# New Applications and Methods in Solid-State NMR

Rosalie Cresswell
Student Number: 4246745

Thesis submitted to The University of Nottingham
for the degree of Doctor of Philosophy

September 2019

# Declaration

The work in this thesis is based on research carried out at the Solid-state NMR Group, School of Chemistry, University of Nottingham, England. No part of this thesis has been submitted for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

Rosalie Cresswell

# Abstract

Solid-state NMR is a powerful, non-destructive, technique which can probe the local chemical structure of both crystalline and amorphous materials. In this thesis several different practical applications and methods of solid-state NMR have been studied.

Solid-state NMR, has been shown to be an effective technique for characterising changes in the molecular architecture of plants during a potential pretreatment process for the production of biofuels. Carbon-13 labelled wheat straw was fermented over 5 days with a fungus, *Aspergilus niger*. Quantitative analysis of 1D $^{13}$C MAS spectra over the fermentation time revealed the production of the cellulase enzymes in the first 24 hours and the steady decrease of the cellulose by 25% over the 5 day period. The 2D analysis also showed clear changes in the xylan shift from a 2-fold conformation, which binds to cellulose microfibrils, to a 3-fold conformation which cannot. It also showed a gradual broadening of the cellulose peaks indicting the cellulose becomes more amorphous throughout the fermentation.

High-field DNP-enhanced NMR has recently become commercially available due to the production of high powered microwave sources such as the gyrotron. DNP-enhanced NMR has been used to study the structure of a pharmaceutical, naloxone where an enhancement of 18 meant a 2D $^{13}$C -$^{13}$C refocused INADEQUATE could be acquired in 16 hours using a $^{13}$C natural abundance sample rather than 216 days using standard NMR. The

2D spectrum was then used to make a complete assignment of the naloxone alkaloid $^{13}C$ NMR shifts.

Finally, improvements have been made to a two-dimensional anisotropic-isotropic experiment which uses an R-sequence to recouple chemical shift anisotropy (CSA) of fluorine sites. The basic R element in the R-sequence was replaced with the $270_0 90_{180}$ composite pulse which has been shown to make the experiment significantly more robust to $B_1$ inhomogeneity. This experiment was the used to extract the CSA parameters for two inseparable polymorphs of casopitant.

# Acknowledgement

I would first and foremost like to acknowledge my PhD supervisor, Dr Jeremy Titman who has been an amazing source of knowledge and support over the last five years. Without his guidance and discussions this thesis would not have been possible. I would also like to acknowledge the facility manager at Nottingham, Dr Huw Williams for all his help with setting up experiments and keeping the NMR spectrometer working throughout my PhD.

I am grateful to EPSRC and the University of Nottingham for funding this PhD as without it I wouldn't have been able to do this work.

I would like to thank the other past and current members of the solid state NMR group; Marco Mais, Dr Lyndsey Knight and Dr Habeeba Miah, for their discussions (NMR related and otherwise) and their company over the past 5 years. I would also like to thank Sue for all of the orange at tea breaks, the electrochemists and friends for all the tea time conversations, pub times and making my PhD at Nottingham enjoyable.

Finally I would like to thank my parents, my brother and the rest of my family and friends for all their support and encouragement throughout and keeping me sane during writing my thesis. A special thanks goes to my husband, Ian who has suffered through proof reading and melt downs. I would finally like to thank my pet rabbits, Booble and Snodge for always managing to cheer me up whilst I wrote this thesis.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Nuclear Magnetic Resonance (NMR) spectroscopy is a powerful technique used to study the chemical environments, structure and dynamics of samples. It has taken decades of research and development to make NMR the effective tool it is today. Some of the key developments that have been made in solid state NMR over the years are outlined below.

## 1.1 A Brief History of NMR

Both solution and solid-state NMR were discovered almost simultaneously in 1946 by Bloch and Purcell, respectively. Purcell *et al.* detected a weak rf signal from a kilogram of paraffin when in the presence of a 0.71 T magnetic field.[17] While Bloch *et al.* observed a resonance from the protons in water using a similar method.[18] For their work, Purcell and Bloch were awarded the 1952 Nobel prize in physics *'for their development of new methods for nuclear magnetic precision measurements and discoveries in connection therewith.'*

The method developed by Purcell and Bloch was originally used by physicists to study the magnetic moments of nuclei as it was assumed that

the resonance frequency was only dependent on the magnetic field strength. However it was later shown to be dependent on the local environment of the nucleus with the discovery of the chemical shift interaction and the Knight shift.[19,20] This discovery meant NMR could be used as a analytical technique to study the different nuclear environments within samples.

The first NMR experiments used a slowly sweeping magnetic field and a constant rf wave to acquire a resonance spectrum. However this was very inefficient and slow method and so in 1949, Hahn demonstrated an alternative method.[21] This used a static magnetic field and rf pulses which could change in frequency depending on the nuclei being observed. This method is still used as the basis of modern NMR spectrometers today. The introduction of rf pulses allowed the development of complicated pulse sequences which could be used to extract a wealth of information about the nuclear environments of a sample.

Initially, solution-state NMR had significantly better resolution than solid-state NMR due to the rapid isotropic tumbling motion of solutions that average all of the anisotropic interactions away. This meant that solution-state NMR was developed more quickly and is now routinely used by chemists and pharmaceutical companies for checking structures of compounds.[22] However, in the past 40 years, huge strides have been made in solid-state NMR to make it a truly effective technique. Methods such as Magic Angle Spinning (MAS), decoupling and Cross Polarisation (CP) have been developed to increase the resolution and sensitivity of the solid-state NMR spectrum.[22]

Magic angle spinning (MAS), first developed in 1959 by Andrew *et al.*, was a major breakthrough in solid-state NMR as it dramatically improved the resolution by mimicking the isotropic tumbling motion seen in solutions.[23] This is achieved by rotating the sample at the magic angle, 54.74°

to the static magnetic field which then removes the orientation dependence of many of the anisotropic interactions. The MAS rate has to be at least the same order of magnitude as the anisotropic interaction involved to completely remove its broadening effect on the spectrum.[22,24] Therefore, there has been a lot of work into developing NMR probes which can spin faster to remove large anisotropic interactions such as the homonuclear dipolar coupling of protons. The highest spinning speeds that can now be achieved exceed 110 kHz using a 0.7 mm probe.[25]

Cross Polarisation (CP) was another key breakthrough in solid-state NMR which was first developed by Pines *et al.* in 1972 on adamantane .[26] CP significantly increased the sensitivity of the NMR signal of dilute or low abundant nuclei such as $^{13}C$ by transferring magnetisation from more abundant spins such as $^{1}H$. The CPMAS experiment was close to follow which increased both the sensitivity and resolution and is routinely used to acquire high resolution solid-state NMR spectra.[27]

Along with these techniques, there has been a lot of research into developing many different experiments to measure the anisotropic interactions that are present in solid-state NMR. The first two-dimensional NMR experiments were demonstrated by Aue *et al.* in 1976 which provided the the basic theory and method for acquiring a multi-dimensional NMR spectrum.[28] Nowdays, there are hundreds of different multi-dimensional NMR experiments that have been developed to measure anisotropic interactions, proximities and dynamics of samples. The broad range of experiments that have been developed mean that solid-state NMR has a broad range of applications such as studying polymorphs in pharmaceuticals, dynamics of polymers, protein structures, hydrogen bond lengths and more recently plant architectures.[4,29–31]

## 1.2 Thesis Overview

In this work a broad range of solid-state NMR applications and methods have been utilised to study both crystalline pharmaceuticals, as well as, amorphous plants and soils. Therefore to understand the results presented in Chapters 4 - 6, the relevant theory and experimental techniques have been described in Chapters 2 and 3.

Chapter 4 describes the study of the composition and molecular architecture of plants and humic acids by solid-state NMR. The majority of the chapter focuses on the in-situ analysis of changes in lignocellulosic material during a pretreatment process for the production of biofuels. Solid-state NMR provides a unique insight into the plant's molecular architecture that no other technique can provide.

Chapter 5 describes the application of DNP-enhanced NMR to study the structure of pharmaceutical compounds. The large enhancement of the NMR signal means that a two-dimensional experiment could be used to make a complete assignment of the $^{13}$C NMR shifts of naloxone alkaloid. The structures and shifts of naloxone alkaloid and naloxone hydrochloride dihydrate are studied and compared.

Chapter 6 introduces a pharmaceutical casopitant which has 2 inseparable polymorphs that can only be distinguished using $^{19}$F NMR. This sample was used in the development of a two-dimensional anisotropic-isotropic experiment for measuring the chemical shift anisotropy for fluorine sites which is more robust to rf inhomogeneity than a previously developed experiment by Martini *et al.*.[32]

# Chapter 2

# Theory of Solid-state NMR

Many features of NMR can be understood using the classical model, but as soon as the nuclear spins interact with each other a quantum approach is required. In this chapter the classical vector model and the basic principles of NMR are first introduced and then the quantum mechanical approach is used to describe the Hamiltonians of the different interactions present.

## 2.1 Basics of NMR

All particles, such as nuclei and electrons, have an intrinsic property known as spin, I. Any nucleus with I $>$ 0 has a magnetic moment, $\boldsymbol{\mu}$ given by;

$$\boldsymbol{\mu} = \gamma I, \tag{2.1}$$

where $\gamma$ is the gyromagnetic ratio of the nucleus.[22,33] Nuclei with a spin, I $>$ 0 are known as NMR active nuclei as the magnetic moment means that the nucleus behaves like a bar magnet in the presence of a magnetic field. Figure 2.1 illustrates the effect of the magnetic field on a sample of nuclei with a non-zero spin.[22]

When there is no magnetic field the nuclear spins are oriented in all

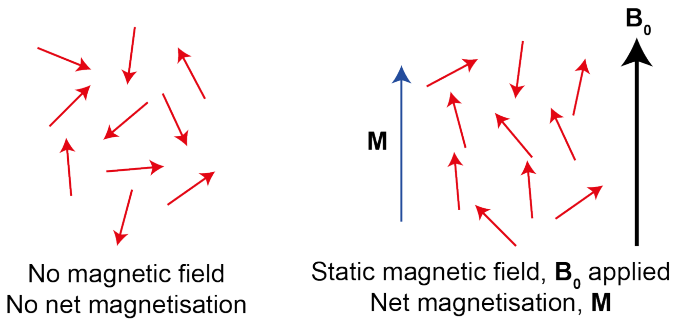


Figure 2.1: When there is no magnetic field the spins are orientated randomly in all directions. When in a magnetic field these spins preferentially align with the direction of the magnetic field leading to a net/bulk magnetisation of the sample.

directions randomly; whereas when in the presence of a magnetic field the spins align with or against the direction of the field. There is still a Boltzmann distribution of orientations but there is a net magnetisation in the direction of the applied magnetic field. The net magnetisation, **M** is the vector sum of all the magnetic moments, $\boldsymbol{\mu_i}$ in the spin system.[22,33]

When a magnetic field, **B** is applied to nuclear spins they precess about the field at a constant rate, $\omega = \gamma$B. In NMR spectroscopy, the large static field applied across the sample is usually referred to as $\mathbf{B}_0$ and the resultant precession rate, $\omega_0$ is called the Larmor frequency given by:[22,24]

$$\omega_0 = -\gamma B_0 \tag{2.2}$$

When a non-zero spin nucleus is in the presence of a magnetic field the nuclear spin energy levels are no longer degenerate and split into 2I+1 energy levels. This is the Zeeman interaction which is the basis of NMR. For a I=1/2 the Zeeman interaction leads to two nuclear spin energy levels; $\alpha$ for m= +1/2 and $\beta$ for m = -1/2. The nuclear spin energy levels are given by the equation;[22,33]

$$\boldsymbol{E} = -m\hbar\gamma B_0 \tag{2.3}$$

which leads to an energy difference of;

$$\Delta E = \hbar\gamma B_0 = \hbar\omega_0 \tag{2.4}$$

## 2.2 Radio Frequency (rf) Pulses

In an NMR experiment, radio frequency (rf) pulses are used to apply a second oscillating magnetic field. This oscillating field interacts with the nuclei in the sample in addition to the static magnetic field. The static magnetic field, $B_0$ is routinely assigned to be along the z axis whilst the oscillating field acts perpendicular to it in the xy plane. The oscillating field is made up of 2 components that rotate in opposite directions, $\pm\omega_{rf}$ where $\omega_{rf}$ is the frequency of the pulse. However, only the component that rotates in the same direction as the net magnetisation precesses about the static field has a significant effect on the nuclear spins. Therefore the oscillating field can be referred to as $\mathbf{B}_1$ with a frequency $\omega_{rf}$.[22,33]

The easiest way to understand the effect of rf pulses on the spin system is to remove the time dependence by transforming the whole system into a rotating reference frame that rotates at $\omega_{rf}$. In this rotating frame of reference the static field $\mathbf{B}_0$ along the z axis becomes $\boldsymbol{B}_0(1 - \frac{\omega_{rf}}{\omega_0})$. This shows that if the oscillating field is set on-resonance where $\omega_{rf} = \omega_0$ then the effect of the static field $B_0$ is zero. In this case, only the rf field affects the nuclear spins causing them to precess about the $\mathbf{B}_1$ field at a frequency, $\omega_1 = \gamma B_1$, which is referred to as the nutation frequency. When the oscillating field is not on-resonance the static field does not go to zero leading to an effective magnetic field, $\mathbf{B}_{eff}$ about which the nuclear spins precess as shown in Figure 2.2.[22]

An rf pulse in an NMR experiment is fully described by a nutation angle

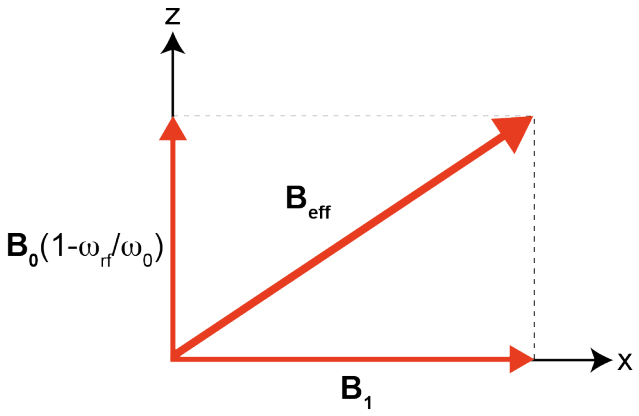


Figure 2.2: A diagram of the effective field acting on the nucleus when in the rotating reference frame and the $B_1$ field acting along the x-axis is not on-resonance.

and a phase. The nutation angle of a pulse, that is on-resonance, is the angle the magnetisation turns during the time, $\tau_{rf}$ of the pulse. Whilst the phase of the pulse defines where on the xy plane the pulse is rotated into. For example a 90° pulse with a phase of 0° leads to an rf pulse along the x-axis which will rotate the magnetisation from the z axis onto the -y axis.[22]

## 2.3 Relaxation

In order to detect an NMR signal the net magnetisation of a sample in a magnetic field $B_0$ is perturbed and studied as it equilibrates. Relaxation describes the means by which the spins return to equilibrium. There are 2 main types of relaxation: spin-lattice relaxation and spin-spin relaxation. The relaxation times vary between samples and are affected by the dynamics of the sample.[24]

Spin-lattice relaxation, also known as $T_1$ relaxation, describes the time in which the longitudinal (z-component) returns to equilibrium after it has been perturbed by an rf pulse.[33] The energy absorbed from the rf pulse is

transferred to its surroundings (lattice) and the energy is dissipated away usually in the form of a small amount of heat energy. The spin lattice relaxation can be described with the following equation,

$$M(t) = M_{eq}(1 - e^{\frac{-t}{T_1}}), \tag{2.5}$$

where M(t) and $M_{eq}$ refers to the net magnetisation vector at time t and at equilibrium respectively and $T_1$ is the relaxation time.[24] The relaxation time is fastest when the motional frequency of the nucleus matches the Larmor frequency. In most cases for solids there is very little movement resulting in long relaxation times.[24,33]

Spin-spin relaxation or $T_2$ relaxation, is the time it takes for the transverse magnetisation to decay or dephase. This relaxation occurs through spin interactions such as dipolar coupling. These interactions vary the local magnetic field around a nucleus which results in nuclei that are precessing at different rates. Over time these different precession rates lead to a dephasing of the transverse magnetisation. The removal of transverse magnetisation through restoring the longitudinal magnetisation also results in spin-spin relaxation.[24]

## 2.4 The Time-Independent Schrödinger Equation

In order to understand the interactions in NMR a quantum mechanical approach is required. A collection of nuclei which are interacting in a designated environment is commonly referred to as a spin system. In quantum mechanics the spin state is described by a wavefunction, $\psi$ or $\Psi$, which depends on both spin and spatial coordinates. Wavefunctions are described

using the quantum numbers that determine the observable parameters of the system.[22]

The wavefunction of a time invariant nuclear spin system is the solution to the time-independent Schrödinger equation given by,

$$\hat{H}\Psi = E\Psi \tag{2.6}$$

where $\hat{H}$ is the total nuclear Hamiltonian operator that acts on the wavefunction, $\Psi$, to give the total energy of the spin system, E. The Schrödinger equation is in the form of an eigenfunction equation, therefore, the wavefunction and the total energy of the spin system correspond to the eigenfunctions and eigenvalues of the equation respectively.[22] The energy of the spin system is a physically observable quantity which means that the associated operator, the Hamiltonian operator, is Hermitian.[34]

Hermitian operators are self adjoint which means that $\hat{A} = \hat{A}^{\dagger}$ where $\hat{A}^{\dagger}$ is the transpose and complex conjugate of the operator $\hat{A}$. This leads to the relationship:

$$\int f^*(\tau)\hat{A}g(\tau)d\tau = [\int g^*(\tau)\hat{A}f(\tau)d\tau]^* \tag{2.7}$$

for arbitrary functions $f(\tau)$ and $g(\tau)$ where the asterix denotes the complex conjugate.[34] This relation can then be used to show that if $f(\tau)$ is defined as an eigenfunction of the operator, $\hat{A}$ and $\lambda$ the corresponding eigenvalue then,

$$\int f^*(\tau)\hat{A}f(\tau)d\tau = \lambda \int f^*(\tau)f(\tau)d\tau = \lambda \tag{2.8}$$

$$\int f^*(\tau)\hat{A}^{\dagger}f(\tau)d\tau = \lambda^* \int f^*(\tau)f(\tau)d\tau = \lambda^* \tag{2.9}$$

where if $\hat{A}$ is associated with a physical quantity, such as energy of a system, the corresponding eigenvalues must be real ($\lambda = \lambda^*$), which is the case if the

operator is self adjoint. This proves that due to the self adjoint nature of the Hermitian operator, all corresponding operators of physical observable quantities such as the total nuclear Hamiltonian must be Hermitian.[34]

The eigenfunction of a Hermitian operator form a complete set of orthonormal functions which means that,

$$\int f_i^*(\tau) f_j(\tau) d\tau = \delta_{ij} \tag{2.10}$$

where $\delta_{ij}$ is the Kroneker delta given by;

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases} \tag{2.11}$$

This means that the wavefunction, $\Psi$, which is the eigenfunction of the nuclear Hamiltonian, can be written as the sum of a basis set of orthonormal functions given by,

$$\Psi = \sum_n c_n \psi_n \tag{2.12}$$

$c_n$ is the coefficients in the eigenfunction, $\Psi$ and $\psi_n$ form the complete basis set to describe $\Psi$.[34]

The observed value of a physically measured parameter, such as the energy of a spin system, is the expectation value of the corresponding operator. Where there are only a discrete number of possibilities, this expectation value is equal to one of the eigenvalues of the operator. However, in reality, when the parameters are measured experimentally, the expectation value is an average of the allowed discrete outcomes. The expectation value for an operator A is given by,[34]

$$\langle A \rangle = \int \Psi^*(\tau) \hat{A} \Psi(\tau) d\tau \tag{2.13}$$

When $\Psi$ is an eigenfunction of the operator $\hat{A}$ (i.e. $\Psi = \psi_n$) then the expectation value is just the corresponding eigenvalue, $\lambda_n$. However, the majority of the time, $\Psi$ is not an eigenfunction of $\hat{A}$ and so using, Equ. (2.12) the expectation value can be shown to be:

$$\langle A \rangle = \sum_{j=1}^{N} c_j^* c_j \lambda_j \tag{2.14}$$

where $c_j^* c_j$ is referred to as the probability that $\lambda_j$ is observed during a single measurement. This shows that whilst only discrete eigenvalues are allowed, the observed expectation value of many measurements of a system can have any value that fulfils the Equation 2.14.[34]

When there are two operators, $\hat{A}$ and $\hat{B}$, of the observable parameters A and B respectively, they are said to be compatible if it does not matter which order these parameters are measured. If the two operators commute such that $\hat{A}\hat{B} = \hat{B}\hat{A}$ then it can be shown that the operators have the same complete set of eigenfunctions. The commutator of two operators, $\hat{A}$ and $\hat{B}$ is defined as

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \tag{2.15}$$

If $[\hat{A}, \hat{B}] = 0$ then the operators have the same eigenfunctions otherwise the Heisenberg uncertainty relation has to be made for the two operators.[34]

## 2.5 Spin Operators and Spin States

In the majority of cases the total nuclear wavefunction $\Psi$ can be written as a product of a spatial term and a spin term. The spatial term only depends on the spatial coordinates of the nucleus whilst the spin term only depends on the nuclear spin coordinates. The total nuclear wave function can be factorised in this way as the spatial nuclear coordinates

are usually unaffected by the nuclear spin coordinates, and conversely the nuclear spin coordinates are unaffected by the spatial coordinates of the nucleus. This is described as the spin and spatial terms of the wavefunction being uncoupled.[22]

To study the spin properties of a nucleus, spin operators are applied to act on the spin coordinates of the nuclear wavefunction. For a single uncoupled spin the spin operators consist of magnitude, $\hat{I}^2$, alongside the x, y and z components given by $\hat{I}_x$, $\hat{I}_y$ and $\hat{I}_z$ respectively. These spin operators are related by:

$$\hat{I}^2 = \hat{I}_x^2 + \hat{I}_y^2 + \hat{I}_z^2 \tag{2.16}$$

and have the cyclic commutation relation:

$$[\hat{I}_x, \hat{I}_y] = i\hbar\hat{I}_z \tag{2.17}$$

At equilibrium, $\hat{I}^2$ commutes with $\hat{I}_z$ which means that they have the same set of eigenfunctions. The eigenfunctions are defined using the quantum numbers I and m denoted using $\psi_{Im}$. For a spin-1/2 nucleus; I=1/2, m=-1/2,+1/2; the eigenfunctions of $\hat{I}_z$ are usually referred to as $\alpha$ (m=+1/2) and $\beta$ (m=-1/2). The time-independent Schrödinger equation is in the form of an eigenfunction equation, and so, solving it gives the eigenvalues of the nuclear wavefunction.[22] The eigenvalues for $\psi_{Im}$ are defined by the eigenfunction equations,

$$\hat{I}^2\psi_{Im} = I(I+1)\hbar\psi_{Im} \tag{2.18}$$

$$\hat{I}_z\psi_{Im} = m\hbar\psi_{Im} \tag{2.19}$$

If the Hamiltonian of a non-interacting nucleus is not purely described

by $\hat{I}^2$ and $\hat{I}_z$ then the nuclear spin wavefunction, $\Psi$ can be written as linear combination of $\psi_{Im}$ functions such as,

$$\Psi = \sum_m c_m \psi_{Im} \tag{2.20}$$

Where $c_m$ is the combination coefficients in the eigenfunction, $\Psi$ and $\psi_{Im}$ form the basis set that equates to the expansion of $\Psi$. When this is substituted into the time-independent Schrödinger equation it gives,

$$\hat{H}\Psi = E\Psi \Rightarrow \sum_m c_m \hat{H}\psi_{Im} = E\sum_m c_m \psi_{Im} \tag{2.21}$$

Each of the coefficients can be calculated using a series of simultaneous equations where $\psi_{Im}$ are orthonormal and normalised resulting in a matrix that can be solved using,

$$\det|\boldsymbol{H} - E\mathbf{1}| = 0 \tag{2.22}$$

For a system containing a set of non-interacting nuclear spins the sum of the $\psi_{Im}$ functions are used to solve the time-independent Schrödinger equation. When a sum of functions is used to describe a system this is generally referred to as the super position of states.[22] This method is good for simple spin systems but can become quite longwinded as spin systems get larger and more complex.[22,34]

### 2.5.1 Dirac Bra-Ket notation

The notation for describing spin systems can quickly get very cumbersome and so the Dirac Bra-Ket notation is regularly utilised to make it more concise and clear. In this notation the wave function can be represented by a 'Ket' function, $|\psi\rangle$ and its complex conjugate is represented by a 'Bra'

function, $\langle\psi|$. The expectation value for the time independent Schrödinger equation can be condensed down to:

$$\int \psi^*_{Im'} \hat{H} \psi_{Im} \, d\tau = \langle I, m'| \, \hat{H} \, |I, m\rangle \tag{2.23}$$

where

$$'Bra' - \langle I, m'| = \psi^*_{Im'} \quad \text{and} \quad 'Ket' - |I, m\rangle = \psi_{Im} \tag{2.24}$$

and the Bra-Ket notation implies integration over all variables.[22,34]

