be diffracted or scattered by the sapphire rotor[193]. Though it remains to seen whether DNP mechanisms with currently available biradicals are effective in a fast MAS regime[177, 178, 180].

For both ^{1}H and ^{13}C detected experiments on GB3, it could be beneficial to explore selective labelling schemes to reduce spectral crowding. For ^{13}C detected experiments, sparse ^{13}C or ^{15}N labelling could be employed, and with ^{1}H detected experiments, sparse ^{15}N labelling would be necessary. Using this technique, it may be easier to unambiguously assign ^{14}N sites to amino acid residues based on the chemical shift of the detected nucleus. In this manner, one could easily determine ^{14}N δ_{Q}^{iso} values at specific sites, from the assigned ^{15}N chemical shift, allowing a more precise analysis of the ^{14}N EFG tensor parameters in the protein. If resolution and sensitivity were high enough, a lineshape analysis as presented in this thesis may also be possible. Site specific resolution of ^{14}N sites may alternatively be attainable by using 3D experiments with a ^{14}N dimension in both ^{13}C and ^{1}H detected experiments. Should this goal be realised, it would be possible and of significant interest to perform an analysis of the ^{14}N EFG tensor parameters found at individual amide sites in proteins as a function of their structure and dynamics. At this point, the method and analysis

could potentially be extended to other proteins of increased biological interest amenable to ssNMR, such as amyloid fibrils and membrane proteins.

REFERENCES

1. Gan, Z.H., *Measuring amide nitrogen quadrupolar coupling by high-resolution N-14/C-13 NMR correlation under magic-angle spinning.* Journal of the American Chemical Society, 2006. **128**(18): p. 6040-6041.
2. Cavadini, S., et al., *Nitrogen-14 NMR spectroscopy using residual dipolar splittings in solids.* Journal of the American Chemical Society, 2006. **128**(24): p. 7706-7707.
3. Middleton, D.A., *Solid-state NMR detection of (14)N-(13)C dipolar couplings between amino acid side groups provides constraints on amyloid fibril architecture.* Magnetic Resonance in Chemistry, 2011. **49**(2): p. 65-69.
4. Jarvis, J.A., et al., *An efficient NMR method for the characterisation of 14N sites through indirect 13C detection.* Phys Chem Chem Phys, 2013. **15**(20): p. 7613-20.
5. O'Dell, L.A., *Direct detection of nitrogen-14 in solid-state NMR spectroscopy.* Progress in Nuclear Magnetic Resonance Spectroscopy, 2011. **59**(4): p. 295-318.
6. Lu, J.X., et al., *Molecular structure of beta-amyloid fibrils in Alzheimer's disease brain tissue.* Cell, 2013. **154**(6): p. 1257-68.
7. Schutz, A.K., et al., *Atomic-resolution three-dimensional structure of amyloid beta fibrils bearing the Osaka mutation.* Angew Chem Int Ed Engl, 2015. **54**(1): p. 331-5.
8. Comellas, G. and C.M. Rienstra, *Protein structure determination by magic-angle spinning solid-state NMR, and insights into the formation, structure, and stability of amyloid fibrils.* Annu Rev Biophys, 2013. **42**: p. 515-36.
9. Andreas, L.B., et al., *Structure and Mechanism of the Influenza A M2 Dimer of Dimers.* J Am Chem Soc, 2015.
10. Park, S.H., et al., *Structure of the chemokine receptor CXCR1 in phospholipid bilayers.* Nature, 2012. **491**(7426): p. 779-83.

11. Wang, S. and V. Ladizhansky, *Recent advances in magic angle spinning solid state NMR of membrane proteins.* Prog Nucl Magn Reson Spectrosc, 2014. **82**: p. 1-26.
12. Mehring, M., *Principles of high-resolution NMR in solids*. 2nd, rev. and enl. ed. 1983, Berlin ; New York: Springer-Verlag. viii, 342 p.
13. Haeberlen, U., *High resolution NMR in solids : selective averaging*. Advances in magnetic resonance : Supplement. 1976, New York: Academic Press. v.
14. Duer, M.J., *Introduction to solid-state NMR spectroscopy*. 2004, Oxford, UK ; Malden, MA: Blackwell. xiv, 349 p.
15. Ernst, R.R., G. Bodenhausen, and A. Wokaun, *Principles of nuclear magnetic resonance in one and two dimensions*. The International series of monographs on chemistry. 1987, Oxford Oxfordshire

New York: Clarendon Press ;

Oxford University Press. xxiv, 610 p.

16. Varshalovich, D.A., A.N. Moskalev, and V.K. Khersonskiĭ, *Quantum theory of angular momentum : irreducible tensors, spherical harmonics, vector coupling coefficients, 3nj symbols*. 1988, Singapore ; Teaneck, NJ, USA: World Scientific Pub. x, 514 p.
17. Samoson, A., *Satellite Transition High-Resolution Nmr of Quadrupolar Nuclei in Powders.* Chemical Physics Letters, 1985. **119**(1): p. 29-32.
18. Frydman, L. and J.S. Harwood, *Isotropic Spectra of Half-Integer Quadrupolar Spins from Bidimensional Magic-Angle-Spinning Nmr.* Journal of the American Chemical Society, 1995. **117**(19): p. 5367-5368.
19. Gan, Z.H., *Isotropic NMR spectra of half-integer quadrupolar nuclei using satellite transitions and magic-angle spinning.* Journal of the American Chemical Society, 2000. **122**(13): p. 3242-3243.
20. O'Dell, L.A., *14N Overtone Magic Angle Spinning NMR.* 2015. **86**: p. 211-236.