## 2.6 The Density matrix operator

An alternative approach to describing a spin system in quantum mechanics is using the density operator. This method is more concise and can more easily include time-dependent terms without it becoming an arduous task.[22] When there are time dependent terms then the wavefunction, $\Psi(t)$ has to be a solution to the time-dependent Schrödinger equation which is given by:

$$\hat{H}(t)\Psi(t) = i\hbar \frac{\partial(\Psi(t))}{\partial t} \tag{2.25}$$

The time-dependent wavefunction, $\Psi(t)$ can be expanded in terms of a set of orthonormal time-independent basis functions, $|i\rangle$,

$$\Psi(t) = \sum_{n=1}^{N} c_i(t) \, |i\rangle \tag{2.26}$$

where the sum runs over all spin states and the $c_i$ coefficients contain all the time dependence.[22,34] The expectation value for an operator, $\hat{A}$ can then be written as,

$$\langle A\rangle = \sum_{i,j} c^*_i(t) c_j(t) \, \langle i| \, \hat{A} \, |j\rangle \tag{2.27}$$

It is convenient to define the product of the coefficients $c_i^*(t)c_j(t)$ as elements in a matrix representation of an operator , $\hat{P}(t)$, with each element given by,[34]

$$P(t)_{ij} = \langle i| \hat{P}(t) |j\rangle = c_i^*(t)c_j(t) \tag{2.28}$$

This operator can then be substituted in to the expectation value Eq. (2.27) to give,

$$\begin{aligned} \langle A\rangle = \sum_{i,j} c_i^*(t)c_j(t) \langle i| \hat{A} |j\rangle = \sum_{i,j} \langle i| \hat{P}(t) |j\rangle \langle i| \hat{A} |j\rangle \\ = \sum_i \langle i| \hat{P}(t)\hat{A} |j\rangle = Tr\{\hat{P}(t)\hat{A}\} \end{aligned} \tag{2.29}$$

Where $Tr\{\}$ is the trace of a matrix which is the sum of the diagonal elements.[34] The result of Eq. (2.29) states that the expectation value is just the trace of the product of $\hat{P}(t)$ and the operator $\hat{A}$ corresponding to the desired observable. $\hat{P}(t)$ describes the state of the system at any point in time. So far, this analysis is only applicable if the whole system can be described using the same wavefunction, this is called a pure state and $\hat{P(t)}$ is the density operator for this pure state spin system.[34] However, usually a spin system is in a mixed state where a sub-ensemble can be described by a wavefunction, $\Psi$, and a probability density function, $\mathcal{P}(\Psi)$, which gives the contribution of each sub-ensemble to the mixed state. The statistical expectation value for the mixed state can be found by averaging over the

whole probability distribution called an ensemble average.[22,34]

$$\langle \bar{A} \rangle = \int \mathcal{P}(\Psi) \langle \Psi | \hat{A} | \Psi \rangle \, d\tau \tag{2.30}$$
$$= \sum_{ij} \int \mathcal{P}(\Psi) c_i(t) c_j^*(t) d\tau \langle i | \hat{A} | j \rangle$$
$$= \sum_{ij} \overline{c_i(t) c_j^*(t)} \langle i | \hat{A} | j \rangle$$

The statistical expectation value shows that only the $\overline{c_i(t)c_j^*(t)}$ factors vary whilst the matrix elements of $\langle i | \hat{A} | j \rangle$ do not change. In Eq. (2.30) the overbar denotes the statistical ensemble average. The coefficients, $\overline{c_i(t)c_j^*(t)}$ of the ensemble average generates a matrix which represents the operator called the density operator, $\hat{\rho}(t)$.[22,34] Hence, the expectation value for an observable corresponding to the operator $\hat{A}$ of an ensemble of spins in a mixed state is given by;

$$\langle \bar{A} \rangle = Tr\{\hat{\rho}(t)\hat{A}\} \tag{2.31}$$

Therefore, to calculate the expectation value of an observable, only the corresponding operator of the observable and the density operator of the system must be known.[34]

## 2.7 The Liouville-von Neumann Equation

The time evolution of the spin system is given by the time evolution of the density operator which is described by the equation:

$$\frac{d\hat{\rho}(t)}{dt} = -i[\hat{H}, \hat{\rho}(t)] \tag{2.32}$$

This is the Liouville-von Neumann equation, which can be derived from the time-dependent Schrödinger equation using Dirac's Bra-Ket notation and the definition of the density operator.[34,35] For a time-independent Hamiltonian, the time evolution of the density operator is given by,

$$\hat{\rho}(t) = exp(-i\hat{H}t)\hat{\rho}(0)exp(i\hat{H}t), \tag{2.33}$$

where $\hat{\rho}(0)$ is the density operator at t=0.[34,35] When the Hamiltonian is time-dependent the equation is impossible to solve, and in most cases it is easier to transform the system into a time-independent system. One way of achieving a time-independent Hamiltonian is by changing the frame of reference to a time-independent frame. This is particularly useful for understanding the effects of an rf pulse on a system which can be made time-independent by transforming the reference frame into a frame rotating at the nutation frequency, $\omega_{rf}$.[22] An alternative method is to make the time-dependent Hamiltonian a sum of many time intervals where, within each interval, the Hamiltonian can be assumed constant. This method is the basis of the NMR simulation software, SIMPSON.[36]

## 2.8 Coherence

The diagonal of the density matrix can be shown to contain the populations of a state for a given basis function. When the spin system is at thermal equilibrium then all the off-diagonal elements are zero. However in NMR, to be able to observe the nuclear spins the spin system has to be perturbed. Any non-zero off-diagonal element in $\hat{\rho}(t)$ indicates a coherence between eigenstates $|i\rangle$ and $|j\rangle$. This means that time dependent phase properties of the spin system correlate via $|i\rangle$ and $|j\rangle$. The coherences are assigned by describing the overall change in the quantum number m. The

matrix elements that have $\Delta m = \pm 1$ are referred to as single-quantum (SQ) coherence whilst those that have $\Delta m = \pm 2$ and 0 are referred to as double-quantum (DQ) and zero-quantum (ZQ) coherence respectively.[24,34]

For a single spin-1/2 nucleus which has two spin states, $\alpha$ and $\beta$, the equilibrium density matrix is given by a $2 \times 2$ matrix:

$$\hat{\rho}(t) = \begin{bmatrix} \frac{1}{2} + \frac{p}{2} & 0 \\ 0 & \frac{1}{2} - \frac{p}{2} \end{bmatrix} \tag{2.34}$$

where the off-diagonal elements are zero and the diagonal elements are related to the populations of each state where p is given by the Boltzmann distribution, $\frac{(\hbar \gamma B_0)}{kT}$.

For a single spin-1/2 system the off-diagonal terms are single-quantum coherences. For a multi-spin system with N spin-1/2 nuclei the number of elements in the density matrix increases by $2^{2N}$. In the case of a two spin-1/2 system the density matrix is a $4 \times 4$ matrix which summarise all 16 combinations the spin states can be in. The off-diagonal elements in the density matrix shown in Eq. (2.35) are labelled with the coherences and spin states for each element and the diagonals are labelled p to represent the terms related to the population of the states.[24,34]

$$\begin{array}{c} \begin{matrix} \alpha\alpha & \alpha\beta & \beta\alpha & \beta\beta \end{matrix} \\ \begin{bmatrix} p & SQ & SQ & DQ \\ SQ & p & ZQ & SQ \\ SQ & ZQ & p & SQ \\ DQ & SQ & SQ & p \end{bmatrix} \begin{matrix} \alpha\alpha \\ \alpha\beta \\ \beta\alpha \\ \beta\beta \end{matrix} \end{array} \tag{2.35}$$

For example, if the top right-hand corner element of the matrix was non-zero then there would be a double quantum coherence between the $\alpha\alpha$

state and the $\beta\beta$ state of the 2 spin system.[24,34]

## 2.9 NMR Interactions

The total nuclear Hamiltonian can be written as a sum of interactions which can be split in to external and internal terms. The external terms are the interactions that are imposed on to the nuclei which, in NMR's case, is the static magnetic field causing the Zeeman splitting and the application of the oscillating rf field. The internal terms relate to nuclear environments and interactions between nuclei within the sample.

### 2.9.1 External Interactions

#### Zeeman Interaction and the Secular Approximation

Zeeman splitting occurs when a large static magnetic field is applied to a sample of I $>$ 0 nuclei causing the degenerate nuclear spin energy levels to split into 2I +1 energy levels. The Hamiltonian for this interaction is given by,

$$\hat{H}_Z = -\hat{\boldsymbol{\mu}} \cdot \boldsymbol{B_0} = -\gamma\hbar\hat{I}_z B_0. \tag{2.36}$$

The Zeeman Hamiltonian is usually the largest energy interaction by an order of magnitude. Therefore any internal interactions within the sample can be designated as a perturbation to the Zeeman energy. This leads to the secular approximation which states that any contribution that is not proportional to $I_z$ is assumed to be negligible.[22]

#### Radio frequency Interaction

The nuclear spin properties of the nucleus are studied by perturbing the spin system using rf pulses. The rf pulses produce an oscillating magnetic

field that interacts with the nuclear spins in the system. Many of the effects the rf pulses have on the nuclear spins have been described using the classical picture. Given the information the classical picture provides, this results in a Hamiltonian for the effect of the rf pulses on the nuclear spins of:[22]

$$\hat{H}_{rf} = -\gamma \hat{I}_x B_1 \cos(\omega_{rf} t) \tag{2.37}$$

### 2.9.2 Internal Interactions

#### Chemical shift

The chemical shift arises as the chemical environment affects the magnetic field perceived by the nucleus. The static magnetic field, $B_0$ generates an electron current with the electrons in bonds that surround the nucleus. This movement of electrons in turn generates a magnetic field around the nucleus. The induced field adds or subtracts (shields or deshields) from $B_0$ giving the local field felt by the nucleus. This shielding effect of the nearby electrons creates differences in the resonant Larmor frequency for nuclei leading to a shift in the spectrum separating different chemical environments.[22,24]

The chemical shielding Hamiltonian acting on a spin I is given by;

$$\hat{H}_{CS} = \gamma \cdot \hat{\boldsymbol{I}} \cdot \boldsymbol{\sigma} \cdot \boldsymbol{B_0} \tag{2.38}$$

where $\boldsymbol{B_0}$ is the static magnetic field acting along the z axis and $\boldsymbol{\sigma}$ is the chemical shielding tensor. The shielding tensor can be split into symmetric and asymmetric components. However, in NMR only the symmetric component has any significant effect on the spectrum. Therefore, in this work, $\boldsymbol{\sigma}$ will only refer to the symmetric component given by;

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \tag{2.39}$$

For a given nucleus it is possible to define a frame such that $\boldsymbol{\sigma}$ is diagonal. This axis frame is referred to as the Principal Axis System (PAS). The diagonal components of the shielding tensor in the principal axis system are called the principal values.[22,24]

$$\boldsymbol{\sigma}^{PAS} = \begin{pmatrix} \sigma_{xx}^{PAS} & 0 & 0 \\ 0 & \sigma_{yy}^{PAS} & 0 \\ 0 & 0 & \sigma_{zz}^{PAS} \end{pmatrix} \tag{2.40}$$

These are commonly described in terms of the isotropic shielding, $\sigma_{iso}$, anisotropy, $\Delta$ and asymmetry, $\eta$ by;

$$\sigma_{iso} = \frac{1}{3}(\sigma_{xx}^{PAS} + \sigma_{yy}^{PAS} + \sigma_{zz}^{PAS}) \tag{2.41}$$

$$\Delta = \sigma_{zz}^{PAS} - \sigma_{iso} \tag{2.42}$$

$$\eta = \frac{\sigma_{xx}^{PAS} - \sigma_{yy}^{PAS}}{\Delta} \tag{2.43}$$

where $\sigma_{iso}$ is the average shielding over the three principal values, the anisotropy $\Delta$ provides information about the electronic environment surrounding the nucleus and the asymmetry $\eta$ defines the symmetry of the nuclear environment.[22] In solution-state NMR only the isotropic shift is seen as the anisotropy is averaged away by the isotropic tumbling of the molecules in solution.[22,24,33]

The chemical shift is orientation dependent and so the chemical shift of a single crystal changes depending on the orientation of the crystal with

respect to the static magnetic field $B_0$. In a powder, all orientations are present creating a line shape referred as a powder pattern seen in Figure 2.3. The shape and width of the powder pattern depends on the principal values of the chemical shift tensor. The discontinuities in the powder pattern relate to the principal values and can be used to extract the isotropic shift, anisotropy and asymmetry of the nuclear spin environment as shown in Figure 2.4.[22,24]

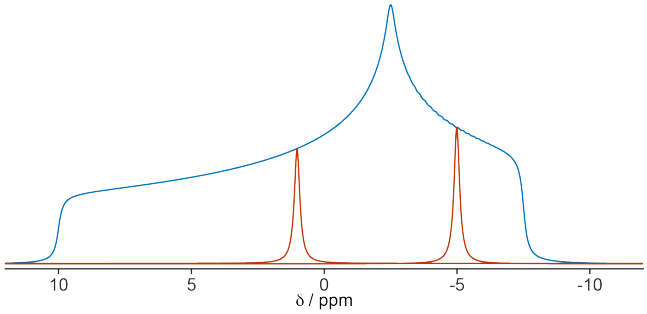


Figure 2.3: In blue is a powder pattern line shape which is a summation of all the chemical shifts for all orientations. A single crystal would result in a single peak within the range with a frequency depending on the orientation. In red are two examples of the peaks that arise from 2 different orientations of a single crystal.

Usually in NMR the chemical shift is defined relative to a reference sample's chemical shift. For example for carbon the shift of TMS is set to 0 ppm and other shifts would be found relative to this value. The chemical shift, $\delta_{iso}$ is given by;

$$\delta_{iso} = \frac{\mu - \mu_{ref}}{\mu_{ref}} \tag{2.44}$$

where $\mu$ is the frequency of the sample shift and $\mu_{ref}$ is the frequency of the reference sample. The shielding tensor can be written in this form making it the chemical shift tensor whose elements are given by;

$$\delta_{ij} = \frac{(\sigma_{ij}(ref) - \sigma_{ij})}{1 - \sigma_{ij}} \tag{2.45}$$

When in the principal axis system for the chemical shielding tensor

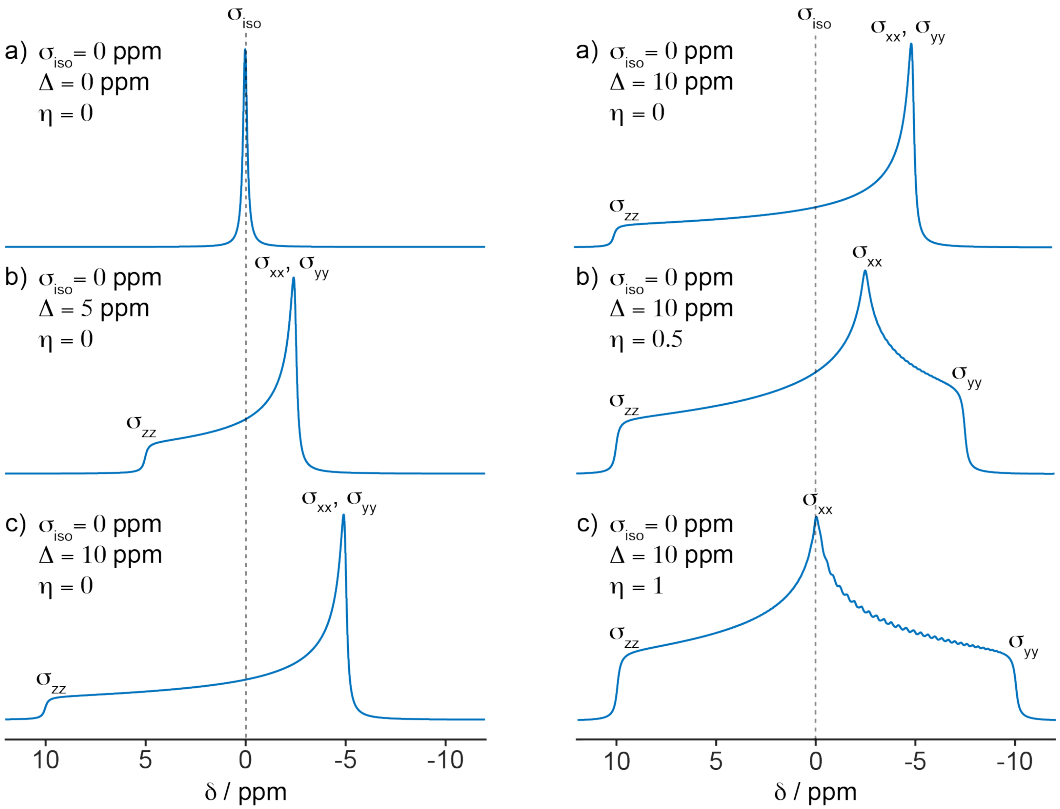


Figure 2.4: Different resulting powder patterns for different $\Delta$ and $\eta$ with an isotropic shift of 0 ppm with the principal axis values indicated on each diagram.

the analogous principal values are given by $\delta_{xx}^{PAS}$, $\delta_{yy}^{PAS}$ and $\delta_{zz}^{PAS}$. The corresponding chemical shift anisotropy, $\zeta_{cs}$ and asymmetry, $\eta_{cs}$ are then defined as,

$$\zeta_{cs} = \delta_{zz}^{PAS} - \delta_{iso} \tag{2.46}$$

$$\eta_{cs} = \frac{\delta_{yy}^{PAS} - \delta_{xx}^{PAS}}{\Delta_{cs}} \tag{2.47}$$

Where the principle axis values are defined such that:

$$\delta_{zz}^{PAS} \geq \delta_{xx}^{PAS} \geq \delta_{yy}^{PAS} \tag{2.48}$$

The powder patterns seen in NMR can be used to extract the chemical shift anisotropy and asymmetry parameters.[22] Unfortunately these powder patterns regularly overlap generating broad indistinct line shapes that can be fitted but can have large errors associated with the extracted parameters.[22,24] There are several methods that have been developed to overcome these issues which will be discussed further in Chapter 3.

### Dipolar Coupling

The dipole-dipole coupling arises from the interaction between the magnetic moments of nearby spins. The nuclear spins in a magnetic field have their own local magnetic field, the local field of a nucleus can interact with the local fields of nearby nuclei giving rise to the dipolar coupling. This interaction in solution-state NMR is not seen in the spectrum as it is completely averaged away by the isotropic molecular tumbling motions. In solid-state NMR, however, this is not the case and the dipolar coupling interaction is a major source of broadening of lines. The interaction strength is proportional to $\frac{1}{r^3}$ where r is the distance between the interacting spins.[22]

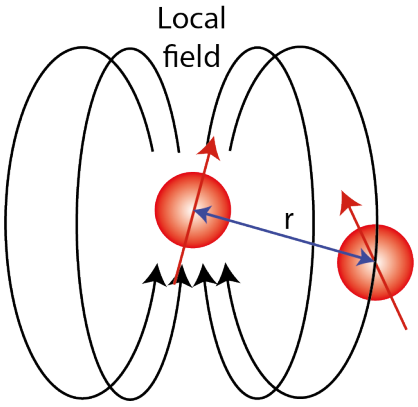


Figure 2.5: A diagram illustrating the dipolar coupling which arises from the interaction of the local field of one nuclei with the local field of a nearby nuclei separated by a distance, r.

Under a classical regime the equation for the energy of the interaction

of two magnetic dipoles, $\mu_1$ and $\mu_2$ is given by;

$$U = (\frac{\boldsymbol{\mu}_1 \cdot \boldsymbol{\mu}_2}{r^3} - 3\frac{(\boldsymbol{\mu}_1 \cdot \boldsymbol{r})(\boldsymbol{\mu}_2 \cdot \boldsymbol{r})}{r^5})\frac{\mu_0}{4\pi} \tag{2.49}$$

However in the quantum mechanical regime the magnetic moment operator is given by:

$$\hat{\boldsymbol{\mu}} = \gamma\hbar\hat{\boldsymbol{I}}. \tag{2.50}$$

When $\hat{\boldsymbol{\mu}}$ substituted in to the classical equation the dipolar coupling hamiltonian becomes:

$$\hat{H}_{dd} = -(\frac{\mu_0}{4\pi})\gamma_I\gamma_S\hbar(\frac{\boldsymbol{I}\cdot\boldsymbol{S}}{r^3} - 3\frac{(\boldsymbol{I}\cdot\boldsymbol{r})(\boldsymbol{S}\cdot\boldsymbol{r})}{r^5}). \tag{2.51}$$

If this is then transformed into spherical coordinates and the scalar products are expanded then the formula can be expressed as:[22]

$$\hat{H}_{dd} = -(\frac{\mu_0}{4\pi})(\frac{\gamma_I\gamma_S\hbar}{r^3})[A + B + C + D + E + F] \tag{2.52}$$

where

$$A = \hat{I}_z\hat{S}_z(3\cos^2\theta - 1), \tag{2.53}$$

$$B = -\frac{1}{4}[\hat{I}_+\hat{S}_- + \hat{I}_-\hat{S}_+](3\cos^2\theta - 1), \tag{2.54}$$

$$C = \frac{3}{2}[\hat{I}_z\hat{S}_+ + \hat{I}_+\hat{S}_z]\sin\theta\cos\theta e^{-i\phi}, \tag{2.55}$$

$$D = \frac{3}{2}[\hat{I}_z\hat{S}_- + \hat{I}_-\hat{S}_z]\sin\theta\cos\theta e^{-i\phi}, \tag{2.56}$$

$$E = \frac{3}{4}[\hat{I}_+\hat{S}_+]\sin^2\theta e^{-2i\phi}, \tag{2.57}$$

$$F = \frac{3}{4}[\hat{I}_-\hat{S}_-]\sin^2\theta e^{-2i\phi}, \tag{2.58}$$

The dipolar coupling can also be written using a similar form as the

chemical shift tensor such as,

$$\hat{H}_{dd} = -2\hat{\boldsymbol{I}} \cdot \boldsymbol{D} \cdot \hat{\boldsymbol{S}} \tag{2.59}$$

where **D** is the dipolar coupling tensor. When the tensor is in the PAS it has diagonal elements of $\frac{-d}{2}, \frac{-d}{2}, +d$ and d is the dipolar coupling constant given by:

$$d = \hbar(\frac{\mu_0}{4\pi})\frac{1}{r^3}\gamma_I\gamma_S \tag{2.60}$$

The tensor describes how the magnetic field of spin S felt by spin I varies with the orientation of the IS internuclear vector in a magnetic field as well as the $r^3$ dependence. The dipolar coupling tensor is traceless which means in the PAS the determinant of the tensor is zero. This means that there is no isotropic component of the the dipolar coupling. Therefore, in solution-state NMR there is no direct effect of dipolar coupling on the NMR spectrum.[22]

There are two main types of dipolar coupling; homonuclear and heteronuclear as discussed in the next two sections.

### Homonuclear Dipolar coupling

Homonuclear dipolar coupling is the coupling between spins of the same nuclei and so the nuclei involved have the same Larmor frequency. In order to understand the effect of the coupling on the NMR spectrum the Hamiltonians given in the section above need to be transformed in to a rotating reference frame that rotates at the Larmor frequency.

For the spin system, where nucleus I is coupled to nucleus S (referred to as the I-S spin system) the total Hamiltonian can be described as:

$$\hat{\boldsymbol{H}} = \hat{\boldsymbol{H}}_{\boldsymbol{0}} + \hat{\boldsymbol{H}}_{dd} \tag{2.61}$$

where $\hat{\boldsymbol{H}}_{\mathbf{0}} = \omega_0 \hat{\boldsymbol{I}}_z + \omega_0 \hat{\boldsymbol{S}}_z$ as both I and S have the same Larmor frequency in the homonuclear case. This is then transformed into the rotating Hamiltonian given by $\hat{\boldsymbol{H}}'$:

$$\hat{\boldsymbol{H}}' = \hat{\boldsymbol{R}}_z^{-1}(\omega_0 t)\hat{\boldsymbol{H}}_{\mathbf{0}}\hat{\boldsymbol{R}}_z(\omega_0 t) + \hat{\boldsymbol{R}}_z^{-1}(\omega_0 t)\hat{\boldsymbol{H}}_{\boldsymbol{dd}}\hat{\boldsymbol{R}}_z(\omega_0 t) - \omega_0 \hat{\boldsymbol{I}}_z - \omega_0 \hat{\boldsymbol{S}}_z \quad (2.62)$$

where $\hat{\boldsymbol{R}}_z(\omega_0 t)$ is the rotation operator that acts on the Hamiltonian to put it in the rotation frame rotating at a frequency of $\omega_0 t$. The term $\hat{\boldsymbol{H}}_0$ is unaffected by the rotation operator leaving only the second term as the others cancel. If the general $\hat{\boldsymbol{H}}_{dd}$ is substituted into the Equation 2.62 it gives:

$$\hat{\boldsymbol{H}}' = -d[A + B] + \hat{\boldsymbol{R}}_z^{-1}(\omega_0 t)[C + D + E + F]\hat{\boldsymbol{R}}_z(\omega_0 t) \quad (2.63)$$

where d is the dipolar coupling constant and A-F are defined as in the general dipolar coupling eq. 2.52 - 2.58. The A and B terms are not affected by the transformation into the rotating frame whilst C-F terms are time dependent. The standard method to avoid the time dependency that the C -F terms bring, is to use average Hamiltonian theory to approximate the Hamiltonian by a sum of successively higher order terms. This average theory is an accurate approximation for describing the NMR spectrum assuming that the magnitude of the dipolar coupling is small relative to the Larmor frequency, $\omega_0$; i.e. the secular approximation holds.[22] The first order contribution can then be found to be:

$$\bar{\boldsymbol{H}}_{dd}^{(0)} = -d(3\cos^2\theta - 1)[\hat{\boldsymbol{I}}_{\boldsymbol{z}}\hat{\boldsymbol{S}}_{\boldsymbol{z}} - \frac{1}{2}(\hat{\boldsymbol{I}}_{\boldsymbol{x}}\hat{\boldsymbol{S}}_{\boldsymbol{x}} + \hat{\boldsymbol{I}}_{\boldsymbol{y}}\hat{\boldsymbol{S}}_{\boldsymbol{y}})] \tag{2.64}$$
$$= -d(3\cos^2\theta - 1)[3\hat{\boldsymbol{I}}_{\boldsymbol{z}}\hat{\boldsymbol{S}}_{\boldsymbol{z}} - \hat{\boldsymbol{I}}.\hat{\boldsymbol{S}}]$$

as the time dependent terms integrate to zero in the average Hamiltonian. This first order average Hamiltonian of the homonuclear dipolar coupling is usually referred to as the secular form.