21. Cavadini, S., *Indirect detection of nitrogen-14 in solid-state NMR spectroscopy.* Progress in Nuclear Magnetic Resonance Spectroscopy, 2010. **56**(1): p. 46-77.
22. Stark, R.E., R.A. Haberkorn, and R.G. Griffin, *14N NMR determination of NH bond lengths in solids.* The Journal of Chemical Physics, 1978. **68**(4): p. 1996.
23. Haberkorn, R.A., et al., *Determination of Bond Distances and Bond Angles by Solid-State Nuclear Magnetic-Resonance - C-13 and N-14 Nmr-Study of Glycine.* Journal of the American Chemical Society, 1981. **103**(10): p. 2534-2539.
24. Mcdowell, C.A., et al., *Determination of the N-14 Quadrupole Coupling Tensors in a Single-Crystal of L-Histidine Hydrochloride Monohydrate by Nmr-Spectroscopy.* Journal of Magnetic Resonance, 1986. **69**(2): p. 283-292.
25. Naito, A. and C.A. McDowell, *Determination of the 14N quadrupole coupling tensors and the 13C chemical shielding tensors in a single crystal of L-asparagine monohydrate.* The Journal of Chemical Physics, 1984. **81**(11): p. 4795.
26. Rothgeb, T.M. and E. Oldfield, *N-14 Nuclear Magnetic-Resonance Spectroscopy as a Probe of Lipid Bilayer Headgroup Structure.* Journal of Biological Chemistry, 1981. **256**(12): p. 6004-6009.
27. O'dell, L.A. and R.W. Schurko, *Fast and Simple Acquisition of Solid-State N-14 NMR Spectra with Signal Enhancement via Population Transfer.* Journal of the American Chemical Society, 2009. **131**(19): p. 6658-+.
28. O'dell, L.A. and R.W. Schurko, *Static solid-state N-14 NMR and computational studies of nitrogen EFG tensors in some crystalline amino acids.* Physical Chemistry Chemical Physics, 2009. **11**(32): p. 7069-7077.
29. O'Dell, L.A. and R.W. Schurko, *QCPMG using adiabatic pulses for faster acquisition of ultra-wideline NMR spectra.* Chemical Physics Letters, 2008. **464**(1-3): p. 97-102.

30. O'Dell, L.A., A.J. Rossini, and R.W. Schurko, *Acquisition of ultra-wideline NMR spectra from quadrupolar nuclei by frequency stepped WURST–QCPMG.* Chemical Physics Letters, 2009. **468**(4-6): p. 330-335.
31. Harris, K.J., et al., *Rapid acquisition of 14N solid-state NMR spectra with broadband cross polarization.* Chemistry, 2013. **19**(48): p. 16469-75.
32. Veinberg, S.L., et al., *Ultra-wideline14N solid-state NMR as a method for differentiating polymorphs: glycine as a case study.* CrystEngComm, 2015.
33. Giavani, T., et al., *A solid-state N-14 magic-angle spinning NMR study of some amino acids.* Journal of Magnetic Resonance, 2004. **166**(2): p. 262-272.
34. Lindstrom, F., P.T.F. Williamson, and G. Grobner, *Molecular insight into the electrostatic membrane surface potential by N-14/P-31 MAS NMR spectroscopy: Nociceptin-lipid association.* Journal of the American Chemical Society, 2005. **127**(18): p. 6610-6616.
35. Doux, J.P., B.A. Hall, and J.A. Killian, *How lipid headgroups sense the membrane environment: an application of (1)(4)N NMR.* Biophys J, 2012. **103**(6): p. 1245-53.
36. Scherer, P.G.a.S., Joachim, *Electric charge effects on phospholipid headgroups. Phosphatidylcholine in mixtures with cationic and anionic amphiphiles.* Biochemistry, 1989. **28**(19): p. 7720--7728.
37. Ramamoorthy, A., et al., *Nitrogen-14 solid-state NMR spectroscopy of aligned phospholipid bilayers to probe peptide-lipid interaction and oligomerization of membrane associated peptides.* Journal of the American Chemical Society, 2008. **130**(33): p. 11023-11029.
38. Bloom, M. and M.A. Legros, *Direct Detection of 2-Quantum Coherence.* Canadian Journal of Physics, 1986. **64**(11): p. 1522-1528.
39. Tycko, R., P.L. Stewart, and S.J. Opella, *Peptide Plane Orientations Determined by Fundamental and Overtone N-14 Nmr.* Journal of the American Chemical Society, 1986. **108**(18): p. 5419-5425.