This average translates to several key effects that are visible in the NMR spectrum. The first term, $\hat{\boldsymbol{I}}_{\boldsymbol{z}}\hat{\boldsymbol{S}}_{\boldsymbol{z}}$, causes a shift in the energy of the energy levels such that they are no longer just separated by the Larmor frequency. The second term, $(\hat{\boldsymbol{I}}_{\boldsymbol{x}}\hat{\boldsymbol{S}}_{\boldsymbol{x}} + \hat{\boldsymbol{I}}_{\boldsymbol{y}}\hat{\boldsymbol{S}}_{\boldsymbol{y}})$, causes a mixing of the $\alpha\beta$ and $\beta\alpha$ terms leading to a split in the degenerate Zeeman levels over a range of $\Delta M = \pm 1$, where M is the total z component of the spin system. This means even without taking into account of orientation dependencies the slightly different energies results in very broad gaussian line shapes. This mixing of the $\alpha\beta$ terms results in broadening of the lines in the spectrum called inhomogeneous broadening.[22]

### Heteronuclear Dipolar coupling

Heteronuclear dipolar coupling is the coupling between two different nuclei such as $^{13}$C and $^{1}$H. The same rotation frame transformation can be applied here as in the homonuclear dipolar coupling case resulting in a transformed Hamiltonian of,

$$\hat{\boldsymbol{H}}^*(t) = -dA + \hat{\boldsymbol{R}}_z^{-1}(\omega_0 t)[B + C + D + E + F]\hat{\boldsymbol{R}}_z(\omega_0 t) \tag{2.65}$$

where in this case only term A remains unaffected by the rotation transformation. This means when applying the Average Hamiltonian theory the first order average Hamiltonian only has one term given by:

$$\hat{\boldsymbol{H}}_{dd}^{hetero} = -d(3\cos^2\theta - 1)\hat{\boldsymbol{I}}_{\boldsymbol{z}}\hat{\boldsymbol{S}}_z. \tag{2.66}$$

The first-order averaged Hamiltonian for the heteronuclear dipolar coupling case is further truncated compared to the homonuclear Hamiltonian as term B also becomes unimportant in the first-order representation. Without the B term in the first order Hamiltonian there is a lack of broadening caused by the heteronuclear interaction due to mixing of $\alpha$ and $\beta$ states. Instead there is an orientation dependence of the dipolar coupling which give different transition frequencies according to:

$$\omega_{dd}^I = \omega_0^I \pm \frac{1}{2}d(3\cos^2\theta - 1) \tag{2.67}$$

This means that when looking at a powder sample a line shape similar to that created by the chemical shift anisotropy is produced. However there are two transitions one which depends on $+(3\cos^2\theta - 1)$ whilst the other depends on $-(3\cos^2\theta - 1)$. This results in two line shapes overlapping as a mirror image with the same isotropic frequency leading to a line shape called a powder Pake pattern with two horns separated by the heteronuclear coupling as shown in Figure 2.6.[22]

### Quadrupolar interaction

Any NMR active nucleus with a spin, I $> \frac{1}{2}$ is referred to as a quadrupolar nucleus. Around $\frac{3}{4}$ of all NMR active nuclei are quadrupolar. A quadrupolar nucleus possesses a nuclear electric quadrupole moment which can interact with an electric field gradient at the nucleus. This means that the

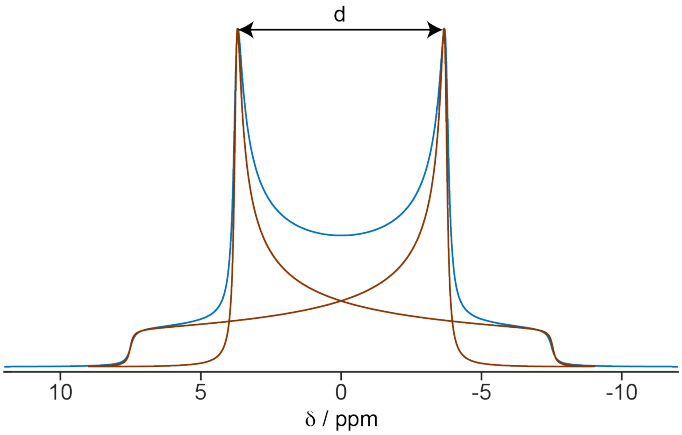


Figure 2.6: An example of the well characterised Pake powder pattern formed in the presence of the orientation dependence of the heteronuclear dipolar coupling where the dipolar coupling constant is the distance between the two 'horns'.

nucleus can not only interact with the magnetic fields present but also the electric fields that are present at the nucleus. This can affect the nuclear spin energy as well as other magnetic interactions at the nucleus. The strength of the quadrupolar interaction depends on the nuclear quadrupole moment and the electric field gradient. The electric quadrupole moment, eQ is a constant for each given nucleus and does not depend on the environment whereas, the electric field gradient does depends on the environment of the nucleus, particularly on the geometry of the bonds surrounding the nucleus.[22,24]

The quadrupolar coupling Hamiltonian is given by:

$$\hat{\boldsymbol{H}}_Q = \frac{eQ}{2I(2I-1)\hbar}\hat{\boldsymbol{I}}.\boldsymbol{V}.\hat{\boldsymbol{I}} \tag{2.68}$$

The quadrupolar anisotropy and asymmetry terms are regularly used

to describe a system and are defined as

$$eq = V_{zz}^{PAS} \tag{2.69}$$

$$\eta_Q = \frac{V_{xx}^{PAS} - V_{yy}^{PAS}}{V_{zz}^{PAS}} = \frac{V_{xx}^{PAS} - V_{yy}^{PAS}}{eq} \tag{2.70}$$

The quadrupolar interaction leads to some more complicated and varied lineshapes. However this is outside of the scope of this work.[22]

### Scalar (J) coupling

Scalar (J) coupling is an interaction between nuclei that are covalently bonded to each other. The Hamiltonian for this interaction between covalently bonded nuclei I and S is given by:

$$\hat{\boldsymbol{H}}_J = \hat{\boldsymbol{I}}.\hat{\boldsymbol{J}}.\hat{\boldsymbol{S}} = 2\pi J \hat{\boldsymbol{I}}.\hat{\boldsymbol{S}} \tag{2.71}$$

Where J is the scalar coupling constant. In solution NMR this interaction does not average away in the isotropic tumbling and causes a splitting of the NMR lines. The size of the splitting is the scalar coupling constant, J. For two nuclei of spin = 1/2 in a magnetic field, the spins can be either aligned with the field or against the field resulting in the two spin energy levels $\alpha$ and $\beta$. When these spins are covalently bonded their spin energy level affects the polarisation of the shared electrons in the covalent bond. Whether the spin states of the nuclei are parallel (i.e. $\alpha\alpha$) or antiparallel (i.e. $\alpha\beta$) leads to two slightly different energies which can be seen in the splitting of the line shape as shown in Figure 2.7.[22,24,33] This interaction tends to be of the order of a few hertz, and so generally can be ignored as the other interactions are usually an order of magnitude bigger. The larger interactions result in NMR peak line widths that are too broad to

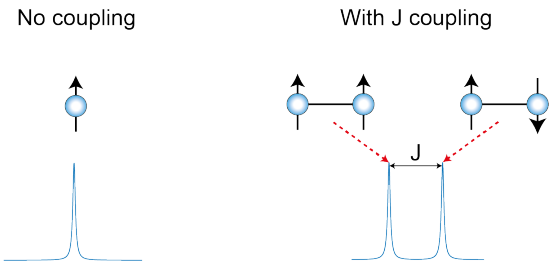


Figure 2.7: Left shows the NMR line of a single uncoupled spin. Right shows the effect of the scalar coupling of two spin-1/2 nuclei on the NMR line. The line is split into two and separated by the scalar coupling constant, J.

observe J coupling splitting. However, this interaction can be used in 2D correlation experiments such as the refocused INADEQUATE which will be described in the next chapter.[22,24]

## 2.10 Essential techniques

NMR is an inherently insensitive technique which arises from the small energy level differences created when the nuclei are placed into a large static magnetic field. In NMR, the experiment is repeated many times (scans) and the signals are added together. Enough scans are recorded to produce an acceptable signal to noise ratio. Another issue that plagues solid-state NMR in particular is the resolution. In solid-state NMR there are many anisotropic interactions present that significantly broaden the line shapes creating a broad indistinct line shape which is not useful. There are several methods such as Magic Angle Spinning (MAS), Cross Polarisation (CP) and Decoupling that have been developed for solid-state NMR.[22] These techniques significantly improve the sensitivity and resolution of solid-state NMR. These are now used routinely in solid-state NMR, and the improvements they have made allows solid-state NMR to be used to study a wide range of complex crystalline and amorphous solid systems such as pharmaceuticals, polymers, plants and catalysts.[4,29–31,37]

### 2.10.1 Magic angle spinning

Solid-state NMR has been plagued with low resolution due to powder patterns of the different environments overlapping resulting in broad and indistinctive line-shapes. As stated previously, all internal hamiltonians can be written in the form of a product of a spatial contribution and spin contribution. The anisotropic chemical shift hamiltonian for example can be written as

$$H = \frac{1}{3}\left\{\sum_{j=1}^{3}\sigma_j + \sum_{j=1}^{3}(3\cos^2\theta_j - 1)\sigma_j\right\}\gamma I_z B_0 \tag{2.72}$$

The spatial term in the Hamiltonian is proportional to $3\cos^2\theta - 1$ where $\theta$ is the angle that describes the transformation from the PAS frame to the laboratory reference, whilst the spin term is proportional to $I_z$.

Magic Angle Spinning (MAS) was first introduced in 1958 by E. R. Andrew *et al.*[23] which mimics the rapid isotropic tumbling seen in solution. It does this by taking advantage of the proportionality constant of $3\cos^2\theta - 1$ in the spatial term of the anisotropic interactions. This constant goes to zero at at angle of 54.74° which is known as the magic angle. In a powder sample which spinning at the magic angle, all angles of $\theta$ and $\beta$ are present, where $\beta$ describes the transformation of the PAS frame into the rotating reference frame, so when the sample is spinning the motionally averaged Hamiltonian becomes proportional to,

$$\bar{H} \propto \int_0^{\pi}(3\cos^2\theta - 1)\sin\theta d\theta = \frac{1}{2}(3\cos^2\beta_R - 1)(3\cos^2\beta - 1) \tag{2.73}$$

where the angle $\beta_R$ acts as a scaling factor for the motionally averaged Hamiltonian which goes to zero at the magic angle.[22,24]

The MAS rate depends on the sample and the information required

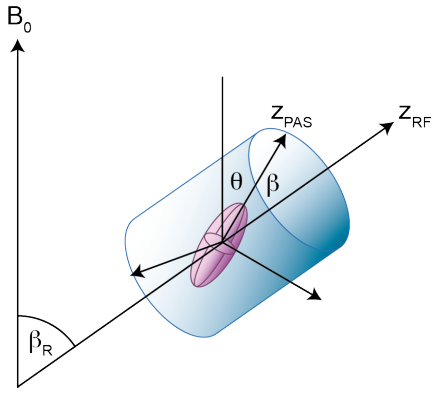


Figure 2.8: A diagram of a rotor (blue) containing a sample (pink) which are at the magic angle, $\beta_R$ with respect to the static field, $B_0$. The angles $\theta$ and $\beta$ describe the transformation of the PAS frame into the laboratory frame and the rotating frame, respectively.

from the sample. If the sample does not rotate at a frequency larger than the anisotropy involved then spinning sidebands arise at multiples of the spinning speed which follow the shape of the powder pattern, as shown in Figure 2.9. These can cause a number of issues as they can overlap with other chemical shifts. In order to identify the spinning sidebands, experiments are usually repeated with different MAS rates so the the spinning sidebands move whilst the isotropic chemical shifts do not.[24] If it is not possible to spin fast enough to remove obtrusive spinning sidebands then pulse sequences such as TOSS (Total Spinning Sideband Suppression) have been developed to remove the sidebands.[38] This pulse sequence is very useful. However, care is needed as this removes intensity of the sidebands but does not put it back to the central isotropic shift peak and so the resulting spectrum is not quantitative.

For protons the homonuclear dipolar coupling tends to be extremely large such that even high speed magic angle spinning does not completely remove its broadening effects. Recent developments of probes with ever decreasing sizes and increasing spinning speeds are beginning to approach point where even these large interactions could be averaged away producing

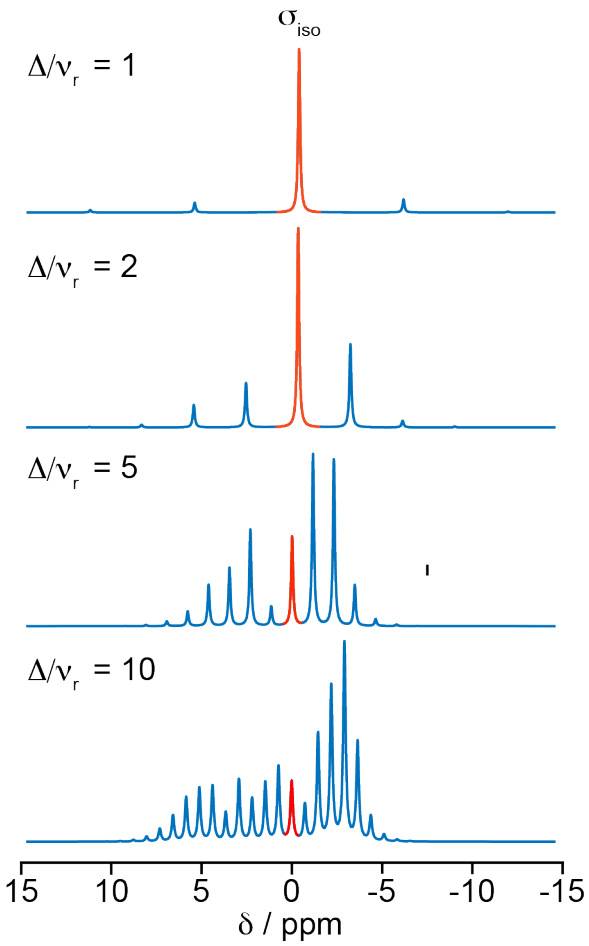


Figure 2.9: An example of a spectrum with an anisotropy of 10 ppm and asymmetry of 0 at different MAS rates.

solution-state NMR like resolution. Magic angle spinning does not average away the homogeneous broadening of heteronuclear dipolar coupling which is where another essential technique called Decoupling was developed.

### 2.10.2 Decoupling

Dipolar coupling causes significant broadening of the NMR spectrum. Magic angle spinning, if fast enough, can average much of this interaction away. Unfortunately, it is not always possible to spin fast enough to completely remove the dipolar coupling so decoupling is needed to acquire a high resolution NMR spectrum.

Only heteronuclear decoupling was used in the NMR experiments in this work and so will be the main focus of this section. Heteronuclear decoupling was developed as a means of overcoming the broadening caused by heteronuclear dipolar coupling. This is routinely used alongside MAS to increase the resolution of the isotropic shifts. For example, when acquiring a $^{13}$C CP MAS spectrum, the $^{1}$H channel is constantly irradiated with a high power rf pulse during acquisition to decouple the heteronuclear dipolar coupling between the proton and carbon nuclei. There are several different types of decoupling schemes developed over the years and the more routinely used methods are described below.

**High power Decoupling**

This is the simplest of the decoupling techniques which involves continuous high power irradiation at the frequency of the abundant spin Larmor frequency. The dilute spin pulse sequence can be applied and the NMR signal acquired whilst the constant rf irradiation is applied to the abundant nucleus. This decouples the interaction as the continuous irradiation effectively causes rapid transitions between the $\alpha$ and $\beta$ levels at a rate that is determined by the strength of the irradiation. This rapid transition between the levels causes the time dependent Hamiltonian for the dipolar coupling to average away to zero if the transition is oscillating fast enough. Generally the strength of the rf irradiation needs to be at least three times the strength of the largest heteronuclear dipolar coupling to completely remove its effect.[22]

**Two pulse phase modulation (TPPM) and SPINAL-64 Decoupling**

These pulse sequences are more involved than the high power decoupling but are more efficient at decoupling for the same power input which puts less strain on the NMR probes. TPPM decoupling is made up of 2 pulses with a pulse length of $\theta_t$ and alternating phases, $(\pm\phi)$. For the optimum decoupling using this method, the transverse excitation is made near resonant which means that:

$$\frac{\omega_p}{2\pi} \approx \omega_{rf} \cos(\phi/2) \tag{2.74}$$

where $\omega_p$ is the frequency of the phase switching given by $2\pi/2\theta_t$ and $\omega_{rf}$ is the rf frequency used. This condition is met if $\theta_t$ is close to a $\pi$ pulse length when $\phi$ is a small angle. The pulse length and phase have to be optimised in each experiment as it they depend on the sample and the MAS rate. The optimum pulse length and phase is usually found to be around 170° and between 10-70° respectively. A train of these two pulses is then applied during acquisition to continuously generate rapid transitions which decouple the heteronuclear dipolar coupling of the abundant nucleus.[39] The line widths achieved by this decoupling scheme can be half that of constant wave decoupling and the method works better at higher spinning speeds.

The small phase incremental alternation with 64 steps (SPINAL-64) decoupling method is an improved version of the TPPM sequence. By changing the phase of the two pulses $(\theta_t(\pm\phi))$ in a stepwise manner to give

two basic elements Q and $\bar{Q}$,

$$Q = \theta_t(10^\circ)\theta_t(-10^\circ)\theta_t(15^\circ)\theta_t(-15^\circ)\theta_t(20^\circ)\theta_t(-20^\circ)\theta_t(15^\circ)\theta_t(-15^\circ) \tag{2.75}$$

$$\bar{Q} = \theta_t(-10^\circ)\theta_t(10^\circ)\theta_t(-15^\circ)\theta_t(15^\circ)\theta_t(-20^\circ)\theta_t(20^\circ)\theta_t(-15^\circ)\theta_t(15^\circ) \tag{2.76}$$

improves the efficiency of the decoupling sequence. These new basic elements can then be super-cycled to improve the efficiency further. SPINAL-64 is made up of 64 pulses in a total cycle which is made up of the basic elements, Q and $\bar{Q}$, in the following way,

$$SPINAL-64 = Q\bar{Q}\bar{Q}Q\ \bar{Q}QQ\bar{Q} \tag{2.77}$$

There are shorter and longer versions of the SPINAL decoupling sequences as it just depend on how many basic elements are used in the super-cycling. SPINAL-64 is most commonly used as it was found to be slightly better than SPINAL-32 and there was no further improvements made by going to SPINAL-128.[6]

### 2.10.3 Cross Polarisation

NMR is an inherently insensitive technique and nuclei such as $^{13}$C have low natural abundance compounding this issue further. Cross polarisation (CP) initially developed in year 1972 by A. Pines *et al.*[26] was developed to improve the sensitivity of dilute and low gamma nuclei. CP transfers magnetisation from abundant nuclei, such as protons, to less abundant nuclei. This is achieved by the CP experiment shown in Figure 2.10.[22,24,26]

A 90° pulse is applied on the abundant nucleus to put the magnetisation of I (usually $^1$H) in to the transverse plane. This magnetisation is kept

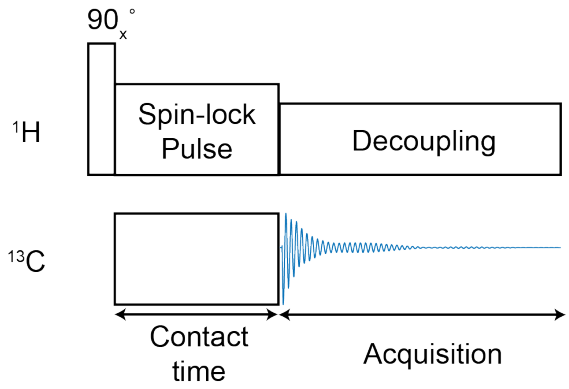


Figure 2.10: The Cross polarisation experiment in the case of transferring magnetisation from proton to carbon nuclei.

along the -y axis by a spin lock pulse. A pulse is also applied to the less abundant nucleus at the same time. The spin lock pulse and the pulse on the less abundant nucleus has to obey the Hartmann-Hahn matching condition given by

$$\gamma_H B_1(^1H) = \gamma_X B_1(X) \tag{2.78}$$

in order to enable the transfer of magnetisation to the less abundant spins whilst maintaining the conservation of energy. The matching condition ensures that the energy levels of both nuclei involved are the same enabling the transfer of magnetisation. This transfer of magnetisation from protons to carbons in standard organic solids can lead to an increase in sensitivity of up to $10^3$ times compared to the NMR signal of the carbon nucleus directly.[26] The spin-lattice relaxation in the rotating frame relates to the mechanism of the relaxation of the transverse magnetisation during an rf spin-lock pulse like that used in a CP experiment. This relaxation follows the same exponential relationship as its none rotating counterpart. The time constant, $T_{1\rho}$ defines the transverse relaxation of the magnetisation on a nuclei in a rotating frame and is a key parameter to measure for each sample to ensure optimum CP transfer. The time at which the experiment

can be repeated is usually a lot shorter than direct detection of dilute nuclei as the repetition rate depends on the $T_1$ relaxation time of the abundant protons which tend to be shorter.

The application of the CP experiment under MAS requires an adjustment to the Hartmann-Hahn condition to

$$\gamma_H B_1(^1H) = \gamma_X B_1(X) \pm n\nu_R \tag{2.79}$$

where n is an integer and $\nu_R$ is the spinning speed.

In order to broaden this matching condition a ramp CP experiment was developed.[40] Whilst more sensitive, CP does not produce a quantitative spectrum as its process is driven by dipolar coupling. Therefore, you mobile environments where this is averaged away and quaternary carbons such as carbonyls tend to be under represented. However, despite its non-quantitative nature this method is routinely used in NMR due to the large increases in sensitivity and the dramatic reduction in experimental times.

# Chapter 3

# Experimental Methods

One benefit of modern NMR is the extensive range of experiments available which can be acquired on standard equipment by just varying the pulses and delays that define the pulse sequence. These experiments allow a wealth of information about samples to be extracted such as chemical environments, proximities and dymanics. In this chapter the main experiments used in the work presented following chapters of this thesis are introduced.

## 3.1 1D NMR

The simplest NMR experiment is the one-pulse experiment, which involves applying a rf pulse to rotate the magnetisation into the xy plane where it can be detected by the receiver. When the magnetisation precesses about the xy plane it induces a current and hence a voltage in the coil that surrounds the sample. This voltage is the NMR signal which is detected in the form of a Free Induction Decay (FID) during the acquisition time of the experiment.[24] The NMR signal from the sample is very small ($\mu$V) and is amplified by the preamplifier, and then digitised by the analogue to

digital converter (ADC).

The NMR signal S(t) can be described as a complex signal with real and imaginary parts given by:

$$S(t) = S_0 \exp(i\omega t) = S_0 \cos \omega t + iS_0 \sin \omega t \tag{3.1}$$

where $\omega$ is the precession frequency. The signal decays over time due to spin-spin relaxation and so the NMR signal can be fully described by,

$$(t) = S_0 \exp(i\omega t) \exp(\frac{-t}{T_2}) \tag{3.2}$$

where $T_2$ is the time constant for the relaxation. The acquired FID is a time domain signal and is usually very complicated as there can be many different precession frequencies within a sample. Therefore, the signal is normally Fourier transformed into the frequency domain to produce the NMR spectrum. When only the real component of the FID is Fourier transformed it produces a absorption mode Lorenzian line shape whilst the imaginary component produces a dispersion mode Lorentzian as illustrated in Figure 3.1.[41]

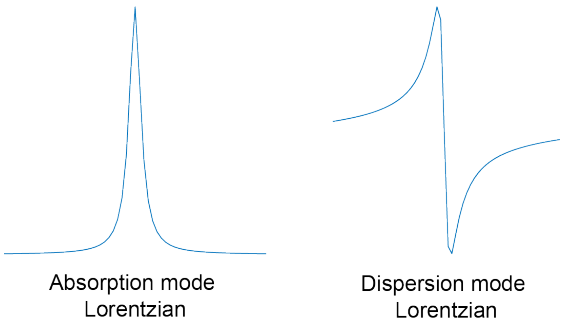


Figure 3.1: A diagram showing the absorption mode and dispersion mode Lorentzian lineshapes produced from Fourier transforming the real and imaginary components of the FID.