40. Tycko, R. and S.J. Opella, *High-Resolution N-14 Overtone Spectroscopy - an Approach to Natural Abundance Nitrogen Nmr of Oriented and Polycrystalline Systems.* Journal of the American Chemical Society, 1986. **108**(12): p. 3531-3532.
41. Tycko, R. and S.J. Opella, *Overtone NMR spectroscopy.* The Journal of Chemical Physics, 1987. **86**(4): p. 1761.
42. Wi, S.S. and L. Frydman, *Heteronuclear recoupling in solid-state magic-angle-spinning NMR via overtone irradiation.* Journal of the American Chemical Society, 2001. **123**(42): p. 10354-10361.
43. Takegoshi, K., et al., *Indirect high-resolution observation of N-14 NMR in rotating solids.* Journal of the American Chemical Society, 2001. **123**(43): p. 10786-10787.
44. Fukazawa, J., et al., *(14)N Quadrupolar Coupling of Amide Nitrogen and Peptide Secondary Structure As Studied by Solid-State NMR Spectroscopy.* Journal of the American Chemical Society, 2010. **132**(12): p. 4290-4294.
45. O'dell, L.A. and C.I. Ratcliffe, *(14)N magic angle spinning overtone NMR spectra.* Chemical Physics Letters, 2011. **514**(1-3): p. 168-173.
46. Nishiyama, Y., et al., *Proton-nitrogen-14 overtone two-dimensional correlation NMR spectroscopy of solid-sample at very fast magic angle sample spinning.* J Magn Reson, 2013. **230**: p. 160-4.
47. Rossini, A.J., L. Emsley, and L.A. O'Dell, *Dynamic nuclear polarisation enhanced (14)N overtone MAS NMR spectroscopy.* Phys Chem Chem Phys, 2014. **16**(25): p. 12890-9.
48. Zhao, X., et al., *Heteronuclear polarization transfer by symmetry-based recoupling sequences in solid-state NMR.* Solid State Nucl Magn Reson, 2004. **26**(2): p. 57-64.
49. Haies, I.M., et al., *(14)N overtone NMR under MAS: signal enhancement using symmetry-based sequences and novel simulation strategies.* Phys Chem Chem Phys, 2015. **17**(9): p. 6577-87.
50. Haies, I., et al., *14N overtone transition in double rotation solid-state NMR.* Phys. Chem. Chem. Phys., 2015.

51. Gullion, T., *Introduction to rotational-echo, double-resonance NMR.* Concepts in Magnetic Resonance, 1998. **10**(5): p. 277-289.
52. Gullion, T. and J. Schaefer, *Rotational-Echo Double-Resonance Nmr.* Journal of Magnetic Resonance, 1989. **81**(1): p. 196-200.
53. Pan, Y., T. Gullion, and J. Schaefer, *Determination of C-N Internuclear Distances by Rotational-Echo Double-Resonance Nmr of Solids.* Journal of Magnetic Resonance, 1990. **90**(2): p. 330-340.
54. Grey, C.P., A.P.A.M. Eijkelenboom, and W.S. Veeman, *N-14 Population Transfers in 2-Dimensional C-13-N-14-H-1 Triple-Resonance Magic-Angle-Spinning Nuclear-Magnetic-Resonance Spectroscopy.* Solid State Nuclear Magnetic Resonance, 1995. **4**(2): p. 113-120.
55. Grey, C.P., W.S. Veeman, and A.J. Vega, *Rotational Echo N-14/C-13/H-1 Triple-Resonance Solid-State Nuclear-Magnetic-Resonance - a Probe of C-13-N-14 Internuclear Distances.* Journal of Chemical Physics, 1993. **98**(10): p. 7711-7724.
56. Gullion, T., *Detecting C-13-O-17 dipolar interactions by rotational-echo, adiabatic-passage, double-resonance NMR.* Journal of Magnetic Resonance Series A, 1995. **117**(2): p. 326-329.
57. Gullion, T., *Measurement of Dipolar Interactions between Spin-1/2 and Quadrupolar Nuclei by Rotational-Echo, Adiabatic-Passage, Double-Resonance Nmr.* Chemical Physics Letters, 1995. **246**(3): p. 325-330.
58. Gan, Z., *Measuring multiple carbon-nitrogen distances in natural abundant solids using R-RESPDOR NMR.* Chem Commun (Camb), 2006(45): p. 4712-4.
59. Bodenhausen, G. and D.J. Ruben, *Natural Abundance N-15 Nmr by Enhanced Heteronuclear Spectroscopy.* Chemical Physics Letters, 1980. **69**(1): p. 185-189.
60. Muller, L., *Sensitivity Enhanced Detection of Weak Nuclei Using Heteronuclear Multiple Quantum Coherence.* Journal of the American Chemical Society, 1979. **101**(16): p. 4481-4484.

61. Hexem, J.G., M.H. Frey, and S.J. Opella, *Influence of N-14 on C-13 Nmr-Spectra of Solids.* Journal of the American Chemical Society, 1981. **103**(1): p. 224-226.
62. Cavadini, S., et al., *Indirect detection of nitrogen-14 in solids via protons by nuclear magnetic resonance spectroscopy.* Journal of Magnetic Resonance, 2006. **182**(1): p. 168-172.
63. Gan, Z.H., *C-13/N-14 heteronuclear multiple-quantum correlation with rotary resonance and REDOR dipolar recoupling.* Journal of Magnetic Resonance, 2007. **184**(1): p. 39-43.
64. Gan, Z.H., J.P. Amoureux, and J. Trebosc, *Proton-detected N-14 MAS NMR using homonuclear decoupled rotary resonance.* Chemical Physics Letters, 2007. **435**(1-3): p. 163-169.
65. Tatton, A.S., et al., *N-14-H-1 Heteronuclear Multiple-Quantum Correlation Magic-Angle Spinning NMR Spectroscopy of Organic Solids.* Zeitschrift Fur Physikalische Chemie-International Journal of Research in Physical Chemistry & Chemical Physics, 2012. **226**(11-12): p. 1187-1203.
66. Tatton, A.S., et al., *Probing hydrogen bonding in cocrystals and amorphous dispersions using (14)N-(1)H HMQC solid-state NMR.* Mol Pharm, 2013. **10**(3): p. 999-1007.
67. Cavadini, S., A. Abraham, and G. Bodenhausen, *Coherence transfer between spy nuclei and nitrogen-14 in solids.* Journal of Magnetic Resonance, 2008. **190**(1): p. 160-164.
68. Antonijevic, S. and N. Halpern-Manners, *Probing amide bond nitrogens in solids using N-14 NMR spectroscopy.* Solid State Nuclear Magnetic Resonance, 2008. **33**(4): p. 82-87.
69. Morris, G.A. and R. Freeman, *Selective Excitation in Fourier-Transform Nuclear Magnetic-Resonance.* Journal of Magnetic Resonance, 1978. **29**(3): p. 433-462.
70. Vitzthum, V., et al., *Broadband excitation and indirect detection of nitrogen-14 in rotating solids using Delays Alternating with Nutation (DANTE).* Journal of Magnetic Resonance, 2011. **212**(1): p. 234-239.

71. Vitzthum, V., et al., *Uniform broadband excitation of crystallites in rotating solids using interleaved sequences of delays alternating with nutation.* J Magn Reson, 2012. **223**: p. 228-36.
72. Siegel, R., et al., *3d H-1-C-13-N-14 Correlation Solid-State Nmr Spectrum.* Journal of Magnetic Resonance, 2008. **193**(2): p. 321-325.
73. Carravetta, M., et al., *Symmetry principles for the design of radiofrequency pulse sequences in the nuclear magnetic resonance of rotating solids.* Chemical Physics Letters, 2000. **321**(3-4): p. 205-215.
74. Cavadini, S., et al., *Line-narrowing in proton-detected nitrogen-14 NMR.* Journal of Magnetic Resonance, 2010. **202**(1): p. 57-63.
75. Nishiyama, Y., et al., *Very fast magic angle spinning H-1-N-14 2D solid-state NMR: Sub-micro-liter sample data collection in a few minutes.* Journal of Magnetic Resonance, 2011. **208**(1): p. 44-48.
76. O'Dell, L.A., R. He, and J. Pandohee, *Identifying H–N proximities in solid-state NMR using 14N overtone irradiation under fast MAS.* CrystEngComm, 2013. **15**(43): p. 8657.
77. Cavadini, S., et al., *Evidence for dynamics on a 100 ns time scale from single- and double-quantum nitrogen-14 NMR in solid peptides.* Journal of the American Chemical Society, 2008. **130**(33): p. 10850-+.
78. Maruyoshi, K., et al., *Identifying the intermolecular hydrogen-bonding supramolecular synthons in an indomethacin-nicotinamide cocrystal by solid-state NMR.* Chem Commun (Camb), 2012. **48**(88): p. 10844-6.
79. Tatton, A.S., et al., *Probing intermolecular interactions and nitrogen protonation in pharmaceuticals by novel 15N-edited and 2D 14N-1H solid-state NMR.* CrystEngComm, 2012. **14**(8): p. 2654.
80. Castellani, F., et al., *Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy.* Nature, 2002. **420**(6911): p. 98-102.
81. Castellani, F., et al., *Determination of solid-state NMR structures of proteins by means of three-dimensional N-15-C-13-C-13 dipolar correlation spectroscopy and chemical shift analysis.* Biochemistry, 2003. **42**(39): p. 11476-11483.

82. Franks, W.T., et al., *Backbone conformational constraints in a microcrystalline U-N-15-Labeled protein by 3D dipolar-shift solid-state NMR spectroscopy.* Journal of the American Chemical Society, 2006. **128**(10): p. 3154-3155.
83. McDermott, A.E., *Structural and dynamic studies of proteins by solid-state NMR spectroscopy: rapid movement forward.* Current Opinion in Structural Biology, 2004. **14**(5): p. 554-561.
84. Zech, S.G., A.J. Wand, and A.E. McDermott, *Protein structure determination by high-resolution solid-state NMR spectroscopy: Application to microcrystalline ubiquitin.* Journal of the American Chemical Society, 2005. **127**(24): p. 8618-8626.
85. Franks, W.T., et al., *Magic-angle spinning solid-state NMR spectroscopy of the beta 1 immunoglobulin binding domain of protein G (GB1): N-15 and C-13 chemical shift assignments and conformational analysis.* Journal of the American Chemical Society, 2005. **127**(35): p. 12291-12305.
86. Nadaud, P.S., J.J. Helmus, and C.P. Jaroniec, *C-13 and N-15 chemical shift assignments and secondary structure of the B3 immunoglobulin-binding domain of streptococcal protein G by magic-angle spinning solid-state NMR spectroscopy.* Biomolecular Nmr Assignments, 2007. **1**(1): p. 117-120.
87. Ulmer, T.S., et al., *Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy.* Journal of the American Chemical Society, 2003. **125**(30): p. 9179-9191.
88. Derrick, J.P. and D.B. Wigley, *The 3rd Igg-Binding Domain from Streptococcal Protein-G - an Analysis by X-Ray Crystallography of the Structure Alone and in a Complex with Fab.* Journal of Molecular Biology, 1994. **243**(5): p. 906-918.
89. Schmidt, H.L.F., et al., *Crystal polymorphism of protein GB1 examined by solid-state NMR spectroscopy and X-ray diffraction.* Journal of Physical Chemistry B, 2007. **111**(51): p. 14362-14369.