The width of the NMR line is determined by how quickly the signal

decays away. Hence, the intrinsic line width at half its maximum height is given by $1/\pi T_2$ Hz. Therefore the faster the signal decays away the broader the NMR line.

### 3.1.1 Phasing and Processing an NMR spectrum

The real and imaginary components of the FID are 90° out of phase which means that if, at t = 0, the real component is at maximum intensity, the imaginary component is at zero. If the real component is at its maximum intensity at t = 0, then the Fourier transform of it is completely absorption mode Lorenzian and conversely the imaginary component is completely dispersion mode Lorentzian.[41]

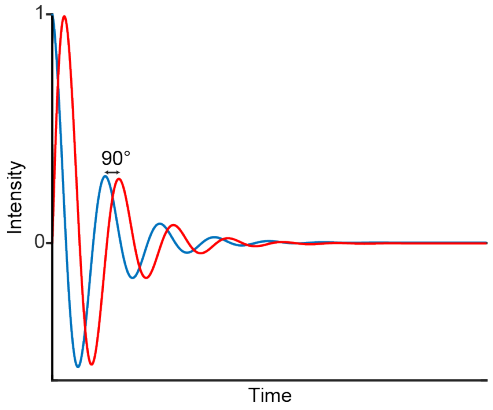


Figure 3.2: A diagram showing the real (blue) and imaginary (red) components of the FID showing that they are 90° out of phase from one another.

In NMR, it is not always possible to acquire the FID when the maximum of the real component is at t = 0 which leads to the line shape being a mixture of absorption and dispersion mode Lorentzian. In this case the signal is said to be phase shifted, or has a phase error, which can be included in the equation for the signal to give;

$$S(t) = S_0 \exp(i\phi) \exp(i\omega t) \exp(\frac{-t}{T_2}) \tag{3.3}$$

where $\phi$ is the phase shift of the spectrum. In order to correct the phase error, the function $\exp(-i\phi)$ can be applied to spectrum which is referred to as a zero-order phase correction. Having the all the peaks in the spectrum in absorption mode Lorentzian provides the better resolution, and so, this phase correction is routinely done on every NMR spectrum.[41]

In an NMR experiment, there is a time before acquisition and after the end of the rf pulses which is referred to as the 'dead time'. This delay before acquisition ensures any effects of the high voltage rf pulses, such as ringing of the probe, have dissipated and so do not interfere with the low voltage signal that is measured from the sample.[24,41]

The dead time delay creates another phase error which is referred to as a first-order frequency dependent phase shift. This arises as resonances that are further away from the transmitter frequency are able to evolve more during the dead time. This phase error can be corrected by applying another function to the spectrum called first-order phase correction or use a spin echo experiment to remove the majority of the first order correction required by using a 180° pulse to refocus the resonances that evolve at different rates. Normally a mixture of both zero and first order corrections are required to generate an in phase NMR spectrum.[24]

Zero filling involves the addition of extra zero points to the end of an FID to improve the digital resolution of the NMR spectrum. The extra points in the FID means that the NMR peaks are represented by more points in the spectrum. The maximum digital resolution is achieved by zero filling to twice the number of points acquired.[24,41]

In an FID the signal is strongest at the start and then decays away until there is just noise. The majority of the noise in the FID comes from the thermal noise of the coil that detects the signal. A minor component of the noise can also come from the amplifiers and electronics in the spectrometer.

If the FID is acquired for longer than it takes for the signal to decay then this additional noise is carried through into the spectrum leading to a poor signal to noise ratio. Therefore, the acquisition time need to be chosen to give the optimal resolution for the sample.[41] Line broadening is commonly used to remove some of the noise in the FID and to improve the appearance of the signal in the spectrum. A weighted function is applied to the FID that decays towards zero and removes some of the noise in the FID resulting in an improved signal to noise ratio in the spectrum. The most commonly used line broadening function in 1D NMR experiments is an exponential function. This weighted function forces the signal to decay faster which increases the line width of the spectrum. Too much line broadening can decrease the resolution of the spectrum, distort the peak line shapes and reduce the signal to noise ratio of the NMR spectrum.[24] It has been found that applying a weighing function with the same line width as the signal in the spectrum produces the optimum signal to noise ratio. This type of weighting function is referred to as a matched filter.[41]

There are many alternative functions that can be used on the FID to improve the signal to noise as well as the resolution, such as a Gaussian function. However, these have not been used in this work.

### 3.1.2 Phase Cycling

NMR experiments are described by their pulse sequence which define the pulses and delays for the experiment. In addition to the pulse sequence it is important to know the phase cycling for the experiment. Phase cycling is where the experiment is repeated several times with a systematic variation of the phases of the rf pulses and the receiver phase to select a particular signal to acquire. The phase cycle selects a particular set of coherences at each pulse to follow the desired coherence transfer pathway

for the experiment.[42]

## 3.2 Quantitative NMR

The main way to acquire a fully quantitative NMR spectrum is to allow the nuclei to relax completely between each scan. To ensure that at least 99.9% of the sample has relaxed to equilibrium a relaxation delay between experiments has to be at least 5 times the longest $T_1$ relaxation time. For proton NMR, this is usually achievable as they usually have short $T_1$ times of a few seconds. However, acquiring a quantitative $^{13}C$ NMR spectrum is more difficult as the abundance is very low and the relaxation times can be of the order of hundreds of seconds making acquiring a single 1D spectrum unreasonably long.[10,43] For $^{13}C$ NMR assigning the chemical environments in the sample is commonly prioritised over quantification, therefore a CP-MAS experiment is usually used due to its increase in sensitivity. When acquiring a standard $^{13}C$ NMR spectrum with just a 90° rf pulse and $^{1}H$ decoupling during acquisition, literature refers to it as a Direct Polarisation (DP) experiment in order to distinguish it from a standard CP experiment.

## 3.3 The Schmidt-Rohr Quantitative Method

The Schmidt-Rohr quantitative method was initially developed in 2000 to study the crystallinity of polyethylene fibers and has since been used to obtain a composition analysis of a wide range of humic acids.[10,44] This method is based on using a CP experiment with a z-filter to measure the $^{13}C$ $T_1$ relaxation times throughout the sample. Initially a $^{13}C$ DP MAS spectrum is acquired using a recycle delay of around 1.3 times the longest $^{13}C$ $T_1$ time in the sample. This recycle delay maximises the sensitivity of

the longest $^{13}C$ $T_1$ contribution for a given measurement of time; however, this does not produce a quantitative spectrum.

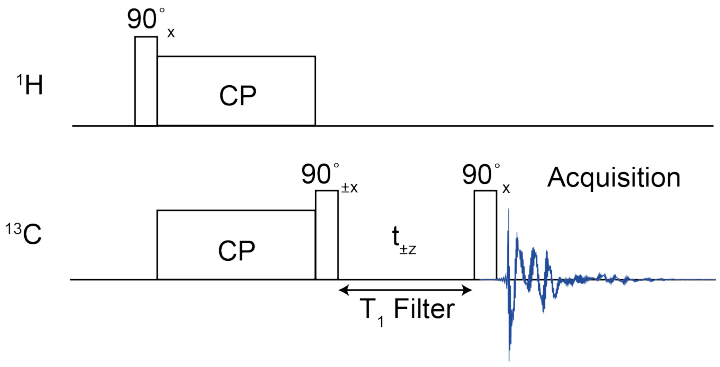


Figure 3.3: The CP / $T_1$ experiment pulse program used is the Schmidt-Rohr method.

The proportion of the signal missing in the $^{13}C$ DP MAS spectrum can be calculated by measuring the $T_1$ relaxation for each site. This is achieved by comparing the intensities of two spectra acquired using the CP/$T_1$ experiment shown in Figure 3.3 with two different delays for the $T_1$ filter time, $t_{\pm z}$. The first $t_{\pm z}$ delay is set as short as possible to achieve a spectrum with the maximum amount of signal whilst the second CP/$T_1$ spectrum is acquired using a $t_{\pm z}$ delay equal to the recycle delay of the initial $^{13}C$ DP MAS spectrum. The proportion of remaining signal in the long $t_{\pm z}$ delay CP/$T_1$ spectrum is equal to the proportion of signal missing in the $^{13}C$ MAS spectrum as shown in Figure 3.4.[10] Both the CP/$T_1$ spectra and the $^{13}C$ MAS spectrum are fitted closely with the same set of Gaussian peaks and the peak regions of the spectrum are defined. The peak areas for each region in all three spectra are calculated and the peak areas of the $^{13}C$ MAS spectrum can be corrected using the proportion calculated from the 2 CP/$T_1$ experiment peak areas.

The $^{13}C$ MAS peak signals are known to follow a relaxation curve shown in Figure 3.4 and are given by:

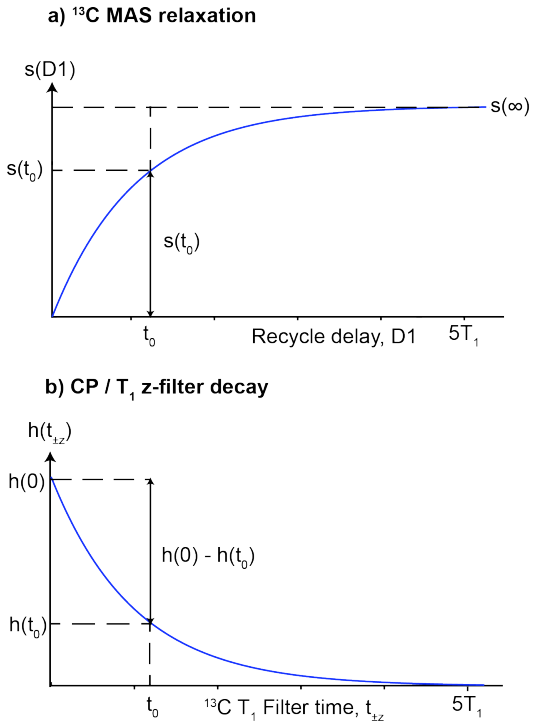


Figure 3.4: A schematic showing how the decay of the CP / $T_1$ experiment can be used to calculate the missing proportion of signal from the incompletely relaxed $^{13}$C MAS spectrum.

$$S(D1) = S(\infty) \times (1 - r(D1)) \tag{3.4}$$

where S(D1) is the signal acquired at a certain recycle delay and S($\infty$) is the signal when the sample has undergone complete relaxation. In general, the $T_1$ relaxation curve for each different chemical environment is an exponential curve making r(D1) = exp(-D1/$T_1$). However, the samples studied in this work have overlapping signals of different environments resulting in an r(D1) that is not a single exponential. Therefore the general case has been considered here. The signal change of the CP/$T_1$ spectra with a certain $t_{\pm z}$ delay is given by a decay shown in Figure 3.4b and is described by equation:

$$h(t_{\pm z}) = h(0) \times r(t_{\pm z}) \tag{3.5}$$

where h(0) is the signal at the shortest z delay and h($t_{\pm z}$) is the signal at the long z delay. In this experiment the longer $T_1$ filter time is equal to the recycle delay in the $^{13}$C MAS experiment indicated as $t_0$ in Figure 3.4 meaning that,

$$r(D1 = t_0) = r(t_{\pm z} = t_0) = h(t_{\pm z} = t_0)/h(0) \tag{3.6}$$

By substituting the $h(t_{\pm z} = t_0)/h(0)$ into Equation 3.4 the corrected signal of the $^{13}$C MAS peak areas can be calculated as:

$$S(\infty) = S(D1 = t_0) \times \frac{h(0)}{h(0) - h(t_{\pm z} = t_0)} \tag{3.7}$$

This means that while this method does not assume a single exponential decay it does make the assumption that the components that contribute to a peak have the same CP transfer efficiency. This assumption, is in the majority of cases, acceptable as sites with the same chemical shift tend to be in very similar chemical environments. In order to minimise the error involved in this quantitative analysis method the recycle delay used needs to be at least 1.3 times the longest $T_1$ relaxation time to ensure peak area corrections are less than 25%.[10]

## 3.4 2D NMR experiments

There is an extensive library of 2D experiments which have been designed to extract information about a sample that is not possible to acquire in a 1D spectrum. Two-dimensional correlation experiments such as the DARR (Dipolar Assisted Rotary Resonance) experiment correlates nuclei that are close in proximity to each other.[45] These types of experiments are useful for understanding the structure and architecture of materials as well as

resolving different sites that otherwise overlap in the 1D spectrum. Other 2D experiments can reintroduce interactions such as the CSA and dipolar couplings that MAS average away to extract further information about the chemical environment whilst still resolving the isotropic shifts.[24] In this work several different 2D experiments have be used and are described in more detail in the following sections.

### 3.4.1 Refocused INADEQUATE

The refocused INADEQUATE experiment was first implemented in solution-state NMR and has, in recent years, been utilised in solid-state NMR.[46] INADEQUATE is an acronym for Incredible Natural Abundance DoublE QUAntum Transfer Experiment. It is a 2D double quantum experiment which uses scalar (J) coupling to mediate magnetisation transfer between directly bonded nuclei.[46] The use of scalar coupling is unusual for solid-state NMR as scalar coupling tends to be orders of magnitude smaller than the more commonly used through space dipolar coupling.[22] However by using scalar coupling it is possible to unambiguously determine covalently bonded nuclei.

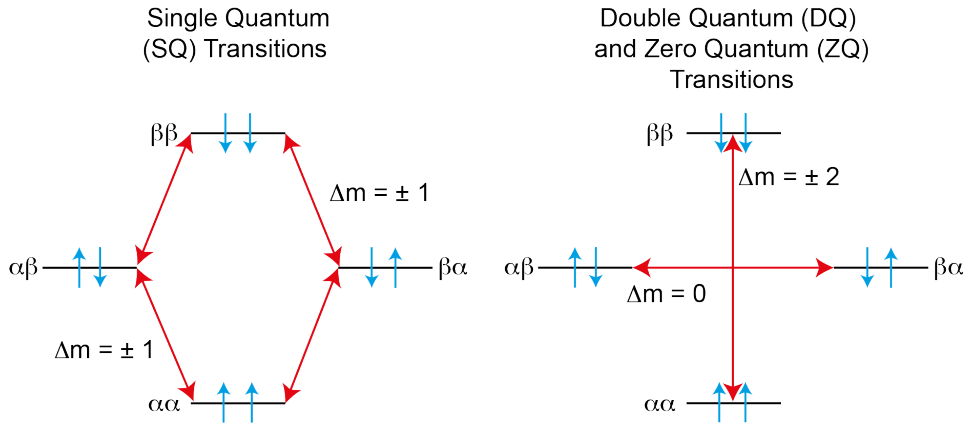


Figure 3.5: A diagram illustrating the Single Quantum (SQ) and Double Quantum (DQ) coherence transitions.

In NMR experiments only Single Quantum (SQ) coherence can be di-

rectly detected and Double Quantum (DQ) coherence is a normally forbidden transition; both of which are illustrated in Figure 3.5. DQ coherence can only be detected indirectly as a 2D experiment whereby it is generated and then converted to a SQ coherence to be detected.[46]

In an INADEQUATE experiment the SQ coherence produced for detection is an anti-phase multiplet whose components are separated by the scalar coupling. When in solution, where the multiplet components have line widths of a few hertz, this anti-phase signal can be resolved. However, in solids with line widths an order of magnitude broader than the scalar coupling, this anti-phase signal results in a cancelling of intensities leading to drastic loss of signal as seen in Figure 3.6. In a refocused INADEQUATE this anti-phase component is converted in to an in-phase signal before detection, which can lead to a large increase in intensity of the broad peaks in solid-state NMR.[46]

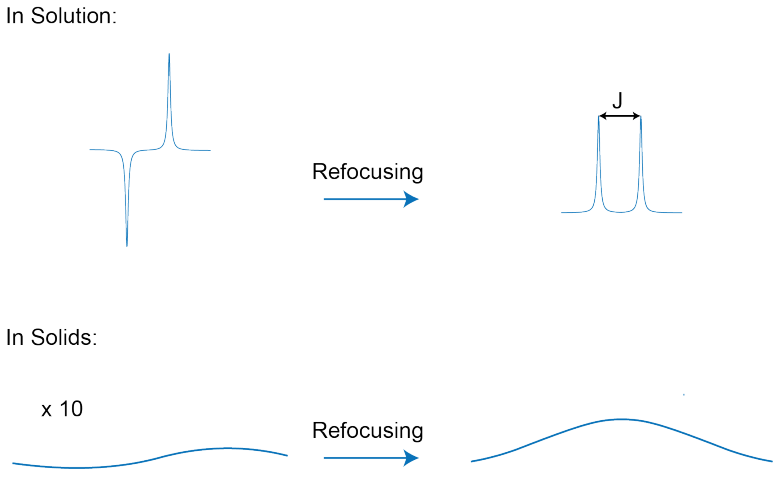


Figure 3.6: The effect of 'refocusing' in the INADEQUATE experiment for solution-state NMR where the lines have widths of a couple of hertz and for solid-state NMR, where the lines are hundreds of hertz wide.

The pulse sequence for the refocused INADEQUATE is shown in Figure 3.7 along side the coherence transfer pathway for the experiment. After CP the first 180° pulse is used to refocus the chemical shift. The first 90° pulse generates the DQ coherence which evolves during the $t_1$ period and then

is transferred back to single quantum coherence by the second 90° pulse. The final $\tau$-180°-$\tau$ is the refocusing of the anti-phase signal to an in-phase signal ready for detection.[46]

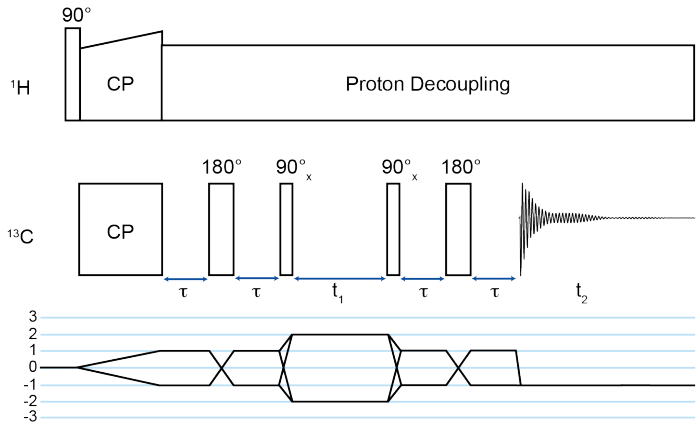


Figure 3.7: The refocused INADEQUATE pulse sequence and coherence transfer pathway.

Implementing this experiment can be difficult, with the appearance of various artefacts in the spectrum. One of the main artefacts which is seen in labelled samples is a 'relayed' peak, which arise from two non-bonded nuclei that are connected via a central nucleus. Another artefact is the interference of some anti-phase coherence which leads a phase twisted spectrum. These artefacts can be minimised by placing a z-filter at the end of the pulse sequence which aids with the refocusing of these antiphase coherences, resulting in the pulse sequence in Figure 3.8.[47]

In the refocused INADEQUATE spectrum the NMR shifts are separated over a second dimension where directly bonded nuclei are at the same double quantum frequency which is given by the sum of the isotropic shifts of these nuclei as illustrated in Figure 3.9. As this experiment is mediated by the J coupling there are no correlations between close, but non-bonded, nuclei.

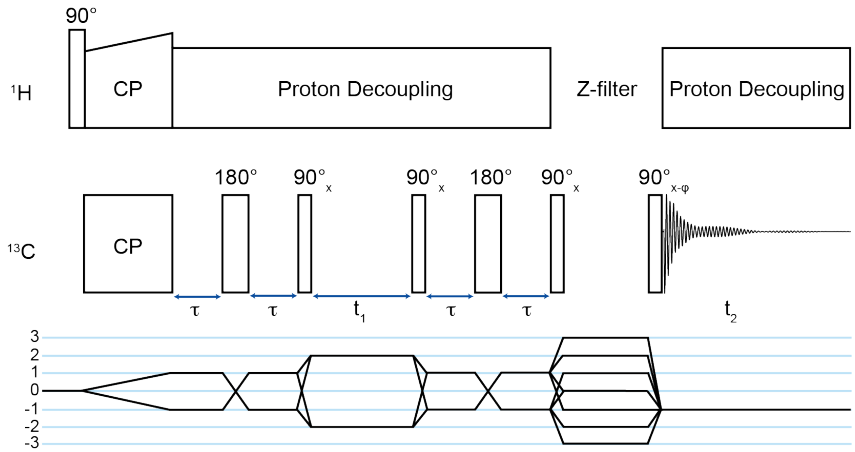


Figure 3.8: The refocused INADEQUATE with the addition of the z-filter pulse sequence and coherence transfer pathway.

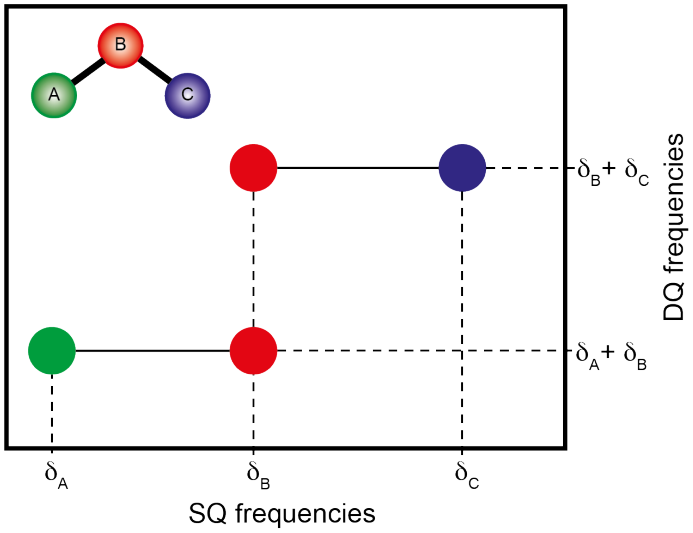


Figure 3.9: A example of a refocused INADEQUATE spectrum for a three atom system. Only the nuclei that are covalently bonded together have cross peaks.

### 3.4.2 Symmetry-Based Pulse sequences

Symmetry based pulse sequences are a set of experiments that use both the rotation of the sample and the rotation of the spins via rf irradiation to manipulate and reintroduce interactions such as the chemical shift anisotropy (CSA). Before the 1980s the majority of solid-state NMR experiment developments treated the rotation of the sample in space and the rotation of nuclear spins separately. In most cases the focus was on the rotation of

nuclear spins and the effect of MAS was treated as a small perturbation.[48] This separation was possible due to the large difference in time scales between the sample rotation and the nuclear spin rotation. It was in the 1980s where it was found that the combination of these two physical manipulations lead to useful effects. The REDOR (Rotational Echo Double Resonance) experiment was first introduced in 1989 which used the MAS rotation alongside carefully timed rf pulses to recouple $^{13}$C - $^{15}$N dipolar coupling of alanine.[49] These experiments also brought the idea of rotary resonance where the nutation frequency of the rf pulse is equal to a multiple of the spinning speed. When the rotary resonance condition is met it generates a recoupling effect of both the heteronuclear dipolar coupling, as well as, the chemical shift anisotropy.

### 3.4.3 R-sequences

R-sequences take advantage of the properties of the average nuclear spin Hamiltonian when in the presence of sample rotation and nuclear spin rotation.

Physical rotation of the sample leads to a time dependent path of the Euler angles, $\Omega_R = \alpha_R, \beta_R, \gamma_R$ and the rf pulses rotate the nuclear spins creating a time dependent path of the Euler angles $\Omega_{rf} = \alpha_{rf}, \beta_{rf}, \gamma_{rf}$. where the Euler angles describe the orientation of the sample in a particular reference frame. Every interaction has a space rank, l and a spin rank, $\lambda$ as described in Table 3.1. The rotation of the sample leads to a mixing of 2l+1 space components, m which take values -l, -l+1, ..., +l. Whilst the rotation of the nuclear spins cause a mixing of $2\lambda$+1 spin components, $\mu$ which likewise takes values -$\lambda$, -$\lambda$+1, ..., +$\lambda$. In the presence of both these rotations the resulting average Hamiltonian, $\hat{H}^{\Lambda}$ for a particular interaction can be described using a superposition of the spin and space components

as defined by the quantum numbers m and $\mu$,

$$\hat{H}^{\Lambda} = \sum_{m=-l}^{+l} \sum_{\mu=-\lambda}^{+\lambda} \hat{H}^{\Lambda}_{lm\lambda\mu} \tag{3.8}$$

In the case of MAS the component l = 2 and m = 0 is zero due the following property of the magic angle:

$$d^2_{00}(\theta_{magic}) = 0 \tag{3.9}$$

| Interaction | Space rank, l | Space component, m | Spin rank, $\lambda$ | Spin component, $\mu$ |
|---|---|---|---|---|
| Isotropic chemical shift | 0 | 0 | 1 | {-1, 0, 1} |
| Chemical shift anisotropy | 2 | {-2, -1, 1, 2} | 1 | {-1, 0, 1} |
| Scalar (J) coupling | 0 | 0 | 0 | 0 |
| Dipole-Dipole coupling | 2 | {-2, -1, 1, 2} | 2 | {-2, -1, 1, 2} |

Table 3.1: Space and Spin ranks and components for the homonuclear interactions of a diamagnetic spin-1/2 system under MAS rotation.