90. Wylie, B.J., et al., *Ultrahigh resolution protein structures using NMR chemical shift tensors.* Proc Natl Acad Sci U S A, 2011. **108**(41): p. 16974-9.
91. Hall, J.B. and D. Fushman, *Characterization of the overall and local dynamics of a protein with intermediate rotational anisotropy: Differentiating between conformational exchange and anisotropic diffusion in the B3 domain of protein G.* Journal of Biomolecular Nmr, 2003. **27**(3): p. 261-275.
92. Lewandowski, J.R., et al., *Direct observation of hierarchical protein dynamics.* Science, 2015. **348**(6234): p. 578-581.
93. Martin, R.W. and K.W. Zilm, *Preparation of protein nanocrystals and their characterization by solid state NMR.* Journal of Magnetic Resonance, 2003. **165**(1): p. 162-174.
94. Yao, L.S., et al., *Site-Specific Backbone Amide (15)N Chemical Shift Anisotropy Tensors in a Small Protein from Liquid Crystal and Cross-Correlated Relaxation Measurements.* Journal of the American Chemical Society, 2010. **132**(12): p. 4295-4309.
95. Yao, L.S., et al., *NMR Determination of Amide N-H Equilibrium Bond Length from Concerted Dipolar Coupling Measurements.* Journal of the American Chemical Society, 2008. **130**(49): p. 16518-+.
96. Morris, G.A. and R. Freeman, *Enhancement of Nuclear Magnetic-Resonance Signals by Polarization Transfer.* Journal of the American Chemical Society, 1979. **101**(3): p. 760-762.
97. Wittekind, M. and L. Mueller, *HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins.* Journal of Magnetic Resonance, Series B, 1993. **101**(2): p. 201-205.
98. Grzesiek, S. and A. Bax, *Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR.* Journal of the American Chemical Society, 1992. **114**(16): p. 6291-6293.

99. Logan, T., et al., *A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments.* Journal of Biomolecular NMR, 1993. **3**(2): p. 225-231.
100. Szeverenyi, N.M., M.J. Sullivan, and G.E. Maciel, *Observation of Spin Exchange by Two-Dimensional Fourier-Transform C-13 Cross Polarization-Magic-Angle Spinning.* Journal of Magnetic Resonance, 1982. **47**(3): p. 462-475.
101. Takegoshi, K., S. Nakamura, and T. Terao, *C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR.* Chemical Physics Letters, 2001. **344**(5-6): p. 631-637.
102. Hohwy, M., et al., *Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: A compensated C7 pulse sequence.* The Journal of Chemical Physics, 1998. **108**(7): p. 2686.
103. Hohwy, M., et al., *Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy.* The Journal of Chemical Physics, 1999. **110**(16): p. 7983.
104. Verel, R., et al., *A homonuclear spin-pair filter for solid-state NMR based on adiabatic-passage techniques.* Chemical Physics Letters, 1998. **287**(3-4): p. 421-428.
105. Verel, R., M. Ernst, and B.H. Meier, *Adiabatic dipolar recoupling in solid-state NMR: The DREAM scheme.* Journal of Magnetic Resonance, 2001. **150**(1): p. 81-99.
106. Bayro, M.J., et al., *Dipolar truncation in magic-angle spinning NMR recoupling experiments.* J Chem Phys, 2009. **130**(11): p. 114506.
107. Weingarth, M., et al., *Improved magnetization transfer in solid-state NMR with fast magic angle spinning.* Chemical Physics Letters, 2009. **469**(4-6): p. 342-348.
108. Hu, B., et al., *Broad-band homo-nuclear correlations assisted by 1H irradiation for bio-molecules in very high magnetic field at fast and ultra-fast MAS frequencies.* J Magn Reson, 2011. **212**(2): p. 320-9.
109. Hou, G., et al., *Spin diffusion driven by R-symmetry sequences: applications to homonuclear correlation spectroscopy in MAS NMR of*

biological and organic solids. J Am Chem Soc, 2011. **133**(11): p. 3943-53.

110. Hou, G., et al., *Broadband homonuclear correlation spectroscopy driven by combined R2(n)(v) sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids.* J Magn Reson, 2013. **232**: p. 18-30.
111. Scholz, I., et al., *MIRROR recoupling and its application to spin diffusion under fast magic-angle spinning.* Chemical Physics Letters, 2008. **460**(1-3): p. 278-283.
112. Baldus, M., et al., *Efficient N-15-C-13 polarization transfer by adiabatic-passage Hartmann-Hahn cross polarization.* Journal of Magnetic Resonance Series A, 1996. **118**(1): p. 140-144.
113. Baldus, M., et al., *Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems.* Molecular Physics, 1998. **95**(6): p. 1197-1207.
114. Fung, B.M., A.K. Khitrin, and K. Ermolaev, *An improved broadband decoupling sequence for liquid crystals and solids.* Journal of Magnetic Resonance, 2000. **142**(1): p. 97-101.
115. Marion, D., et al., *Rapid Recording of 2d Nmr-Spectra without Phase Cycling - Application to the Study of Hydrogen-Exchange in Proteins.* Journal of Magnetic Resonance, 1989. **85**(2): p. 393-399.
116. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-93.
117. Morcombe, C.R. and K.W. Zilm, *Chemical shift referencing in MAS solid state NMR.* Journal of Magnetic Resonance, 2003. **162**(2): p. 479-486.
118. Wishart, D.S., et al., *H-1, C-13 and N-15 Chemical-Shift Referencing in Biomolecular Nmr.* Journal of Biomolecular Nmr, 1995. **6**(2): p. 135-140.
119. Wishart, D.S. and B.D. Sykes, *The C-13 Chemical-Shift Index - a Simple Method for the Identification of Protein Secondary Structure*

Using C-13 Chemical-Shift Data. Journal of Biomolecular Nmr, 1994. **4**(2): p. 171-180.

120. Wishart, D.S., B.D. Sykes, and F.M. Richards, *The Chemical-Shift Index - a Fast and Simple Method for the Assignment of Protein Secondary Structure through Nmr-Spectroscopy.* Biochemistry, 1992. **31**(6): p. 1647-1651.
121. Hafsa, N.E. and D.S. Wishart, *CSI 2.0: a significantly improved version of the Chemical Shift Index.* J Biomol NMR, 2014. **60**(2-3): p. 131-46.
122. Schwarzinger, S., et al., *Sequence-Dependent Correction of Random Coil NMR Chemical Shifts.* Journal of the American Chemical Society, 2001. **123**(13): p. 2970-2978.
123. Daviso, E., et al., *Efficient resonance assignment of proteins in MAS NMR by simultaneous intra- and inter-residue 3D correlation spectroscopy.* J Biomol NMR, 2013. **55**(3): p. 257-65.
124. Cavadini, S., et al., *Indirect detection of nitrogen-14 in solid-state NMR spectroscopy.* Chemphyschem, 2007. **8**(9): p. 1363-1374.
125. States, D.J., R.A. Haberkorn, and D.J. Ruben, *A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in 4 Quadrants.* Journal of Magnetic Resonance, 1982. **48**(2): p. 286-292.
126. van Beek, J.D., *matNMR: A flexible toolbox for processing, analyzing and visualizing magnetic resonance data in Matlab((R)).* Journal of Magnetic Resonance, 2007. **187**(1): p. 19-26.
127. Bertani, P., J. Raya, and B. Bechinger, *15N chemical shift referencing in solid state NMR.* Solid State Nucl Magn Reson, 2014. **61-62**: p. 15-8.
128. Veshtort, M. and R.G. Griffin, *SPINEVOLUTION: A powerful tool for the simulation of solid and liquid state NMR experiments.* Journal of Magnetic Resonance, 2006. **178**(2): p. 248-282.
129. Harbison, G., J. Herzfeld, and R.G. Griffin, *N-15 Chemical-Shift Tensors in L-Histidine Hydrochloride Monohydrate.* Journal of the American Chemical Society, 1981. **103**(16): p. 4752-4754.

130. Bak, M., et al., *Specification and visualization of anisotropic interaction tensors in polypeptides and numerical simulations in biological solid-state NMR.* Journal of Magnetic Resonance, 2002. **154**(1): p. 28-45.
131. Cheng, V.B., H.H. Suzukawa, and Wolfsber.M, *Investigations of a Nonrandom Numerical-Method for Multidimensional Integration.* Journal of Chemical Physics, 1973. **59**(8): p. 3992-3999.
132. Strohmeier, M., D.W. Alderman, and D.M. Grant, *Obtaining molecular and structural information from C-13-N-14 systems with C-13 FIREMAT experiments.* Journal of Magnetic Resonance, 2002. **155**(2): p. 263-277.
133. Rabbani, S.R., et al., *Measurement of the N-14 Quadrupole Coupling-Constants in Glycine, Diglycine, Triglycine, and Tetraglycine and a Comparison with Calculation.* Journal of Magnetic Resonance, 1987. **72**(2): p. 230-237.
134. Hunt, M.J., A.L. Mackay, and D.T. Edmonds, *Nuclear-Quadrupole Resonance of N-14 in Imidazole and Related Compounds.* Chemical Physics Letters, 1975. **34**(3): p. 473-475.
135. Märker, K., et al., *A New Tool for NMR Crystallography: Complete 13C/15N Assignment of Organic Molecules at Natural Isotopic Abundance Using DNP-Enhanced Solid-State NMR.* Journal of the American Chemical Society, 2015.
136. Antonijevic, S. and G. Bodenhausen, *High-resolution NMR spectroscopy in solids by truly magic-angle spinning.* Angewandte Chemie-International Edition, 2005. **44**(19): p. 2935-2938.
137. Ashbrook, S.E. and S. Wimperis, *Satellite-Transition MAS NMR of Spin I=3/2, 5/2, 7/2, and 9/2 Nuclei: Sensitivity, Resolution, and Practical Implementation.* Journal of Magnetic Resonance, 2002. **156**(2): p. 269-281.
138. Dib, E., T. Mineva, and B. Alonso, *Recent Advances in 14N Solid-State NMR.* 2015.