R-sequences were developed by Levitt *et al.*[48] and use defined symmetry principles to reintroduce (recouple) particular nuclear interactions that have been averaged away by MAS. These sequences are written as $\mathrm{RN}^{\nu}_{n}$ where N, n and $\nu$ are called symmetry numbers of the sequence and are all even. The symmetry numbers for the R-sequence define a set of selection rules which determines which interactions of the average Hamiltonian are recoupled or suppressed by the R-sequence experiment.[48]

### 3.4.4 Constructing $\mathbf{RN}^{\nu}_{n}$ sequences

First a sequence of rf pulses that rotate the nuclear spins about the x axis by 180° is selected and called element $\mathcal{R}$. Next an element $\mathcal{R}'$ is defined as an 180° pulse with the rf phases set to have the opposite signs (i.e. phase inverted) to element $\mathcal{R}$. Now the rf amplitude is set such that an

even number, N, of $\mathcal{R}$ elements fit in the same time as n number of rotor periods. Finally the $RN_n^\nu$ sequence is made up of N/2 pairs of $\mathcal{R}_\phi\mathcal{R}'_{-\phi}$ elements where the phase shift is given by $\phi = \frac{\pi\nu}{N}$ radians.

### 3.4.5 $RN_n^\nu$ First-order Selection Rules

The defined symmetry properties of the Euler angles for R-sequences mean that a selection rule can be shown to apply such that:

$$\bar{H}^{(1)}_{lm\lambda\mu} = 0 \; if \; mn - \mu\nu \neq \frac{N}{2} Z_\lambda \tag{3.10}$$

where $Z_\lambda$ is 0, $\pm2$, $\pm4$, ... if $\lambda$ is even and $\pm1$, $\pm3$, $\pm5$, ... if $\lambda$ is odd.

A useful way of visualising what this selection rule means for different R-sequences is using Space-Spin Selection diagrams (SSS diagrams). These are a quick and convenient way to find which interactions are suppressed or recoupled by each R-sequence. For example, the main R-sequences used in this work is $R12_5^4$ and the respective SSS diagrams show whether a component of each interaction is symmetry allowed or forbidden by the sequence. In this case, it can be seen that the chemical shift anisotropy is symmetry allowed and so is recoupled by this sequence whereas the homonuclear dipolar coupling is forbidden and so suppressed by this sequence's symmetry.

The SSS diagrams show that only certain parts of a particular interaction are symmetry allowed by the R-sequence. This means when recoupling an interaction using these experiments, it leads to only a fraction of the interaction being reintroduced in the experiment. Thus, when measuring the recoupled interaction from experimental data there is a scaling factor, $\kappa_{lm\lambda\mu}$has to be taken into account to get the true interaction strength which can be calculated by the following equation:[48]

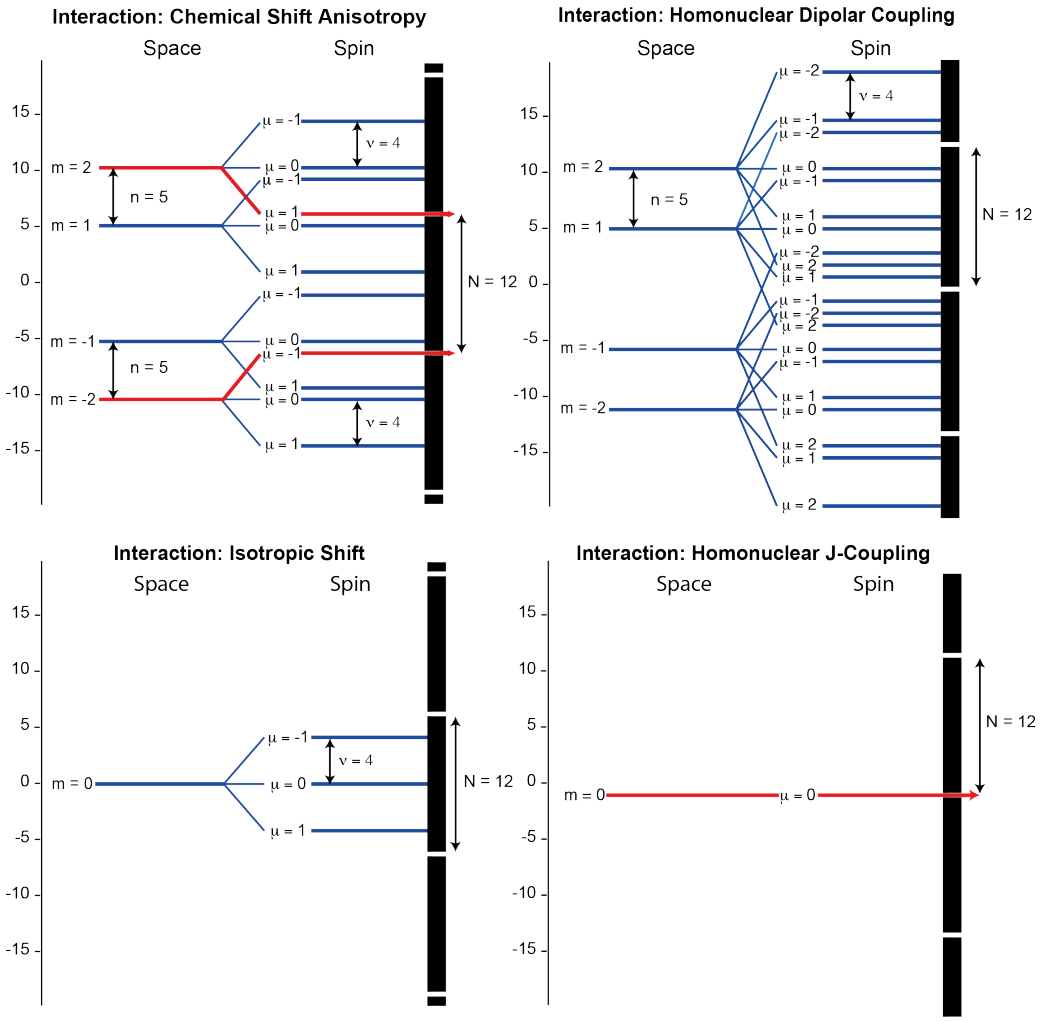


Figure 3.10: SSS diagrams for the $R12^4_5$ for all the homonuclear interactions for a diamagnetic spin 1/2 spin system. Showing that only the chemical shift anisotropy and the homonuclear J coupling is recoupled using this symmetry whilst the homonuclear dipolar coupling and the isotropic shift is suppressed.

$$\kappa_{lm\lambda\mu} = \tau^{-1} d^l_{m0}(\beta_{RL}) \times \int_0^\tau dt\; d^\lambda_{\mu 0}(-\beta_{rf}(t)) \exp i(\mu\gamma_{rf}(t) - \mu\frac{\pi\nu}{N} + m\omega_r t) \tag{3.11}$$

where $\tau$ is the the time of a single R-element, $\omega_r$ is the MAS rate, $\beta_{rf}(t)$ and $\gamma_{rf}(t)$ are rf Euler angles and $\beta_{RL}$ is the spatial Euler angle of the rotor which is at the magic angle (54.74°) in this work. The $d^l_{m'm}(\beta)$ is the notation for the reduced Wigner matrix element which depend on the

space rank of the system. This equation is not trivial to calculate for each sequence therefore in this work SIMPSON simulations of the sequences were used to calculate the scale factors.

### 3.4.6 Rf Inhomogeneities

The R-sequence is made up of of a train of 180° pulses making them extremely sensitive to rf imperfections. There are several different ways that rf imperfections can arise in NMR, such as, due to $B_1$ inhomogeneity in the probe, large spin-spin interactions and small missets of the rf field during optimisations.

$B_1$ inhomogeneity in the probe arises from the fact the coil that surrounds the sample that applies the rf pulse is not infinitely long relative to the rotor. This means that the rf pulses are not homogeneous across the whole of the sample. Hence a 90° pulse is different at the centre of the sample to a 90° pulse at the ends of the rotor. Usually the ends of the rotors experience a rf pulse that is weaker than the nominal pulse in the centre of the sample. This causes issues as it creates a distribution of rotation of the nuclear magnetisation across the sample. If the inhomogeneity of the probe is significantly large, additional background signal from the probe can also be detected which can cause baseline issues in the NMR spectrum. The $B_1$ inhomogeneity of a probe can be measured using the Odedra and Wimperis expeiment[8] which is a nutation experiment carried out with a z-gradient across the sample. This experiment shows how the homogeneity varies across the sample and the nominal frequency that is felt across the sample. In modern probes the $B_1$ inhomogeneity tends to be small. However, in an R-sequence with many pulses the errors from $B_1$ inhomogeneity build up creating significant issues in the experiment such as a large DC offset in the experiment.

Another rf imperfection can be created if there are some spin-spin interactions that are of the order of the rf field amplitude used. This means that the spin-spin interactions can interact and affect the rf field felt by the sample. The main interaction that can be big enough to cause this issue is usually the quadrupolar coupling, however, dipolar coupling and chemical shift anisotropy (CSA) can occasionally be large enough to cause this issue. The samples studied in this work however do not experience this effect.[50]

Finally the third main imperfection is an offset of the rf pulse which can occur is due to all the different sites have slight different Larmor frequencies meaning the nutation frequencies for each site can be slightly different. Therefore, when setting the experiment up on samples a compromise is usually made to get all sites as close to a the desired pulse length as possible.[50]

These offsets and inhomogeneities can severely hamper R-sequence spectrum and a method to make these sequences more robust to these effects has been widely researched. The main methods that have been developed is the use of composite pulses and/ or super cycles.[50,51]

### 3.4.7 Composite Pulses

In order to make R-sequences more robust to the rf inhomogeneities the simple 180° rf pulses can be replaced with a composite pulses. A composite pulses is a sequence of pulses with different phases that have the same overall rotation as a simple rf pulse. These composite pulses have certain compensating mechanisms which make them more robust to certain rf imperfections compared to simple rf pulses.[50] composite pulses can be split into three main types, broadband, narrowband and passband. Broadband composite pulses are the main focus in this work as they are created to be more robust to $B_1$ inhomogeneity than their simple rf pulse counterparts.[52]

Broadband composite pulses can replace simple rf pulses without any additional changes to the experiment making it very easy to implement. Care has to be taken when choosing the composite pulse as the rf field strength required for high MAS rates can exceed the limits of the probe. The change to composite pulses in the R-sequence can also affect the scaling factor, $\kappa$, of the experiment as the total rotation of the R-element is a component of the scaling factor equation.[50]

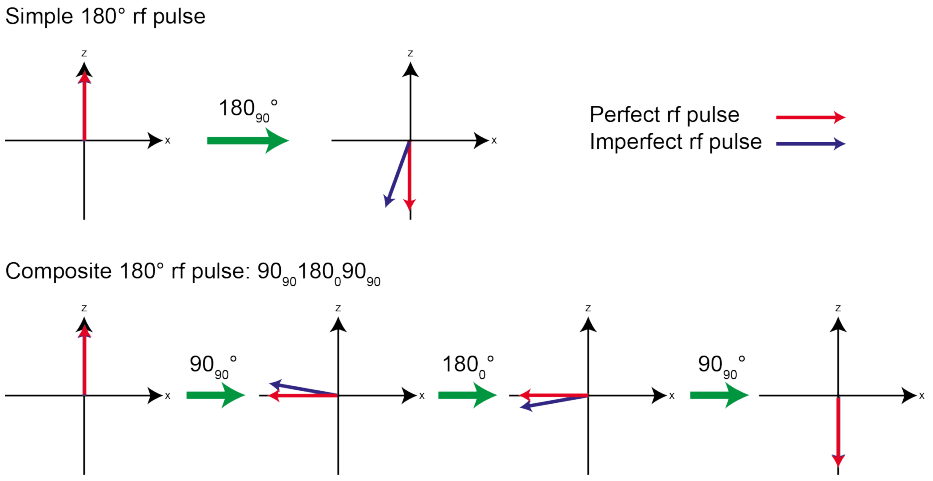


Figure 3.11: A diagram showing how the $B_1$ inhomogeneity affect the rf pulse and how the composite pulse $90°_{90}180°_{0}90°_{90}$ compensates for this inhomogeneity in-comparison to a simple 180° pulse.

Figure 3.11 illustrates how a composite pulse such as $90°_{90}180°_{0}90°_{90}$ is able to compensate for the rf inhomogeneity. However, the diagram does not illustrate that in the case of weaker rf pulses the composite pulse results in a small amount of phase distortion. This can occur in a lot of composite pulses and care is needed to ensure that this does not affect the experiment.[50]

## 3.5 Dynamic Nuclear Polarisation NMR

In 1953 Overhauser introduced a method that used the large polarisation of electrons to enhance the NMR signal of metal nuclei.[53] This was the

basis of Dynamic Nuclear Polarisation (DNP) NMR which, was proven experimentally soon after by Slichter *et al.*, in lithium-7 NMR.[54] Over time NMR moved towards using higher magnetic field strengths due to the increased sensitivity and resolution they provide. Until recently, DNP was limited to low fields ($< 3.5$ T) due to the lack of high powered microwave sources to excite the electronic spins.[55] However, the production of high powered gyrotrons alongside the construction of cryogenic MAS probes has meant DNP is no longer restricted to low fields, resulting in renewed interest into the new applications high field DNP NMR could offer.[55]

DNP takes advantage of the higher polarisability of electrons over nuclei. The enhanced NMR signal is obtained by transferring the large polarisation of the electrons to nearby nuclei. This is achieved by irradiating the electron-nuclear transitions using microwaves. There are several different mechanisms for the polarisation to be transferred in DNP; the Overhauser effect (OE), the Solid effect (SE), the Cross effect (CE) and finally thermal mixing (TM). The dominant method depends on the radical involved, the field strength and the temperature.[55,56]

The first mechanism that was theorised and observed was the Overhauser effect (OE).[53,54] The free electrons within the sample are irradiated by microwaves saturating the electron-nuclear transitions. The time dependent hyperfine coupling enables electron-nuclear cross relaxation leading to nuclear spin hyperpolarisation.[56]

The solid effect (SE) is a mechanism that is dominant in insulating solids where dipolar coupling between the electron and nucleus causes a partial mixing of the nuclear spin state. This enables normally forbidden DQ and ZQ electron-nuclear transitions to be excited. These transitions occur at the sum or difference of the electron $\omega_L$ and nuclear $\omega_{0I_n}$ Larmor

frequencies and so can be excited using microwaves with frequency $\omega_{\mu w}$:[56]

$$\omega_{\mu w} = \omega_L \pm \omega_{0I_n} \tag{3.12}$$

The excitation of these forbidden transitions results in nuclear hyperpolarisation. The enhancement of the NMR signal from the DQ and ZQ transitions have the opposite signs. If the EPR line has a breadth larger than the $\omega_{0I_n}$ then the enhancements of the DQ and ZQ transitions cancel each other out. Therefore, this method requires radicals with small g-anisotropy i.e. a narrow EPR line to avoid this cancellation effect.[56]

The cross effect and thermal mixing are the two main DNP enhancement mechanisms used in the current commercial DNP setup.

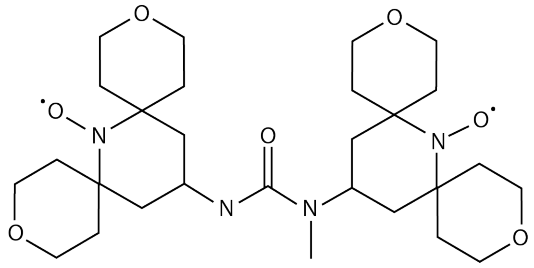


Figure 3.12: The structure of the DNP radical AMUpol.(registered trademark?)

The cross effect uses a pair of electrons usually in the form of a bi-radical such as TOTAPOL or AMUPOL shown in Figure 3.12 and the nuclear spin to transfer the electron polarisation. This is described using a 3 spin model and is only applicable when the following condition is met:

$$\Delta > \omega_{0I} > \delta \tag{3.13}$$

where $\Delta$ is the breadth of the EPR spectrum, $\omega_{0I}$ is the Larmor frequency of the nucleus and $\delta$ is the line width of the EPR spectrum. If this condition is met and $\omega_{0I} = \omega_{0S_1} - \omega_{0S_2}$ is satisfied then a three spin polarisation

transfer is possible where $\omega_{0S_1}$ and $\omega_{0S_2}$ are the 2 electrons in the biradical which have Larmor frequencies that are separated by the Larmor frequency of the nucleus.

To understand the three spin transition more clearly it is easier to visualise the energy level diagram and their populations before and after the electron transition is irradiated with microwaves. Figure 3.13 shows the energy levels of the two electrons that are separated by an energy corresponding to nuclear spin energy $\omega_{0I}$. The energy level diagrams show that before irradiation the populations for each energy level, illustrated by black circles, can be defined according to the Boltzmann distribution. The two levels labelled $\beta\alpha\alpha$ and $\alpha\beta\beta$ have the same energy. It is these matching energy levels that are key to the enhancement of the NMR signal. When electron 1 is irradiated then the populations of the levels change as shown in Figure 3.13b. This causes a population difference between the two matching energy levels. This gives a higher population in the NMR spin of the $\beta$ state which leads to a positive enhancement of the NMR signal. Alternatively if the second electron was irradiated then the populations become as shown in Figure 3.13c where the populations of the nuclear spin $\beta$ state is diminished leading to a negative enhancement of NMR signal. The transition is one where both electron spins are flipped as well as the nuclear spin and is called the flip-flip-flop transition.[55,56]

The theoretical maximum enhancement is given by the ratio of the gyromagnetic ratio of the electron and nuclear spins. This gives a theoretical maximum of $\epsilon = 658$. In order to obtain a reasonable enhancement experimentally a number of parameters have to be optimised some of which require compromise. The cross effect increases with increasing field strength, whilst higher fields require higher frequency microwave sources. Another consideration is which biradical is used and what concentration of radical

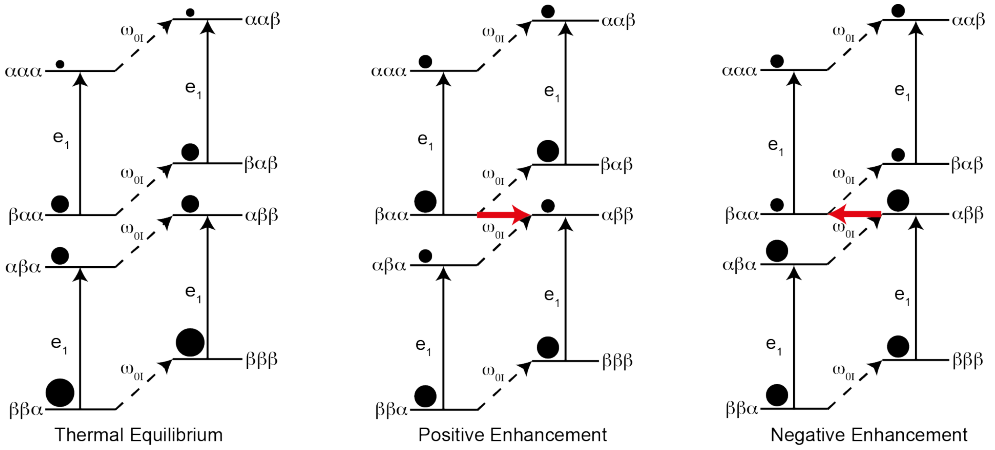


Figure 3.13: A diagram of the cross effect energy levels with the populations illustrated as circles. a) Show the initial populations of energy levels before irradiation. b) Shows electron 2 being saturated by microwave irradiation and the respective population changes. The red arrow shows the direction of equilibration of the 2 matching energy levels leading to a positive NMR signal enhancement. c) Shows the population changes when electron 1 is saturated. This leads to a movement in the populations of the matching energy levels in the opposite direction which results in a negative enhancement of the NMR signal enhancement.

produces the best enhancements. The concentration can have a big effect, as too little and there will not be enough radical to effectively polarise the sample and too much can cause a quenching effect on the sample signal rendering it unobservable. It is also important to choose a solvent that the radical can dissolve into but the sample being observed does not. This is because the radical should surround the sample and not encase it. If the sample dissolves into the solvent the conformation and structure of the solid sample could also changed significantly.[55,56]

# Chapter 4

# Using Solid-state NMR to Study Wheat Straw Composition and Changes During Fermentation

The study of plant architecture presents a difficult challenge due to the complex mix of amorphous and crystalline structures. Standard methods, such as acid hydrolysis completely destroy the plant's structure meaning any morphological information is lost. In acid hydrolysis components are separated and analysed to obtain the overall composition of the plant.[57,58] This method is not ideal as the efficiency of the extraction varies considerably and can result in significant errors in the analysis. Solid-state NMR can provide an alternative means of studying plant samples, allowing *in situ* analysis which can not only provide the composition but also an insight into the internal plant architecture.

In this chapter, the use of solid-state NMR to study the composition and molecular architecture of wheat straw is investigated. In this work

a quantitative NMR method is used to study the composition of wheat straw whilst 2D NMR experiments are used to shed light on the molecular structure of the cellulose and hemicellulose within the straw. These experiments are then used to study the changes in the wheat straw composition and structure during a 5 day fermentation process with a fungus, *Aspergillus niger*.

The project was proposed by Dr. Jolanda von Munster as the use of solid-state NMR, particularly 2D NMR experiments, were beginning to be used to study the molecular architecture of plants by the Dupree and M. Hong groups. Previous work into the use of Aspergillus niger as a pretreatment method was investigated by a PhD student, Dr. Nattha Pensupa in 2015 using standard saccharification techniques to study the changes in the wheat straw during this process.[3] This project aimed to see if any additional information such as, changes in molecular structure of the wheat straw during this fermentation process could be revealed using solid-state NMR. At around the same time the wheat straw project had been proposed, two different groups at Sutton bonnington, Nottingham approached our group with 2 different Humic acid samples. They wanted to know the composition of these two humic acids. The compositional analysis of these humic acid samples was used as a means to validate the Schmidt-Rohr quantitative method which was then used for the quantitative analysis of the wheat straw.

## 4.1 The Composition of Lignocellulose Biomass

Lignocellulose biomass is generally comprised of two main carbohydrate polymers, cellulose and hemicellulose (xylan and xyloglucan) alongside a non-carbohydrate polymer lignin. Different lignocellulose materials contain

different proportions of these three main components which can also vary depending on the plant organ, age and the stage of growth of the plant studied. This means for each plant there tends to be a range of compositions identified in literature.[3]

*Cellulose*

Cellulose is the main component found in both the primary and secondary cell walls of most plants and is thought to provide much of the load bearing structure. It is a polymer comprised of D-glucose monomers which are connected via a $\beta$-(1-4) linkage. The cellulose chains are known to have a 2-fold ($2_1$) helical conformation, and the helices stack together with intra- and intermolecular hydrogen bonds to form large structures known as microfibrils.[4]

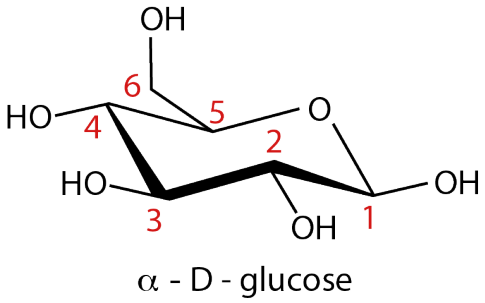


Figure 4.1: The cellulose monomer, $\alpha$-D-glucose with the carbon numbering used in the text.

Initial solid-state NMR studies of plants found that there are at least two different cellulose environments present within the microfibril structure.[59] 1D NMR experiments showed that there are multiple NMR shifts for the C4 (84.0 and 89.0 ppm) and C6 (62.0 and 66.0 ppm) carbons of cellulose and the ratio of these signals indicate that they do not arise from different sites within a unit cell.[59] The assignment of these two distinct cellulose environments has changed a lot over the years and still remains somewhat ambiguous. They were initially assigned to well known crystalline forms of cellulose called I$\alpha$ and I$\beta$ but have since been more gen-

erally defined as either amorphous and crystalline or surface and interior sites.[4]

The I$\alpha$ and I$\beta$ forms of cellulose have been produced synthetically and characterised using solid-state NMR and diffraction techniques. It is known that I$\alpha$ crystals have a triclinic unit cell using one cellulose chain whereas I$\beta$ has a monoclinic unit cell that contains two parallel cellulose chains.[60] Initially, the NMR shifts of these crystalline forms were compared to extracted cellulose from plants and fungi. In these comparisons the cellulose structures did not appear too dissimilar and the NMR shifts of the extracted cellulose NMR could be explained by it being a mixture of the two crystalline forms.[60,61] It was thought that these two crystalline forms made up different parts of the cellulose microfibril or could even form separate microfibrils within the plant. Many studies have tried to elucidate how these two crystalline forms were oriented in the plant.[60,62] An x-ray diffraction study suggested that they are stacked together in the same microfibril. They propose that the stacking of the two crystalline forms of cellulose could lead to faults in the structure that create the different cellulose environments. The addition of these faults into the model produced a simulated diffraction pattern that fitted the finer details of the experimental diffraction pattern of powdered supporting this stacking theory.[62] In higher plants such as wheat straw it was thought that its cellulose is predominantly in the form of I$\beta$ with around 25% in the I$\alpha$ form.[59] The variable difference in the intensities of the different C4 and C6 environments between plants and conflicting studies meant that it was not possible to consistently assign the two cellulose environments using these two crystalline forms.[4]

Hong *et al.* has shown in computational simulations of a cellulose microfibril that changes in the glycosidic C6 conformation and carbon bond angles can affect the NMR shifts of the C4, C5 and C6 carbons.[1] Changes

in the conformation of the C5-C6 bond can shift the C6 and C4 peaks by around 4 ppm, which is roughly the difference seen in the experimental cellulose environments. How these different conformations of cellulose occur within the cellulose microfibril is still debated. Different groups assign the 2 environments as either surface and interior or crystalline and amorphous sites as both could result in differences in the glycosidic bond conformation.[1,4] In the NMR spectrum of plants it was found that neither peak of the two cellulose environments is broader than the other. This infers that the surface and interior assignments are more probable which is the designation that the Dupree *et al.* tend to adopt in their recent work on the *Arabidopsis thaliana* plant.[4,63]

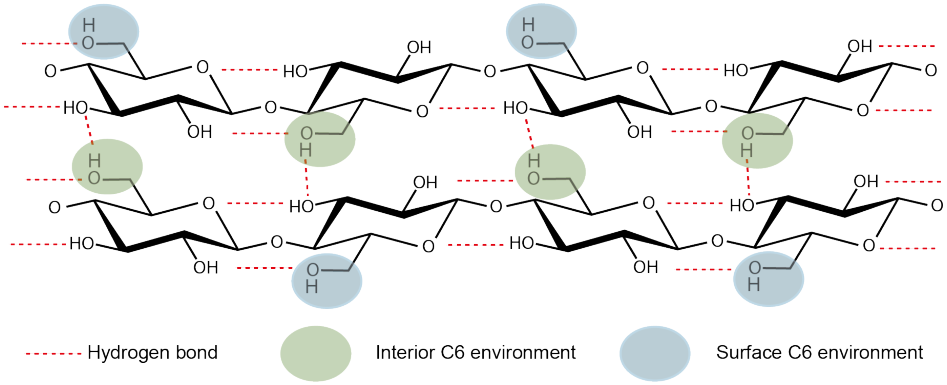


Figure 4.2: A diagram illustrating how the inter- and intramolecular hydrogen bonds hold the cellulose chains together. In between the two chains the C6 hydroxyl group is fixed in position by two hydrogen bonds whereas when on the surface of the microfibril this is only held by one bond meaning it is able to to rotate more freely about this C5-C6 bond. This rotation of the C6 hydroxyl group has been shown in simulations to cause significant differences in the NMR shifts for carbons 4, 5 and 6.[1]

Due to the numerous theories relating to the cellulose structure throughout the literature there is also a debate of which of these two NMR signals should be assigned to surface and interior cellulose environments. In 1995, Newman *et al.* measured the $^{13}C$ spin-lattice relaxation times of the two distinct cellulose sites. It was concluded that the shorter relaxation times

were those sites relating to the surface cellulose (C4 - 84.0 ppm) and the longer relaxation times with the interior cellulose (C4 - 89.0 ppm).[64] More recently, in 2015, a 3D DARR (Dipolar Assisted Rotary Resonance) experiment was acquired on the *Arabidopsis thaliana* plant where it was found there was a stronger dipolar correlation between the C4 site at 89.0 ppm and the protein resonance than that at 84.0 ppm. This lead to a conclusion that the cellulose C4 site at 89.0 ppm is likely to be surface cellulose to enable it to have a stronger correlation with proteins on the outside of the microfibril.[4] This designation of surface and interior sites was altered yet again in a subsequent 2016 paper by Simmons *et al.* where the C4 at 89.0 ppm is referred to as the interior site.[63]

The number of cellulose chains per microfibril is poorly defined and a number of different theories put it between 10 - 36 chains.[65–68] Recently, this number of chains within a microfibril has settled on around 24 chains per microfibril based on some wide angle X-ray diffraction data of spruce wood.[65] One other explanation for the discrepancy in microfibril sizes could be due to different plant systems having their own individual microfibril structure. The habit (i.e. the shape) that these chains form in the microfibril is also poorly understood with numerous computational studies developed in order to understand which habit is most energetically favourable. Kubicki *et al.*[2] recently studied three potential habits of an 18 chain microfibril; 6 $\times$ 3, 234432 and 34443. The 34443 habit shown in Figure 4.3 was found to be most energetically favourable according to computational results by an average of 120 kJ/mol depending upon the DFT method.[2] The cellulose microfibrils are thought to selectively bundle together in the plant to form large cellulose fibres that provide the load bearing structure of the plant.[4]

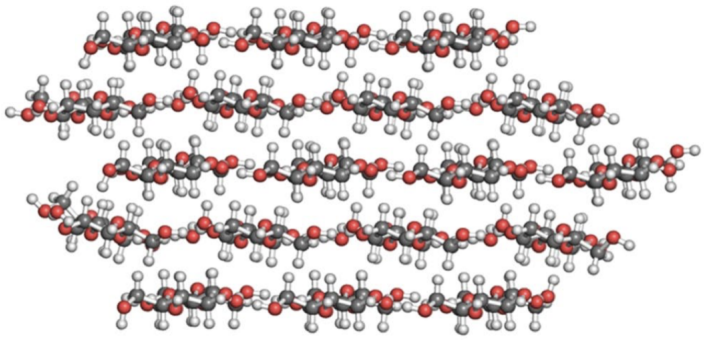

Figure 4.3: An example of a microfibril cross section with the habit 34443 studied by Kubicki *et al.*.[2]

*Hemicellulose*

The second most prevalent component is hemicellulose, which is made up of a wide range of polysaccharides. Within the primary cell wall the main hemicelluose component is xyloglucan which has a glucose backbone connected via $\beta-(1-4)$ linkage with substitutions of xylose, galactose and acetate along the chain. In the secondary cell wall the main hemicellulose component is xylan which has a backbone made of xylose monomers which are connected via a $\beta-(1-4)$ linkage that are decorated with acetate and galacturonic acid. Other polysaccharides such as arabinan and mannan are also present in smaller quantities as part of hemicellulose.[3,4]

In materials like wheat straw the secondary cell wall makes up the majority of the plant so xylan is the predominant hemicellulose component. Xylan in solution and when it is made synthetically is known to have a 3-fold ($3_1$) helical conformation. This has been found by solution-state NMR to have X4 and X5 shifts of 77.2 and 63.7 ppm respectively. However, the solid-state NMR of plants reveals that the X4 and X5 peaks shift downfield to 82.2 and 64.3 ppm respectively. This indicates that the xylan has a different conformation in the plant and when interacting with the cellulose.[63,69]

A recent study using solid-state NMR by Dupree *et al.*[63] has shed light

on how the xylan chains could lie across the cellulose microfibrils. This was achieved by studying the differences in xylan correlations found in the refocused INADEQUATE spectra of a wild type and mutant of the *Aradopsis Thaliania* plant. The mutant was a strain that had much of the cellulose production inhibited leading to a dwarf plant due to the lack of cellulose. The refocused INADEQUATE spectrum of the mutant plant had a xylan X4 shift at 77.4 ppm which is close to the known solution-state shift of xylan in its 3-fold helical conformation. The X4 shift in the wild type plant is at 82.2 ppm, which is a shift of nearly 5 ppm indicating a clear change in the xylan conformation compared to the mutant. It was thought that the cellulose reduction in the mutant could lead to most of the xylan unable to interact with the cellulose microfibril leading to the xylan having a solution-like 3-fold helical conformation.[63]

Dupree *et al.* went further and hypothesised how xylan could interact with the cellulose microfibril. Previous work has shown that the substitutions of xylan are found to be evenly spaced residues leading to a theory that if the xylan was in a 2-fold conformation the acetate could be along one side making binding to cellulose possible.[69] Therefore the X4 and X5 peaks at 82.2 ppm and 64.3 ppm were assigned to xylan bound to cellulose in a 2-fold helical conformation.[63]

*Lignin*

Lignin is a large non-carbohydrate polymer that is generally made up of 3 monomers known as p-hydroxyphenyl (H), syringyl (S) and guaiacyl (G) shown in Figure 4.4.

The monomers are connected via a number of different ester and carbon bonds. The ratio of the H, S and G monomers in a biomass depends on the type and age of the plant. Lignin is formed in mature plants via a process known as lignification and provides the plant with mechanical

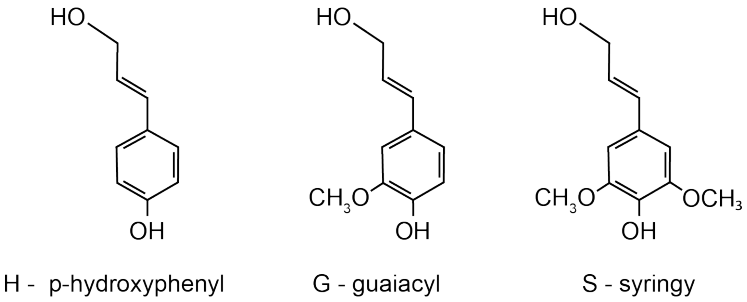


Figure 4.4: The 3 monomers of lignin H, G and S.

strength and recalcitrance.[70] Lignin is extremely difficult to study due to its large complex structure, insolubility and lack of uniformity. The lignin is theorised to encase the cellulose and hemicellulose which makes extraction of fermentable sugars for biofuels more difficult.[71] There is some debate on how lignin encases the cellulose and hemicellulose especially regarding how the lignin interacts with them. There have been some studies that suggest that the galacturonic acid on the xylan can form a linkage with some of the lignin making it more difficult to remove the cellulose and hemicellulose.[72]

Within plants there are also small amounts of proteins and pectins which are generally found in the primary cell wall. A schematic of the key polymers of the plant is illustrated in Figure 4.5 alongside a diagram of how the 3 main components are thought to be arranged in the plant in Figure 4.6 .

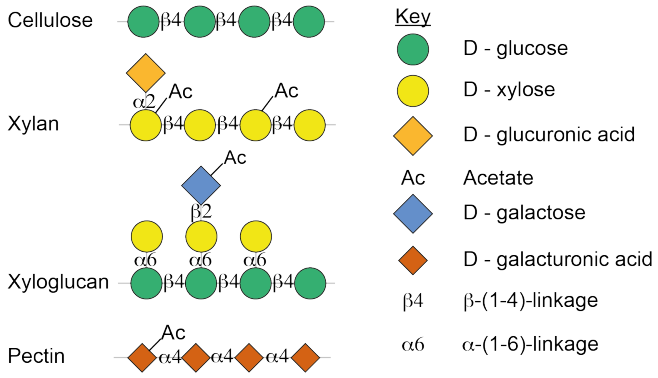


Figure 4.5: A schematic of the main polymers that make up plants.

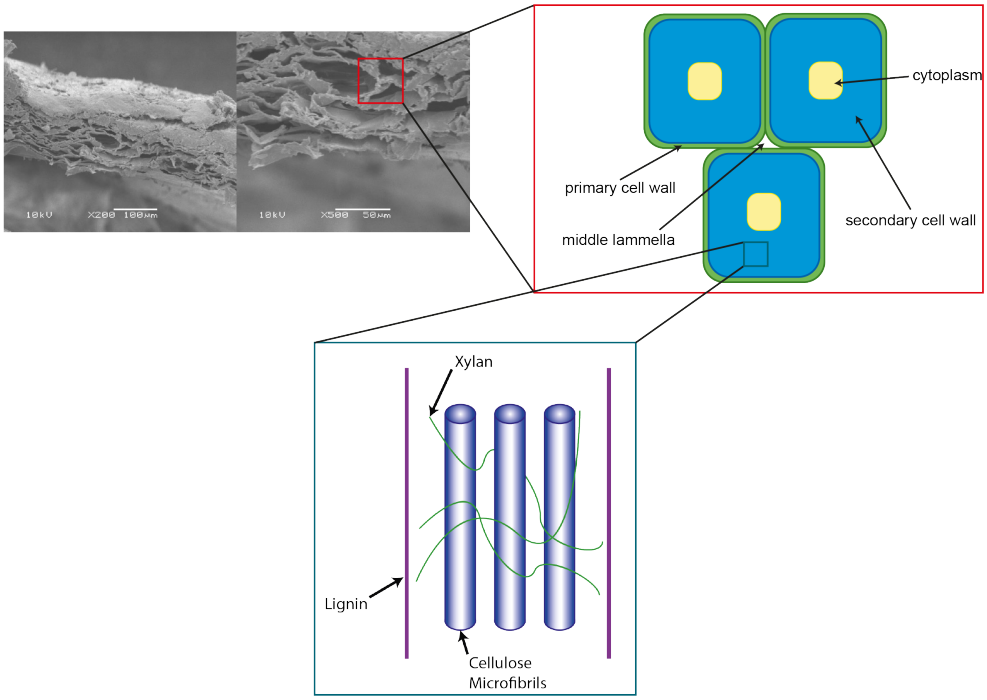


Figure 4.6: Top Left: Scanning electron microscopic images of untreated wheat straw with electron potential of 10 kV and with magnification of 200 and 500 times.[3] Top right: A simplified diagram of the cells within the wheat straw. Bottom: A diagram of the proposed structure of the secondary cell wall.[4]

### 4.1.1 Composition of Wheat straw

Measuring the compositions of plants is a well researched topic with a range of analysis methods developed over the past 100 years. In the literature the main method used to measure the cellulose and hemicellulose content is complete acid hydrolysis developed by Saeman *et al.*.[73] This method adds a strong acid to the plant samples and is left to incubate for an hour at 37°C after which the acid is then diluted and the resulting solution is incubated at 100°C for a further 2 hours. The concentrations of the monosaccharide extracted are found using High Performance Liquid Chromatography

(HPLC) using a standard calibration curve to identify the concentrations and retention times.[73]

The standard method for lignin analysis uses the acetyl bromide method developed by Fukushima *et al.* in 2001.[58] This method uses acetyl bromide in acetic acid which is added to the sample and left to incubate at 50°C for 2 hours with continuous mixing. The lignin acetyl mixture is centrifuged and the optical density (OD) of the sample is measured at 280 nm, then concentration calculated using a standard calibration curve. Lignin is more difficult to analyse as its very nature makes it hard to breakdown and extract from the plant efficiently.[58,71]

The measured composition of a plant can depend on its strain, age, the conditions it is grown in and the analytical methods used. For straw, there are at least three main parts, leaves, stem, and nodes. The compositions in each are also slightly different and so the overall composition will also depend on which part of the wheat straw is analysed.[3] There are numerous composition studies of wheat straw in literature with a significant range of values for each component. The cellulose content of wheat straw in literature ranges from 31% to 48.5%[3,14] while the hemicellulose content is calculated to be between 16.7% and 49%.[16,74] Finally the lignin is expected to make up around 8 - 25%.[13,74] The other constituents such as protein and pectin are very minor components of the wheat straw. It is clear to see that there is a broad range for the composition of wheat straw and so it is key to know the composition a particular sample rather than assuming literature values.

### 4.1.2 Humic Acids

The Schmidt-Rohr quantitative method was originally used to study a wide range of humic acids and so in order to validate this quantitative method

the compositions of two different humic acids were studied. Therefore, a brief description of what humic acids are and the main models for their composition are given in the following section.

Natural humic substances are the major organic component of humus that make up soil, peat, and coal. Humic substances are a mixture of predomiently Humic acids and fulvic acids. Humic acids are a complex polymer mixture that is generally found in soils, sediment and waters. The most concentrated source of humic acids is found in a substance called Leonardite which is also referred to as brown coal. The difference between coal and Leonardite is its increased oxidation degree.[75] Humic acids play a pivotal role in many agricultural and environmental reactions such as sequestering metal ions which then in turn can be absorbed by plants. They have since been found to aid in increasing the water capacity in soils and are used in water treatment methods as well as having anti-bacterial and anti-fungal properties.[75,76]

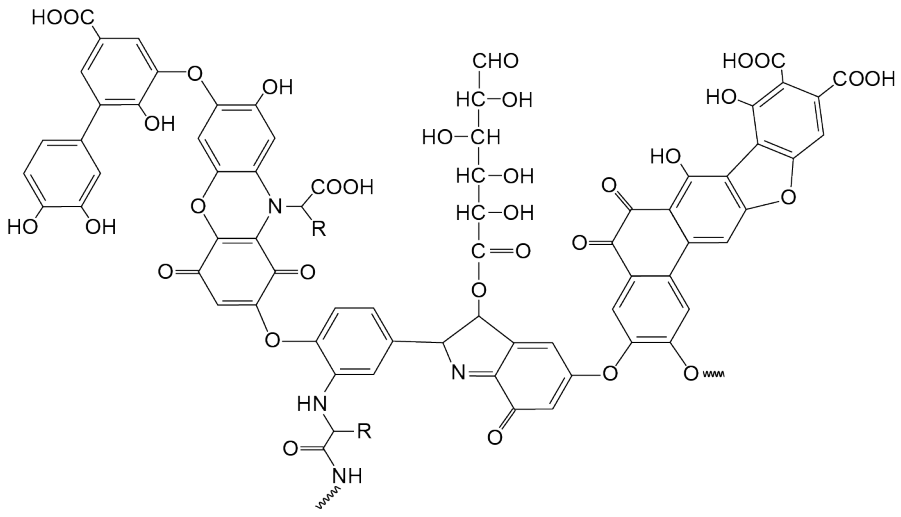


Figure 4.7: An example of the complex polymer structure of humic acid.

The composition of humic acids from different sites varies drastically, since it depends on both the nature of the organic matter which has been broken down and the conditions under which this process occurred. The

compositions and structures of the humic acids are thought to provide an insight into their properties and reactivity.[76] The basic soluble component of humic acids have been studied in more depth in literature using a range of techniques such as solution-state NMR, mass spectrometry and UV/vis spectrometry. It has been found that the ratio of the $sp^2$ and $sp^3$ carbons in humic acids can correlate with some of their key properties. An example of the complex structure of humic acids is shown in Figure 4.7.[10,77]

*Humic Acid Models*

Due to the complex structure of humic acids many different models have been proposed with differing amounts of $sp^2$ and $sp^3$ carbon environments as an attempt to understand their properties and reactivity.[76] A few of the main models that are most relevant to the humic acid samples analysed are introduced in this section.

The first attempt to model humic acid was made by Fuchs' in 1931. This model contains 3-4 carboxyl groups, 3-4 hydroxyl groups, a methoxy and a carbonyl group. However, it does not have an aliphatic component and was quickly found to be a too simple a model to encompass the variation seen in humic acids.[78]

In 1968 Dragunov's model was proposed which is polymeric in nature and is predominantly made-up of aromatic rings which are bridged with O, $CH_2$, NH, N and S. When broken down into the major carbon environments in this model works out to be almost evenly split between quinone, aromatic and carbohydrate groups with a small amount of aliphatic groups. This model is better than Fuchs as it at least contains an aliphatic component but it does not account for any of the carboxyl groups present in the majority of humic acids.[10,75]

Steelink's 1985 model is a further improvement and contains aromatic rings, phenol, quinones, aliphatics and carboxylic groups.[10] The aliphatic

groups have the majority of OH groups attached and the caboxylic groups are solely placed on the aliphatic groups. This model still has a lack of $OCH_3$ and anomeric carbons. Jansen modified this model generating building block with 7 chiral centres that have 128 stereoisomers with ketones or aldehydes instead of quinones. Unfortunately, Jansen's model still does not include anomeric carbons. The Steelink and Jansen's models while flawed, are thought of as a reasonable basis for the humic acid structure.[10]

More recent models, such as the Leenheer model developed in 1989, is made up of three different blocks which are each form from two aromatic rings and a tetrahydrofuran ring with methyl terminated side groups and ketones.[75,79] This structure leads to conformational constraints to intermolecular H-bonding between phenol and carbonyl functional groups.[10] The model's composition has approximately equal amounts of phenol and aromatic carbonyl groups that can interact via hydrogen bonding. The phenol groups adjacent to the aromatic carbonyl groups are thought to lead to the phenol groups behaving as weaker acids than expected.[10,79]

Finally, Schulten and Schnitzer's model developed in 1993 is a complex model produced reflecting the results of mass spectrometry and CPMAS NMR experimental data. The resulting model is predominantly made up of aliphatic and aromatic environments with some carboxylic groups. The CPMAS spectrum have since been shown to be flawed due to the none quantitative nature of the CP experiment. This discrepancy with the previous NMR data is reflected in its particularly low $sp^2$/ $sp^3$ ratio compared to all previous models.

### 4.1.3 Introduction to Biofuel Production

In recent years, with the concern of diminishing fossil fuel resources the need for a sustainable replacement has been a major priority. Currently around

| Model | $sp^2/sp^3$ ratio |
|---|---|
| Steelink | 2 |
| Jansen | 1.75 |
| Fuchs | 1.75 |
| Dragnov | 1.4 |
| Leeheers | 1.1 |
| Schulten and Schitzen | 1.0 |

Table 4.1: A table of the $sp^2/sp^3$ ratio for the main humic acid models.

80% of the world's energy is provided by fossil fuels which cannot sustain the level of energy consumption for much longer than a further 50 years.[80] This has lead to an enormous amount of research into the generation of biofuels for transportation which currently makes up 32% of Europe's energy consumption.[80] At the moment, biofuels are primarily produced using food crops such as corn, wheat and sugar beet.

In the mid 1970s oil prices were at an all time high and Brazil's import costs had become unsustainable. In order to reduce the dependence on oil imports the drive to produce bioethanol from large amounts of cheap sugarcane bagasse was developed. This bioethanol production has grown and now provides Brazil with around 30% of all its transportation fuel.[80,81]

However, the use of food crops to produce fuel can be seen as unethical as leads to fuel being in direct competition with the food supply and has resulted in negative public perceptions of biofuels. Ideally an economical alternative that produces biofuel from waste materials such as grasses, straws and wood would solve this issue. There is a huge source of these waste materials in the agricultural industry that is currently not utilised. In Europe alone, over 122 million tonnes of crop residues are produced every year that could be used for sustainable biofuel production.[82]

First generation biofuels are produced using a three stage process: hydrolysis, fermentation and distillation. The hydrolysis stage extracts the cellulose and hemicellulose components in the form of fermentable sug-

ars such as glucose and xylose. The two main hydrolysis methods used are chemical and enzymatic.[3] Chemical hydrolysis is achieved by subjecting the biomass to strong acid treatments that breakdown the cellulose and hemicellulose structures. The extent to which this method extracts the fermentable sugars depends on the sample, the strength of acid and the temperatures used.[3,83] The alternative method is enzymatic hydrolysis which uses selective enzymes to breakdown the cellulose and hemicellulose. This second method is generally preferred as it requires mild conditions and, unlike acid hydrolysis, does not produce any byproducts such as fufural that are known to inhibit the fermentation process.[84] Fermentation of these extracted sugars then produces the bioethanol which can be collected via a distillation process.[3]

Unfortunately, generating biofuels from waste materials such as straw using these methods developed on food crops produces a very low yield that is not economically viable. For a new fuel to overtake first generation biofuels and fossil fuels it needs to be profitable and more sustainable.[3] It was determined that low biofuel production yield is due to the presence of lignin.[4,85,86] To overcome this issue numerous pretreatment methods have been suggested to improve the efficiency of the production of biofuels from these second generation sources.

### 4.1.4 Pretreatment Methods

The aim of the pretreatment stage is to break up the recalcitrant structure of the plant in order to make it easier to extract fermentable sugars during the subsequent hydrolysis stage. A cost effective pretreatment method should be low energy and should not produce any byproducts that inhibit the hydrolysis and fermentation processes.[86]

In the literature, a wide range of different pretreatment methods have

been designed and trialed for use in the production of second generation biofuels. They can generally be split into 3 main categories; physical, chemical and biological.

### Physical Pretreatment Methods

Physical pretreatment methods such as milling, grinding and freezing, use physical force to break down the plant structure and increase its surface area. These methods have been tested on a range of different lignocellulose materials with varying success.

Depending on the biomass, milling can give an increase in bioethanol production of between 40 and 90%.[87–89] The use of milling is seen to be quite effective and is generally used alongside another pretreatment methods. The freezing method exploits the expansion of water when forming ice which, in the plant, can cause structural damage and an increase in the plant surface area. In 2011 this method was trialled on rice straw leading to a 4 fold increase in the production of glucose sugars.[90] Other methods include irradiating the sample using electromagnetic and mechanical waves such as micro- and ultrasonic waves.[91,92] However, these are less well studied due the the specialist equipment needed and the energy costs. Physical pretreatment methods tend to be very simple and generally involve a reduction in the samples particle size to increase the surface area. However, particularly on the industrial scale, these methods require large amounts of energy. In order to make these methods more viable they are usually applied alongside other methods such as chemical pretreatments.[3]

### Chemical Pretreatment Methods

Chemical pretreatments use chemical reagents to disrupt the lignocellulose structure. The chemicals used are selected as they are known to affect

certain bonds within the plants structure. These methods include acid or alkali treatments as well as oxidative delignification and hydrothermal treatments.[3]

Acid pretreatments use either sulphuric acid or hydrochloric acid as it was found in 1883 that these can depolymerise the cellulose chains.[93] This method is used commercially due to its ease of applicability, particularly on the industrial scale, and its efficiency to depolymerise the cellulose. The acid treatments are usually split in to two types; strong and dilute acid methods.