139. Janssen, H., et al., *Microcoil high-resolution magic angle spinning NMR spectroscopy.* Journal of the American Chemical Society, 2006. **128**(27): p. 8722-8723.
140. Kentgens, A.P., et al., *High-resolution liquid- and solid-state nuclear magnetic resonance of nanoliter sample volumes using microcoil detectors.* J Chem Phys, 2008. **128**(5): p. 052202.
141. Reif, B., et al., *1H-1H MAS correlation spectroscopy and distance measurements in a deuterated peptide.* J Magn Reson, 2001. **151**(2): p. 320-7.
142. Andreas, L.B., et al., *High-resolution proton-detected NMR of proteins at very fast MAS.* J Magn Reson, 2015. **253**: p. 36-49.
143. Brown, S.P., *Applications of high-resolution 1H solid-state NMR.* Solid State Nucl Magn Reson, 2012. **41**: p. 1-27.
144. Akbey, U., et al., *Optimum levels of exchangeable protons in perdeuterated proteins for proton detection in MAS solid-state NMR spectroscopy.* J Biomol NMR, 2010. **46**(1): p. 67-73.
145. Zorin, V.E., S.P. Brown, and P. Hodgkinson, *Origins of linewidth in 1H magic-angle spinning NMR.* J Chem Phys, 2006. **125**(14): p. 144508.
146. Penzel, S., et al., *Protein resonance assignment at MAS frequencies approaching 100 kHz: a quantitative comparison of J-coupling and dipolar-coupling-based transfer methods.* J Biomol NMR, 2015. **63**(2): p. 165-86.
147. Marchetti, A., et al., *Backbone assignment of fully protonated solid proteins by 1H detection and ultrafast magic-angle-spinning NMR spectroscopy.* Angew Chem Int Ed Engl, 2012. **51**(43): p. 10756-9.
148. Ishii, Y. and R. Tycko, *Sensitivity enhancement in solid state N-15 NMR by indirect detection with high-speed magic angle spinning.* Journal of Magnetic Resonance, 2000. **142**(1): p. 199-204.
149. Zhou, D.H., et al., *Proton-detected solid-state NMR Spectroscopy of fully protonated proteins at 40 kHz magic-angle spinning.* Journal of the American Chemical Society, 2007. **129**(38): p. 11791-11801.

150. Nishiyama, Y., et al., *Studies of minute quantities of natural abundance molecules using 2D heteronuclear correlation spectroscopy under 100 kHz MAS.* Solid State Nuclear Magnetic Resonance, 2015. **66–67**: p. 56-61.
151. Nishiyama, Y., et al., *Accurate NMR determination of C–H or N–H distances for unlabeled molecules.* Solid State Nuclear Magnetic Resonance.
152. Zhang, R., et al., *A Novel High-Resolution and Sensitivity-Enhanced Three-Dimensional Solid-State NMR Experiment Under Ultrafast Magic Angle Spinning Conditions.* Scientific Reports, 2015. **5**: p. 11810.
153. Zhang, R. and A. Ramamoorthy, *Selective excitation enables assignment of proton resonances and 1H-1H distance measurement in ultrafast magic angle spinning solid state NMR spectroscopy.* The Journal of Chemical Physics, 2015. **143**(3): p. 034201.
154. Barbet-Massin, E., et al., *Rapid proton-detected NMR assignment for proteins with fast magic angle spinning.* J Am Chem Soc, 2014. **136**(35): p. 12489-97.
155. Wang, S., et al., *Nano-mole scale sequential signal assignment by (1)H-detected protein solid-state NMR.* Chem Commun (Camb), 2015. **51**(81): p. 15055-8.
156. Nieuwkoop, A.J., et al., *Sensitivity and resolution of proton detected spectra of a deuterated protein at 40 and 60 kHz magic-angle-spinning.* J Biomol NMR, 2015. **61**(2): p. 161-71.
157. Hogben, H.J., et al., *Spinach - A software library for simulation of spin dynamics in large spin systems.* Journal of Magnetic Resonance, 2011. **208**(2): p. 179-194.
158. Laage, S., et al., *Fast acquisition of multi-dimensional spectra in solid-state NMR enabled by ultra-fast MAS.* J Magn Reson, 2009. **196**(2): p. 133-41.
159. Kotecha, M., N.P. Wickramasinghe, and Y. Ishii, *Efficient low-power heteronuclear decoupling in 13C high-resolution solid-state NMR*

under fast magic angle spinning. Magn Reson Chem, 2007. **45 Suppl 1**: p. S221-30.

160. O'Dell, L.A. and A. Brinkmann, *14N overtone NMR spectra under magic angle spinning: experiments and numerically exact simulations.* J Chem Phys, 2013. **138**(6): p. 064201.
161. Pandey, M.K. and Y. Nishiyama, *Proton-detected 3D N/N/H isotropic shift correlation experiment mediated through H-H RFDR mixing on a natural abundant sample under ultrafast MAS.* J Magn Reson, 2015. **258**: p. 96-101.
162. Ashbrook, S.E., et al., *Motional broadening: an important distinction between multiple-quantum and satellite-transition MAS NMR of quadrupolar nuclei.* Chemical Physics Letters, 2002. **364**(5-6): p. 634-642.
163. Thrippleton, M.J., M. Cutajar, and S. Wimperis, *Magic angle spinning (MAS) NMR linewidths in the presence of solid-state dynamics.* Chemical Physics Letters, 2008. **452**(4-6): p. 233-238.
164. Chan-Huot, M., et al., *Deuterium MAS NMR studies of dynamics on multiple timescales: histidine and oxalic acid.* Chemphyschem, 2015. **16**(1): p. 204-15.
165. Tycko, R., *Solid-State NMR Studies of Amyloid Fibril Structure.* Annual Review of Physical Chemistry, Vol 62, 2011. **62**: p. 279-299.
166. Wei, Y.F., A.C. de Dios, and A.E. McDermott, *Solid-state N-15 NMR chemical shift anisotropy of histidines: Experimental and theoretical studies of hydrogen bonding.* Journal of the American Chemical Society, 1999. **121**(44): p. 10389-10394.
167. Li, S. and M. Hong, *Protonation, tautomerization, and rotameric structure of histidine: a comprehensive study by magic-angle-spinning solid-state NMR.* J Am Chem Soc, 2011. **133**(5): p. 1534-44.
168. Webber, A.L., et al., *Identifying guanosine self assembly at natural isotopic abundance by high-resolution 1H and 13C solid-state NMR spectroscopy.* J Am Chem Soc, 2011. **133**(49): p. 19777-95.