The strong acid method uses acids such as concentrated sulphuric acid at moderate temperatures of 50-60°C for several hours.[3] During this time much of the cellulose is broken down into its oligomers. This means that the subsequent enzymatic hydrolysis stage is not required and so just the fermentation process is needed to produce the biofuel. Whilst this is an attractive process it is not with out its drawbacks. The strong acid creates a containment issue due to its corrosive nature and also produces byproducts, such as furfural which inhibit the fermentation process. In order to make this method more cost effective a means of recycling the acid would need to be developed.[94,95]

Dilute acid treatments use very dilute acids alongside high temperatures between 160 - 220 °C for short periods of time.[3] This method seems to have a similar effect on the plant structure as the strong acid treatments but is less corrosive and produces fewer inhibitory compounds. An increase in both acid concentration and temperature improves the depolymerisation of cellulose but also increases the production of compounds such as furfural. This means a compromise between the concentration inhibitory compounds produced and the amount of depolymerisation of the cellulose and hemicellulose in the plant.[95,96]

The increase in sugar yields found that there was not much difference between using dilute or concentrated acid treatments leading to more studies using the dilute method as they do not require corrosive resistant containers and are less time consuming. It was found that the dilute acid treatment enables up to 96% of the glucose was removed from barley straw along side 57% of xylose by using 1.5% sulphuric acid (w/v) at 121°C for an hour.[97] This success of extracting the majority of fermentable sugars using this method is mirrored in many studies on different lignocellulose materials.[98–100]

Alkali treatments, like those with acid, are split in to dilute and strong alkali methods which are applied using a range of temperatures (50 - 135 °C). The alkali targets bonds in the lignin but can also affect cellulose and hemicellulose structures. Lignin degradation is the primary process that occurs using alkali treatments but when higher temperatures are used then significant degradation of the cellulose and hemicellulose can also be seen.[101–104] Alkali treatments take longer than their acid counterparts but they have been seen to degrade up to 77% of the lignin in switch grass which leads to a 4 fold increase in the sugar yield.[105] When alkali pretreatments are undertaken using mild conditions, no detectable inhibitory compounds are produced.[106] However higher temperatures and concentrations lead to greater degradation of lignin but more inhibitory byproducts such as formic and acetic acids are produced, which have to be removed before hydrolysis and fermentation.[3,107]

Oxidative delignification uses an oxidising agent such as hydrogen peroxide to degrade the lignin. It has been found that using this method in conjunction with alkali pretreatment methods can result in the removal of 80% of lignin and an 90% increase in the sugar yield.[108] Other methods such as using ozone to breakdown the lignin are not economical due to the large

amounts of ozone that are needed to obtain similar results. Others have also tried a number of different organic solvents however their effectiveness depends on the solvent, the plant and the temperatures used.[3]

Hydrothermal treatments use high temperature water (180 - 200°C) to break the plant structure in a similar way to dilute acid hydrolysis. This method uses high temperatures but tends to require less time than acid hydrolysis, and produces no inhibitory compounds like furfural.[109] It shows a significant increase in the sugar yield of over 90%, however, the high energy costs need to be taken into account.[3,109]

### Biological Pretreatment methods

Biological pretreatment uses microorganisms, such as fungi, to disrupt and break down the biomass structure. Fermentation processes have the benefit that they use very mild conditions but it can also take a long time to obtain enough degradation to make a difference in the production of fermentable sugars. Many studies have been made using different strains of fungus which are designed to degrade different components of the plant.[110] These processes have been studied over a wide range of timescales with studies over just few days up to several months.[111,112] Biological methods such as fermentation are industrially scaleable and are relatively safe compared to chemical treatments. The enzymes produced during the fermentation process could also be collected and used in the hydrolysis stages of the biofuel production process. While time consuming this method requires very little energy and does not produce any inhibitory compounds making this a promising pretreatment for lignocellulose material.[3]

A key part of this work focuses on studying the changes in wheat straw during a fermentation process using a fungus known as *Aspergillus niger* as a potential pretreatment method.

### 4.1.5 *Aspergillus niger*

*Apergillus niger* is a common fungus which is also known as black mould. It is referred to as a model fungus and is used in industry to produce citric acid. *Aspergillus niger* is regularly found in soil and plays a key part in the carbon cycle. It is known to produce many degradative enzymes called amyloglucosidase that break down the cellulose and hemicellulose in lignocellulosic biomass. Understanding the fermentation process and the enzymes this fungus produces could provide key understanding in how it degrades lignocellulosic material which in turn is useful for the efficient production of biofuels. In this work, *Aspergillus niger* has been used as a pretreatment method for breaking down wheat straw and so understanding the composition of the fungus is important as the NMR spectrum could be affected by the growth of the fungus. Unlike wheat straw, the fungus cell walls comprise of 50-60% glucan chains which are mainly joined via $\beta$-1-3-linkages however up to 10% can be connected via $\beta$-1-6 and $\beta$-1-4-linkages. The glucan chains which are joined together via $\beta$-1-3-linkages have a different conformation than the cellulose in wheat straw and so have different NMR shifts. The rest is comprised of 20-30% glycoproteins and 10-20% of a carbohydrate called chitin. It was found using solid-state NMR that the glucan chains and chitin are in close proximity and form a rigid scaffold which is hydrophobic and is surrounded by a hydrophilic $\beta$-1-3-linkage glucan matrix. The NMR shifts for the main constituents of the fungus have been assigned by Kang *et al.* and are compared with the cellulose shifts of wheat straw in the table below.[9] The table of shifts shows that whilst there are some differences between the cellulose shifts and that of the $\beta$-1-3-linkages glucan. Due to these differences it is thought to be possible to distinguish the shifts of the fungus from the wheat straw changes particularly at the later stages of the fermentation period.

| Carbohydrate | C1 | C2 | C3 | C4 | C5 | C6 | CO | $CH_3$ |
|---|---|---|---|---|---|---|---|---|
| Cellulose 'Surface' | 105 | 72 | 75 | 84 | 75.5 | 62.5 | - | - |
| Cellulose 'Interior' | 105 | 72 | 75 | 89 | 72.5 | 65 | - | - |
| Glucan $\beta$-1-3-linkage | 102.5 | 73.0 | 84.5 | 69 | 77.0 | 60.5 | - | - |
| Chitin | 103.0 | 55.5 | 73.0 | 83.0 | 76 | 61 | 175 | 22 |

Table 4.2: A table of the shift of cellulose and the main constituents of the fungus taken from literature.[4,9]

### 4.1.6 Aims of this project

In this work two different approaches to quantitative analysis of organic material have been investigated. Their resultant compositional analysis of a wheat straw sample have been compared. The composition and architecture of the cell walls of wheat straw have been studied using 1D and 2D NMR experiments. These same experiments were then used to investigate the changes in the cell wall structure over a 5 day fermentation process with the fungus, *Aspergillus niger*.

### 4.1.7 Experimental details

Samples:

The Schmidt-Rohr quantitative method was initially tested on 2 different Humic Acid samples provided by the University of Nottingham, Sutton Bonnington campus.

All experiments were used to study wheat straw samples which were uniformly $^{13}$C-enriched labelled (>97% $^{13}$C, leaf and stems, Triticum aestivum) bought from IsoLife in the Netherlands that were milled by the supplier using a Culatti micro hammer mill and sieved to 0.5 mm.

Solid-state fermentation conditions: The $^{13}$C labelled wheat straw was used in the solid-state fermentation process. A brief description of the fungus preparation and fermentation process is given here. However, further details can be found in the paper by Daly *et al.*.[113] The Aspergillus min-

imal media (AMM) was prepared with a pH of 6.5 using NaOH. Glucose cultures 1% w/v were inoculated with *Aspergilus niger* N402 spores with a final concentration of $1 \times 10^6$ $ml^{-1}$ and were incubated for 48 hours at 28 °C. The resulting mycelia were collected by filtering through a Miracloth and 1.5 g wet weight was transferred to flasks containing the labelled wheat straw to give a 1% w/v final concentration of AMM to straw. The cultures were incubated with complete cultures harvested at 9, 12, 24, 48, 72 and 120 hours alongside a control sample. The control sample was a sample of wheat straw without the addition of fungus which was then incubated for the same duration as the fermentation period. The cultures were then lyophilised (freeze dried) to halt further fermentation processes before being packed into 4 mm MAS NMR zirconia rotors.

Solid state NMR:

The experiments on the humic acid samples were acquired on a Varian 300 MHz spectrometer with a magnetic field strength of 7.05 T corresponding to Larmor frequencies of 300.130 MHz and 75.468 MHz for $^1$H and $^{13}$C respectively. Adamantane was used as an external reference with the downfield shift set to 38.4 ppm in compliance with the IUPAC convention.[114] The $^{13}$C MAS spectra for the humic acid samples were acquired using a 4 mm triple resonance probe with an MAS rate of 12.5 kHz ± 5 Hz to ensure the spinning sidebands were minimised. A spin echo was used with short delay of 20 $\mu$s to reduce any baseline distortions and the 90° carbon pulse was optimised to 3 $\mu$s. A recycle delay of 25 s was used with SPINAL-64 heteronuclear decoupling during the acquisition time.[6] The first sample was acquired with 13312 scans whilst the second was acquired with 12000 scans leading to total experimental times of 92.4 and 83.3 hours. The CP/$T_1$ experiments were acquired on a 7 mm double resonance probe with a MAS rate of 6 kHz ± 10 Hz with Total Sideband Suppression (TOSS)

pulse sequence[38] before acquisition to ensure the spinning sidebands did not obscure the line shape. Each experiment was acquired using a contact time of 1.5 ms along with $^{1}$H and $^{13}$C 90° pulse lengths of 5 $\mu$s and 4.5 $\mu$s. During acquisition SPINAL-64 heteronuclear decoupling was applied to the $^{1}$H channel.[6] The experiments had a recycle delay of 2 s and were acquired using 2528 scans for both samples. The $T_1$ filter delay was set to 1 $\mu$s for the short z filter delay and 25 s for the long z filter delay giving total experimental times of 1.40 hours and 1.46 days respectively.

All experiments on the wheat straw samples were acquired on a 600 MHz Bruker Avance III solid-state NMR spectrometer operating at a magnetic field of 14.1 T corresponding to Larmor frequencies of 600.130 MHz and 150.903 MHz for $^{1}$H and $^{13}$C respectively. The wheat straw samples were packed in a 4 mm zirconia rotor and an $E^{free}$ 4 mm triple resonance MAS probe with a MAS rate of 12 kHz ± 5 Hz was used. Adamantane was used as an external reference with the downfield shift set to 38.4 ppm in compliance with the IUPAC convention.[114] The 90° pulse lengths used in the experiments were 3 $\mu$s for $^{1}$H and 5.5 $\mu$s for $^{13}$C with SPINAL-64 decoupling.[6]

When using the Schmidt-Rohr method to study the fermentation of wheat straw all 1D $^{13}$C MAS experiments were acquired with 1184 data points, an acquisition time of 9.95 ms and 800 scans. The spectral width was 59.5 kHz and the recycle delay was set to 30 s giving a total experiment time of 6.67 hours. The 1D $^{13}$C CP MAS experiments were acquired using 1024 points, an acquisition time of 8.60 ms and 800 scans. The spectral width was 59.5 kHz and a recycle delay of 3 s giving a total experimental time of 40 minutes. All CP experiments were acquired using a contact time of 1.2 ms with a CP ramp of 90-100 %. For the CP/$T_1$ experiments the $T_1$ dependent delay was set to 0.01 s in one and 30 s in the second to match the

$^{13}$C MAS experiment. All 1D experiments were Fourier transformed after zero-filling to 2048 data points using 100 Hz of exponential line broadening.

The 2D $^{13}$C-$^{13}$C MAS refocused INADEQUATE[46] was acquired using 1024 data points in the $F_2$ dimension and 192 increments in the $F_1$ dimension with 736 scans per increment. The spectral width was set to 48 kHz in both dimensions and a z-filter was used to remove any zero-quantum frequencies which can lead to phase distortions and artefacts in the spectrum.[47] The cross polarisation was achieved with a contact time of 1.2 ms and the evolution $\tau$ delay was set to 2.2 ms. The $^1$H decoupling was SPINAL-64 and the TPPI method was used for sign discrimination.[115] A recycle delay of 2 s was used giving a total experimental time of 79.9 hours. This was Fourier transformed in both dimensions after zero-filling to 512 points in $F_1$ with 100 Hz of exponential line broadening and 4096 in $F_2$ with 100 Hz of exponential line broadening. All experiments have been processed using Bruker TopSpin 3.5 and all fittings of spectra for quantitative analysis were undertaken using DMFit.[5]

## 4.2 Quantitative NMR

Solid-state NMR spectroscopy is a powerful technique for determining the structures of substances particularly in complex amorphous samples where diffraction techniques are not applicable. However NMR is an inherently insensitive method due to the small energy differences between nuclear spin states. This insensitivity is compounded when studying low abundance nuclei such as $^{13}$C which has a natural abundance of 1.1%.

Obtaining a fully quantitative NMR spectrum usually requires a recycle delay that is sufficiently long for the nuclear magnetisation to fully relax to thermal equilibrium between each scan. In order to achieve this the recycle

delay has to be at least 5 times the longest spin-lattice relaxation time, $T_1$. Biological and organic materials such as humic acids and plants can have relaxation times of the order of hundreds of seconds.[10] This, alongside the insensitivity of the carbon nucleus, results in quantitative experiments that require an unreasonably long time. In this section, the Schmidt-Rohr method has been chosen for the quantitative analysis.

### 4.2.1 Using the Schmidt-Rohr Quantitative Method to Study Humic Acid Compositions

Humic acids and plants are made up of a complex mixture of polymers with broad $^{13}C$ NMR line shape and if the sample has low paramagnetic content then the relaxation times can be very long. This means acquiring a quantitative spectrum of these samples is a time intensive task. Schmidt-Rohr *et al.* developed a method to calculate a quantitative analysis which make this process less demanding on experimental time. The quantitative method developed by Schmidt-Rohr and co-workers was first developed for studying the crystallinity of polymers but was soon applied to a wide range of humic acid samples. They used the Schmidt-Rohr approach, to calculate the $sp^2/sp^3$ ratio of each of the humic acid samples and used this as a measure to compare with the many proposed humic acid models. Schmidt-Rohr *et al.* showed that humic acid compositions vary considerably and no single model represents the composition of all humic acids.[10] In this work the Schimdt-Rohr quantitative method has been used to acquire compositional analysis of two different humic acid samples.

The $^{13}C$ DP (Direct Polarisation) MAS spectrum for the first humic acid sample is shown in Figure 4.8. The spectrum is extremely broad with the majority of the peaks overlapping significantly. The broad line

shape illustrates the complex amorphous structure of humic acids. This broad line shape can cause several issues when attempting to extract the quantitiative intensity of a particular environment. This means care has to be taken when defining the different shift regions appropriately. The samples studied here have a low paramagnetic content and so have long relaxation times.

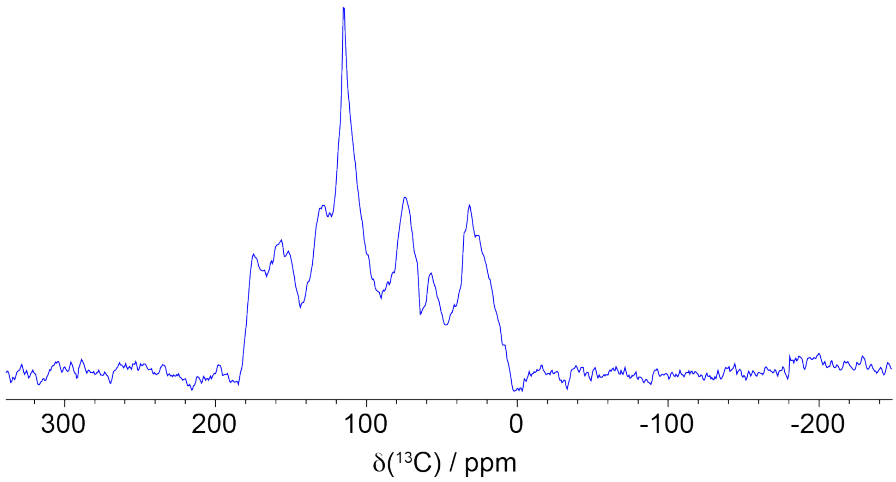


Figure 4.8: A $^{13}C$ (DP) MAS spectrum of the first humic acid sample, recorded with a recycle delay of 10 s which is insufficient for quantitative results.

The CP/$T_1$ spectra recorded with a short $T_1$ filter time of 1 $\mu$s and long $T_1$ filter time of 25 s to match the $^{13}C$ MAS experiment recycle delay are shown in Figure 4.9. The spectra show that the majority of the aliphatic sites have relaxed over the 25 s period. However, the carbonyl and aromatic components in the 100 - 200 ppm range still have a significant amount of signal remaining. This comparison is a good representation of how relaxation times can vary over the different chemical environments in the sample. Therefore, the signal in the $^{13}C$ DP MAS spectrum for the aliphatic region does not require much correction whilst the carbonyl and aromatic regions require significant correction of the peak areas for quantitative analysis.

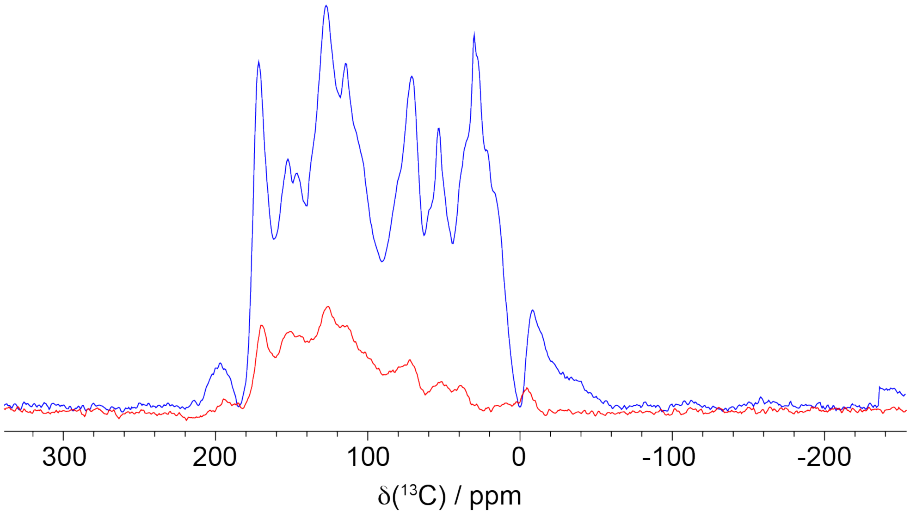


Figure 4.9: A comparison of two $^{13}C$ CP/$T_1$ spectra for the first humic acid sample, recorded with delays of 0.001 s (blue) and 25 s (red). Note the decay of intensity in the latter results from $^{13}C$ $T_1$ relaxation

Figure 4.10 shows a striking difference between the line shapes of the $^{13}C$ DP MAS spectrum and the CP/$T_1$ spectrum recorded with a short $T_1$ filter time. This comparison highlights how the CP transfer efficiencies of different chemical environments can affect the line shape, with much of the aliphatic region being over represented in the CP spectrum.

The $^{13}C$ MAS spectrum and the two CP/$T_1$ spectra were fitted visually using DMfit[5] with the same 75 Gaussian peaks as seen in Figure 4.11. The number of Gaussian peaks used is large however the Gaussian peaks had fixed width and position and only the intensity was allowed to change for each spectrum. The large number of Gaussian peaks enables the difference in the three spectra to be accounted for by changes in peak intensities and aids in the separation of the spectrum into 11 peak regions with the same peaks being associated with each region. These 11 regions are defined in the Table 4.1. The Gaussian peaks in each region were added together and the ratio of the two CP/$T_1$ peaks areas was used to correct the peak areas of the $^{13}C$ DP spectrum.

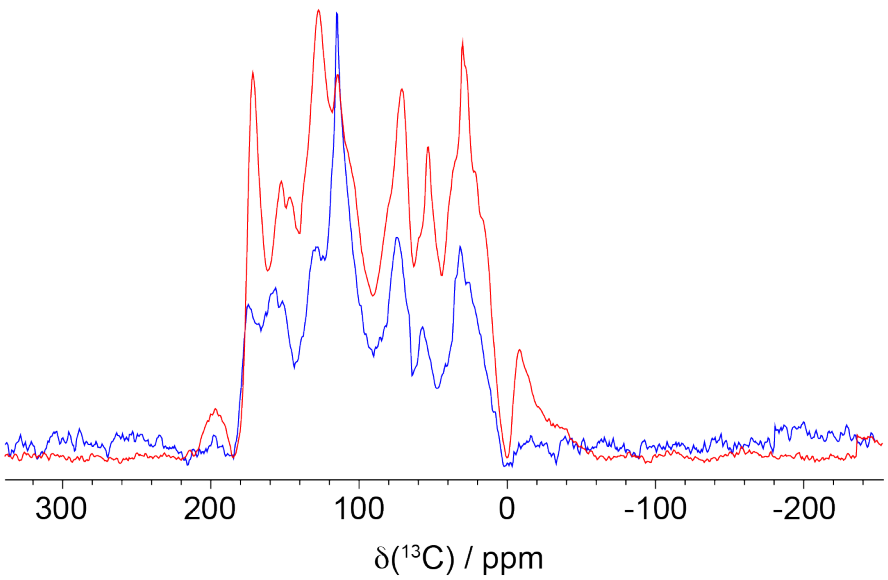


Figure 4.10: A comparison of the $^{13}C$ DP MAS spectrum (blue) and a $^{13}C$ CP/$T_1$ MAS spectrum with a short z-filter of 0.001 s (red) the first humic acid sample.

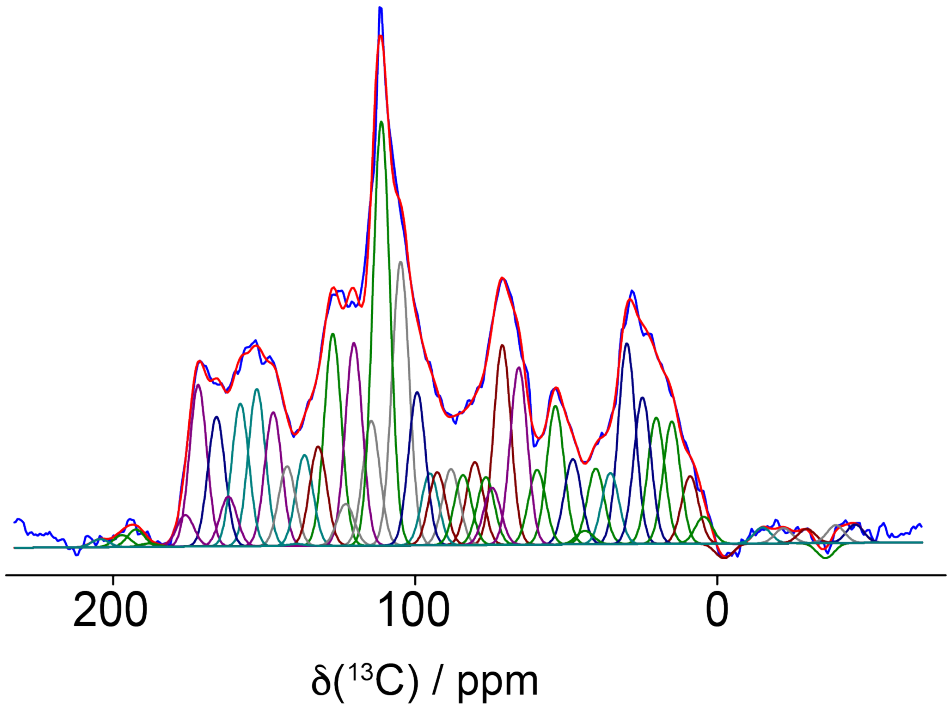


Figure 4.11: The $^{13}C$ MAS spectrum with the DMfit[5] fitting of 75 Gaussian peaks used for the quantitative analysis.

| No. | Range / ppm | Chemical Environment |
|---|---|---|
| 1 | >190 | Ketone, quinone, aldehyde |
| 2 | 162 - 190 | Carboxyl, ester, quinone |
| 3 | 145 - 162 | Phenolic |
| 4 | 120 - 145 | Aromatic |
| 5 | 108 - 120 | Aromatic |
| 6 | 96 - 108 | Anomeric, aromatic carbon neighbouring phenolic carbons |
| 7 | 60 - 96 | Saccharide, alcohol, ether |
| 8 | 50 - 60 | Methoxy, methyne, quaternary |
| 9 | 35 - 50 | Complex aliphatic |
| 10 | 25 - 35 | Methylene in simple aliphatics |
| 11 | 0 - 25 | Methyl |

Table 4.3: The chemical shift ranges of the 11 peak areas with a assignment of their general chemical environment.[10]

The corrected peak areas that have been used to calculate the quantitative analysis are shown in the Table 4.2. The ratio of the $sp^2$/ $sp^3$ hybridized carbons was also calculated since this ratio is known to provide an insight into the properties and reactivity of the humic acid. This is calculated by finding the ratio of carboxylic and aromatic groups area (¿ 108 ppm) with that of the carbohydrate and aliphatic peak areas (¡ 108 ppm). The $sp^2/sp^3$ ratio of 0.9 $\pm$ 0.1 means that there is a higher proportion of carbohydrate and aliphatic groups within this humic acid than carboxylic and aromatic groups. There are many humic acid composition models some of which have been briefly introduced earlier in the chapter. Using the list of $sp^2/sp^3$ ratios of the main humic acid models compiled by Schmidtt-Rohr *et al.*[10] this humic acid's composition could fit with either the Leeheer's or the model developed by Schulten and Schnitzer.[75,76,79] However, when looking in greater detail into the predicted model compositions the Leeheer model agrees best as it has a higher carbohydrate content than the other. The Leeheer structural model suggests that this composition has approximately equal amounts of phenol and aromatic carbonyl groups that can interact via hydrogen bonding. This is thought to result in phenol groups

| Functional group | Chemical Shift Range | Composition % |
|---|---|---|
| Aliphatic | 0 - 50 ppm | 17.0 ± 2.4 |
| Carbohydrate | 50 - 108 ppm | 30.0 ± 1.7 |
| Aromatic | 108 - 162 ppm | 42.0 ± 1.4 |
| Carboxylic | > 162 ppm | 10.0 ± 0.6 |
| $sp^2/sp^3$ ratio | area( > 108 ppm)/area(0 - 108 ppm) | 0.9 ± 0.1 |

Table 4.4: The quantitative analysis of the first humic acid sample.