169. Maly, T., et al., *Dynamic nuclear polarization at high magnetic fields.* J Chem Phys, 2008. **128**(5): p. 052211.
170. Ni, Q.Z., et al., *High Frequency Dynamic Nuclear Polarization.* Accounts of Chemical Research, 2013. **46**(9): p. 1933-1941.
171. Barnes, A.B., et al., *High-Field Dynamic Nuclear Polarization for Solid and Solution Biological NMR.* Appl Magn Reson, 2008. **34**(3-4): p. 237-263.
172. Sengupta, I., P.S. Nadaud, and C.P. Jaroniec, *Protein Structure Determination with Paramagnetic Solid-State NMR Spectroscopy.* Accounts of Chemical Research, 2013. **46**(9): p. 2117-2126.
173. Pintacuda, G. and G. Kervern, *Paramagnetic Solid-State Magic-Angle Spinning NMR Spectroscopy.* Modern Nmr Methodology, 2013. **335**: p. 157-200.
174. Jaroniec, C.P., *Structural studies of proteins by paramagnetic solid-state NMR spectroscopy.* J Magn Reson, 2015. **253**: p. 50-9.
175. Parthasarathy, S., Y. Nishiyama, and Y. Ishii, *Sensitivity and Resolution Enhanced Solid-State NMR for Paramagnetic Systems and Biomolecules under Very Fast Magic Angle Spinning.* Accounts of Chemical Research, 2013. **46**(9): p. 2127-2135.
176. Thurber, K.R. and R. Tycko, *Theory for cross effect dynamic nuclear polarization under magic-angle spinning in solid state nuclear magnetic resonance: The importance of level crossings.* Journal of Chemical Physics, 2012. **137**(8).
177. Mentink-Vigier, F., et al., *Fast passage dynamic nuclear polarization on rotating solids.* J Magn Reson, 2012. **224**: p. 13-21.
178. Hu, K.N., *Polarizing agents and mechanisms for high-field dynamic nuclear polarization of frozen dielectric solids.* Solid State Nucl Magn Reson, 2011. **40**(2): p. 31-41.
179. Hu, K.N., et al., *Dynamic nuclear polarization with biradicals.* Journal of the American Chemical Society, 2004. **126**(35): p. 10844-10845.

180. Sauvee, C., et al., *Highly efficient, water-soluble polarizing agents for dynamic nuclear polarization at high frequency.* Angew Chem Int Ed Engl, 2013. **52**(41): p. 10858-61.
181. Akbey, U., et al., *Dynamic nuclear polarization of deuterated proteins.* Angew Chem Int Ed Engl, 2010. **49**(42): p. 7803-6.
182. Otting, G., *Protein NMR using paramagnetic ions.* Annu Rev Biophys, 2010. **39**: p. 387-405.
183. Wickramasinghe, N.P., et al., *Sensitivity enhancement in (13)C solid-state NMR of protein microcrystals by use of paramagnetic metal ions for optimizing (1)H T(1) relaxation.* J Magn Reson, 2007. **184**(2): p. 350-6.
184. Wickramasinghe, N.P., et al., *Progress in 13C and 1H solid-state nuclear magnetic resonance for paramagnetic systems under very fast magic angle spinning.* J Chem Phys, 2008. **128**(5): p. 052210.
185. Lamley, J.M., et al., *Solid-state NMR of a protein in a precipitated complex with a full-length antibody.* J Am Chem Soc, 2014. **136**(48): p. 16800-6.
186. Mroue, K.H., et al., *Acceleration of natural-abundance solid-state MAS NMR measurements on bone by paramagnetic relaxation from gadolinium-DTPA.* J Magn Reson, 2014. **244**: p. 90-7.
187. Ullrich, S.J., S. Holper, and C. Glaubitz, *Paramagnetic doping of a 7TM membrane protein in lipid bilayers by Gd(3)(+)-complexes for solid-state NMR spectroscopy.* J Biomol NMR, 2014. **58**(1): p. 27-35.
188. Bennett, A.E., et al., *Heteronuclear decoupling in rotating solids.* The Journal of Chemical Physics, 1995. **103**(16): p. 6951.
189. Zhou, D.H. and C.M. Rienstra, *High-performance solvent suppression for proton detected solid-state NMR.* J Magn Reson, 2008. **192**(1): p. 167-72.
190. Lewandowski, J.z.R., et al., *Enhanced Resolution and Coherence Lifetimes in the Solid-State NMR Spectroscopy of Perdeuterated Proteins under Ultrafast Magic-Angle Spinning.* The Journal of Physical Chemistry Letters, 2011. **2**(17): p. 2205-2211.

191. Fricke, P., et al., *Studies on the MxiH protein in T3SS needles using DNP-enhanced ssNMR spectroscopy.* Chemphyschem, 2014. **15**(1): p. 57-60.
192. O'Dell, L.A., *The WURST kind of pulses in solid-state NMR.* Solid State Nuclear Magnetic Resonance, 2013.
193. Kubicki, D.J., et al., *Amplifying dynamic nuclear polarization of frozen solutions by incorporating dielectric particles.* J Am Chem Soc, 2014. **136**(44): p. 15711-8.