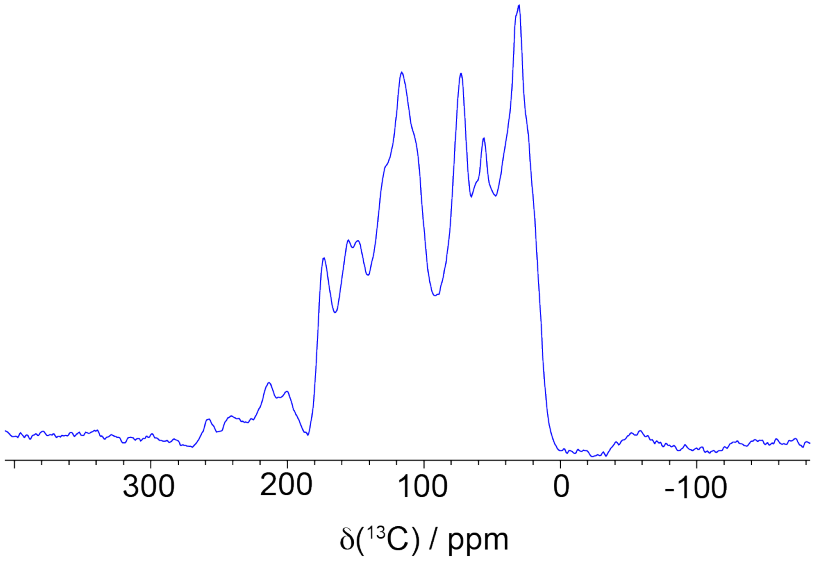


Figure 4.12: An $^{13}C$ DP MAS spectrum of the second humic acid sample recorded with a recycle delay of 25 s.

adjacent to the aromatic carbonyl groups that lead to the phenol groups behaving as weaker acids than expected.[10,79]

The second humic acid spectrum appears completely different showing the variation that can be observed from different humic acid samples. The $^{13}C$ DP MAS spectrum is shown in Figure 4.12 and the overlay of the CP/$T_1$ spectra is shown in Figure 4.13. The spectrum for this humic acid is again very broad and the relaxation times, whilst shorter than the first, are still long enough to warrant the Schmidt-Rohr method for quantitative analysis.

The Schmidt-Rohr method was applied in the same way with the same 11 peak regions and fitted visually with 68 Gaussian peaks in DMfit.[5] The

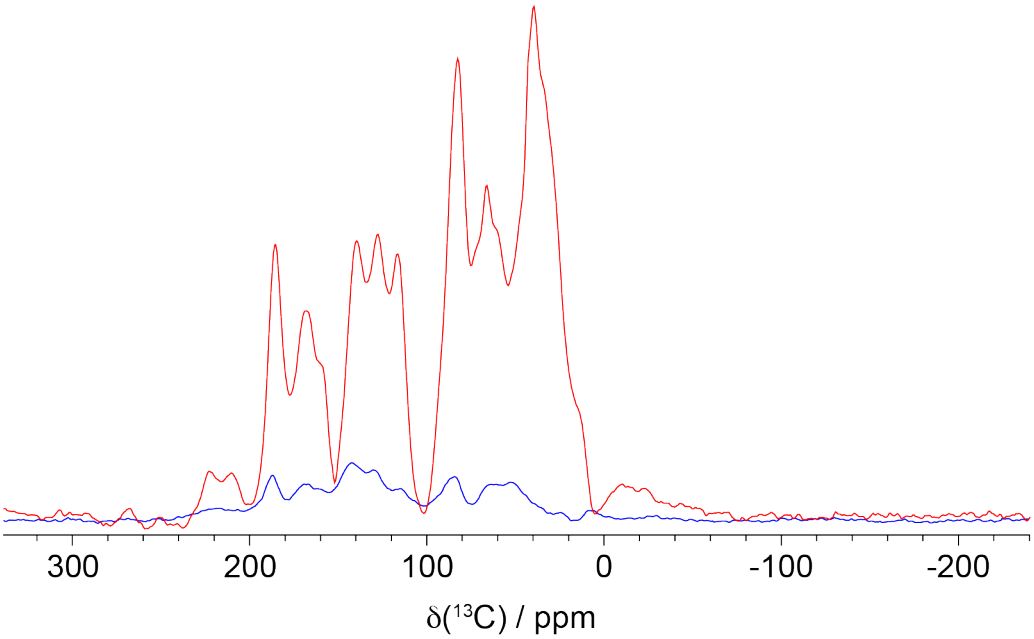


Figure 4.13: A comparison of two $^{13}C$ CP/$T_1$ spectra for the second humic acid sample, recorded with delays of 0.001 s (blue) and 25 s (red)

| Functional group | Chemical Shift Range | Composition % |
|---|---|---|
| Aliphatic | 0 - 50 ppm | 11.0 ± 1.5 |
| Carbohydrate | 50 - 108 ppm | 30.0 ± 1.1 |
| Aromatic | 108 - 162 ppm | 43.0 ± 1.3 |
| Carboxylic | > 162 ppm | 15.0 ± 0.7 |
| $sp^2/sp^3$ ratio | area( > 108 ppm)/area(0 - 108 ppm) | 1.4 ± 0.05 |

Table 4.5: The quantitative analysis of the second humic acid sample.

corrected peak areas were calculated and the quantitative analysis is shown in Table 4.5.

The quantitative analysis shows that this humic acid has more carboxylic content than the first and this is reflected in an larger ratio of $sp^2/sp^3$ indicating this structure is more likely to fit the model developed by Jansen *et al.* or the Dragonov model. Again when looking closer at the constituent break down of both models the Dragonov model fits best as the ratio between carbohydrate sites and aromatic sites is similar to the sample. This structural model suggest the main structure has bridged aromatic rings that are joined together via covalent bonding with carbohy-

drate residue, indicating a much more covalent structure than that of the the first humic acid analysed.[10]

## 4.3 Wheat Straw Composition Analysis by Solid-state NMR

This section focuses on using solid-state NMR to analyse the composition of wheat straw and using it to identify the key components within the straw. Composition of plants can vary depending on strain, age, stage of growth and part of the plant.[3] This means that throughout the literature there is a range of compositions quoted for wheat straw. It is therefore necessary, especially for further work, to obtain a compositional analysis of this wheat straw sample rather than assuming literature values. The composition obtained using solid-state NMR will be compared with the range of values known in literature. The composition of wheat straw varies with age, strain and part of the straw and so there is a significant range in literature. Typically in wheat straw, cellulose content is found to be between 31-48.5 %, hemicellulose between 16.7-49 % and liginin between 8-25 % as well as small amounts of protein and pectins.

An initial 1D $^{13}$C CP MAS spectrum was acquired to obtain a preliminary assignment of the straw shifts. There was some concern that there would be too much overlap of chemical shifts to be able to identify many of the key components of the straw, therefore limiting the information that can be extracted from the quantitative analysis. Figure 4.14 shows the 1D $^{13}$C CP MAS spectrum where some assignments have been made using literature predominantly based on the *Arabidopsis thaliana* model plant.[4]

In general the spectrum can be separated in to 5 main regions; carbonyl (170-190 ppm), aromatic (110-170 ppm), neutral carbohydrate (60-

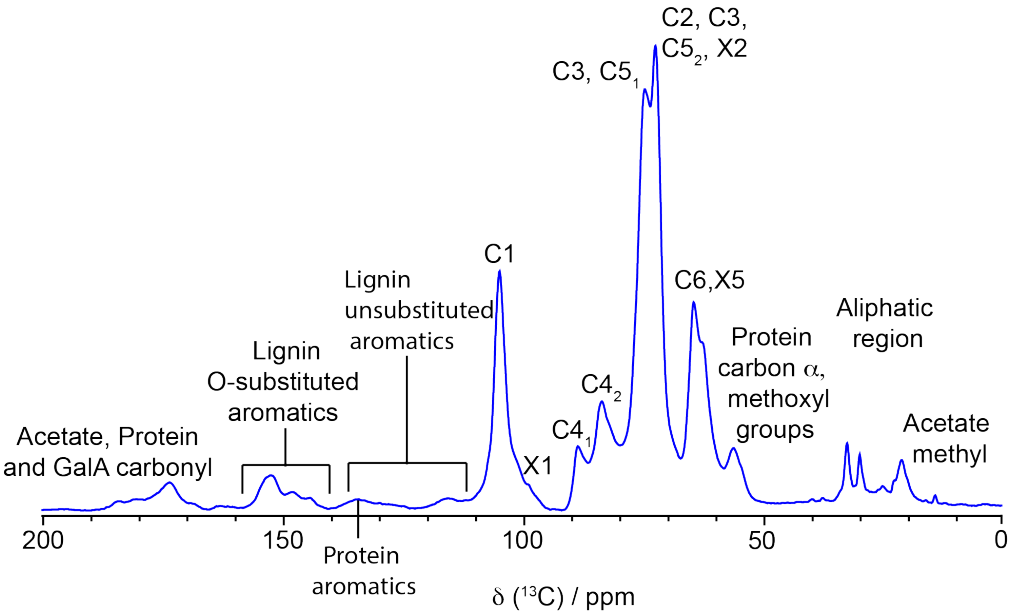


Figure 4.14: 1D $^{13}$C CP MAS spectrum of wheat straw with assignments of the key shifts.

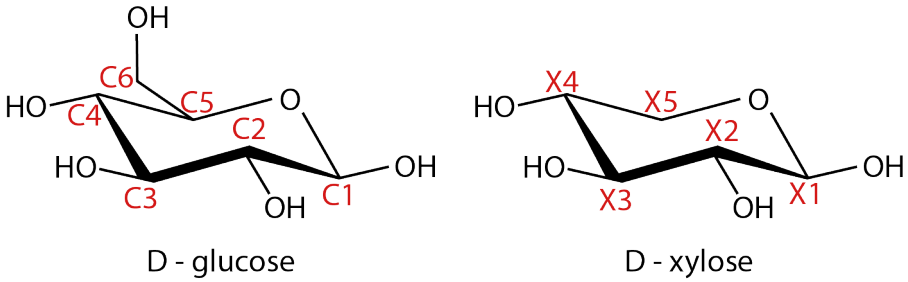


Figure 4.15: Glucose and xylose are the monomers of the prevalent cellulose and xylan polymers within the straw. The carbons in the monomers have been labelled to aid with understanding of the assignment of the $^{13}$C NMR shifts in the spectrum.

110 ppm), methoxyl groups/ protein C$\alpha$ (50-60 ppm) and aliphatic (0-40 ppm).

In the carbonyl region there is a small peak that stretches across the 20 ppm range and appears to have several contributions of signal from different constituents in the straw. The largest contribution of this carbonyl signal is the carbonyl of the acetate that decorates the sides of pectin and the main hemicellulose component in the straw, xylose which makes up between 15-30% of the wheat straw. There are also likely to be contributions from the

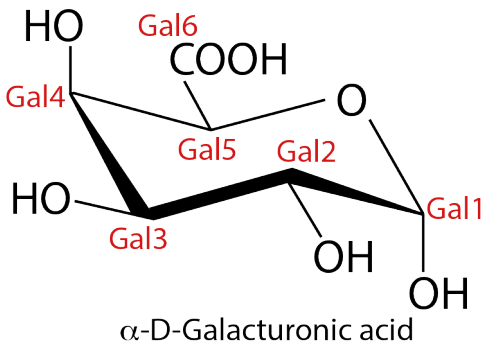


Figure 4.16: Galacturonic acid structure that forms the the backbone of pectin in the wheat straw. The carbons in the molecule have been labelled to aid with understanding of the assignment of the $^{13}C$ NMR shifts in the spectrum.

carbonyls associated with the small amount of protein in the straw and the carbon 6 of uronic acids such as galacturonic acid (Gal6) that form the backbone of pectin which make up around 2-5% of the wheat straw.[4] The majority of the peaks for lignin can be found in the 110-160 ppm range, where 140-160 ppm is assigned to O-substituted lignin aromatics.[4,116?] Unsubstituted lignin aromatics can be assigned to the 120-140 ppm range with a small amount of protein aromatics underneath in the range 125-140 ppm.[4,116,117]

The neutral carbohydrate region, 60-110 ppm, is very congested with peaks from the cellulose and hemicellulose that form the majority of the straw. It can be seen in Figure 4.16 that the monomers for cellulose and xylose are very similar which leads to overlapping signals causing problems when attempting to assign the exact shifts in the spectrum. Due to the main constituent in both primary and secondary cell walls being cellulose the majority of the xylan shifts are masked by the large cellulose signal. Many of the cellulose peaks have been assigned with C1 at 105.1 ppm, C4 at 83.0 and 89.0 ppm and C6 at 58.5 and 67.5 ppm. The signals for C2, C3 and C5 are very close together and fall within the 67.5-80 ppm range.[4,118,119]

The $2_1$ helical cellulose forms large microfibril structures held together with an extensive network of intra and inter molecular hydrogen bonding. Within these structures it is been shown that there are at least two main cellulose environments present despite the debate about their origin.[120] It is possible to observe these two different cellulose environments in the 1D $^{13}$C CP MAS spectrum since there are two signals for the C4 (83.8 and 88.8 ppm) and C6 (62.8 and 64.7 ppm) sites. The C5 carbon is also known to shift from 72.0 ppm to 75.0 ppm, however, it is not possible to observe this here, because of the large number of peaks within the 70-80 ppm range. The $2_1$ helical conformation of the cellulose means the C4, C5 and C6 carbon environments are all along one side of the 2-fold helix. This means they all could be affected by the same changes in the environment. It has been shown that small differences in the C6 glycosidic bond conformation can affect the shifts of all 3 carbon environments by several ppm.[4,118,121] The cellulose signals overlaps with the majority of the shifts for the main hemicellulose constituent xylan. It is possible to distinguish the X1 carbon at around 102.6 ppm as a small shoulder on the C1 carbon signal. This X1 shift has been recently assigned to the 3-fold conformation of xylan, whilst the 2-fold conformation xylan has an X1 shift of 105.0 ppm like the cellulose.[63]

At 50-60 ppm there is a single peak which can be assigned to the $\alpha$ carbon of the protein along with the methoxyl groups associated with lignin. Finally, within the aliphatic region, the main peak that can be assigned is at 21.0 ppm which is the methyl group of acetate that decorates xylan and pectin. The rest of the aliphatic shifts are usually ignored in the literature but are generally assigned to glycoprotein side chains.[122]

Despite the significant overlap of sites in the wheat straw spectrum it was still possible to apply the Schimdt-Rohr quantitative analysis method.

| | Cellulose and Xylose | Lignin | Protein |
|---|---|---|---|
| Frequency range / ppm | 60-110 | 140-160 | 50-60 |
| Literature values | 50-80 % | 8-20 % | 3-5 % |
| Calculated values | 77.8 % | 7.8 % | 3.2 % |

Table 4.6: Quantitative analysis of wheat straw results using solid-state NMR compared to literature ranges.[11–16]

The spectra were fitted manually with 28 Gaussian peaks in DMfit[5]. The $^{13}C$ MAS spectrum peak areas were then extracted and corrected for incomplete relaxation effects. The composition analysis was calculated from these corrected peak areas and the final results are summarised in Table 4.4. The table includes a range in which previous literature composition data has been found for wheat straw.[11–16] The majority of the quantitative analysis obtained using the solid-state NMR method are within the literature value ranges, with the exception for lignin which is at the minimum. However, the amount of lignin depends on the strain of wheat straw studied and the age of the plant. The process of creating lignin in the plant (known as ligninfication) only occurs in mature plants causing older plants to have a larger lignin content.[70,123] Lignin is a rigid polymer and so has an extremely long $T_1$ time, which means in the quantitative method the lignin signals require a larger correction to their peak areas, this could result in an underestimate of the lignin in the straw.

The majority of the 1D spectrum has overlapping shifts which makes it difficult to fully characterise the wheat straw. The wheat straw sample studied in this work is uniformly $^{13}C$ labelled making 2D $^{13}C$-$^{13}C$ correlation experiments feasible. In this case it was chosen to acquire a 2D refocused INADEQUATE spectrum as it correlates directly covalently bonded carbons and separates overlapping signals in a second double quantum frequency dimension.

The full 2D refocused INADEQUATE spectrum can be seen in Figure

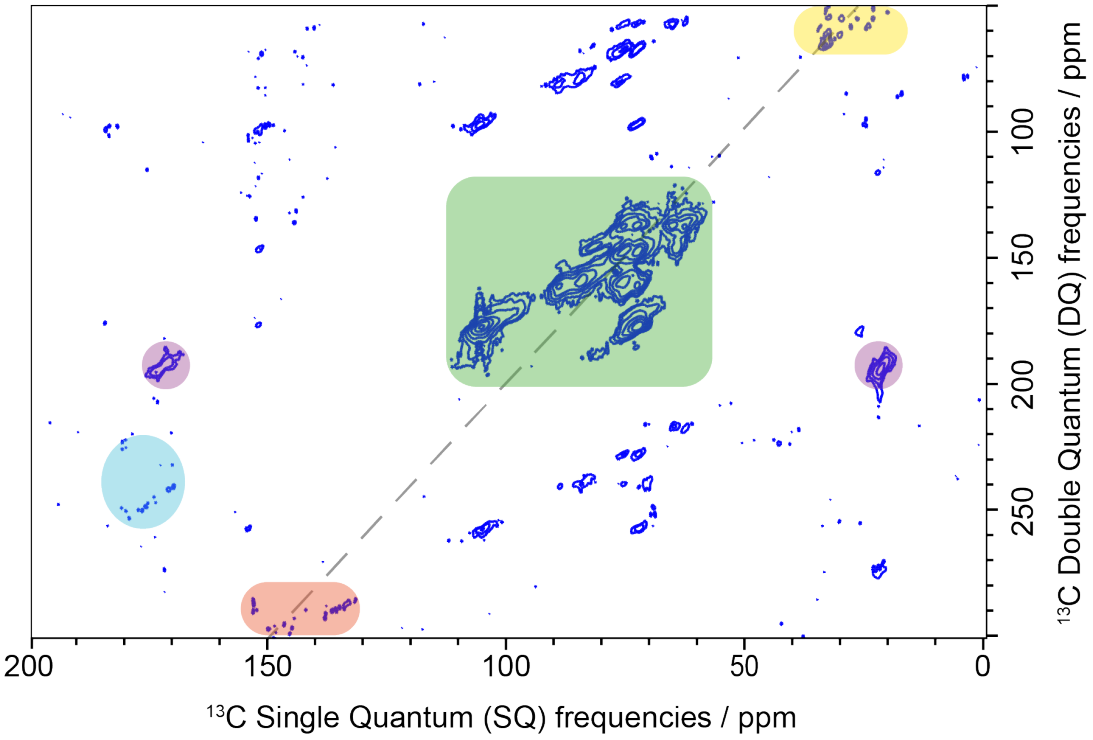


Figure 4.17: Experimental $^{13}$C MAS 2D refocused INADEQUATE spectrum with key areas of the spectrum shaded shows the aliphatic region(yellow), neutral carbohydrate region (green), acetate carbonyl and methyl correlation (purple), protein carbonyls (blue) and the aromatic peaks (orange)

4.17 with the main peak areas highlighted. The correlation between the acetate carbonyl at 172.5 ppm and methyl at 21.0 ppm can be clearly seen at a double quantum frequency of 193.5 ppm. The peak carbonyl of protein can be seen at 174.0 ppm and the corresponding double quantum shift at 250.0 ppm indicates that the corresponding peak should be at 76.0 ppm which is unfortunately where the spinning sidebands of the neutral carbohydrate region fall in the spectrum. The refocused INADEQUATE used CP to increase sensitivity which means quaternary carbon sites found in lignin tend to be under represented. This results in no observable correlations of lignin in the spectrum. In this region there are only a few low intensity aromatic correlations at 135.0 -153.0 ppm and DQ frequency of 287.0 ppm, which are assigned to the aromatic carbons of protein.[4] The neutral carbohydrate region is where the majority of the correlations reside in the spectrum. It is more convenient to look at this region in more detail to be able to identify and assign all of the correlations as shown in Figures 4.18, 4.19 and 4.20.

The second dimension separates almost all the cellulose and hemicellulose peaks that were otherwise unresolved in the 1D spectrum. The correlations of the 2D NMR allows all of the cellulose carbon signals within 70-80 ppm to be assigned with the C2 at 72.2 ppm, C3 at 74.8 ppm and C5 at 72.5 and 75.3 ppm. The complete assignment of all the cellulose shifts for both cellulose environments can be made by using the correlations in the 2D spectrum. The shifts for the 2 cellulose environments have been identified in Figures 4.18 and 4.19

The first cellulose environment shown in Figure 4.18 is assigned as C1 at 105.0 ppm, C2 at 71.8 ppm, C3 at 75.0 ppm, C4 at 84.0 ppm, C5 at 75.3 ppm and C6 at 62.5 ppm which is generally called the surface cellulose environment. The second cellulose environment shown in Figure

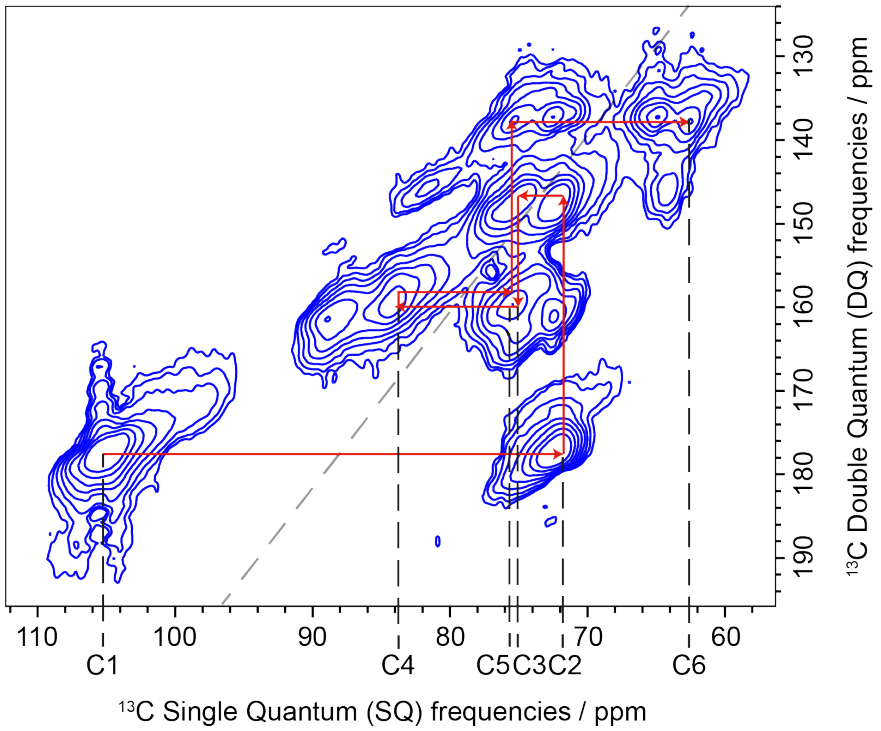


Figure 4.18: $^{13}$C CP MAS 2D refocused INADEQUATE spectrum neutral carbohydrate region with the first main cellulose environment shifts labelled

4.19 has the same C1, C2 and C3 peaks, but C4 and C6 signals have shifted downfield to 88.5 ppm, and 66.0 ppm, respectively. The C5 carbon peak also shifts upfield from 75.3 ppm to 73.0 ppm. This cellulose site is generally referred to as the interior cellulose site.

The additional resolution of the 2D NMR experiment also enables several xylose carbon shifts to be identified despite being obscured by the large cellulose peaks in the 1D spectrum. The xylose peaks have been assigned in Figure 4.20 with X1 at 102.6 ppm and X2 at 69.5 ppm assigned to the 3-fold conformation of xylan as the 2-fold X1 and X2 shifts are the same as the cellulose shifts. Whilst the 2-fold xylan X4 and X5 shifts can be resolved at 82.2 ppm and X5 at 64.0 ppm respectively. Unfortunately it is still not possible to resolve either X3 shift, but a literature value for the $^{13}$C shift this site in xylan has been found to be at 72.9 ppm.[63,69] The large intensity of the xylan X4 and X5 peaks indicate that the majority of the

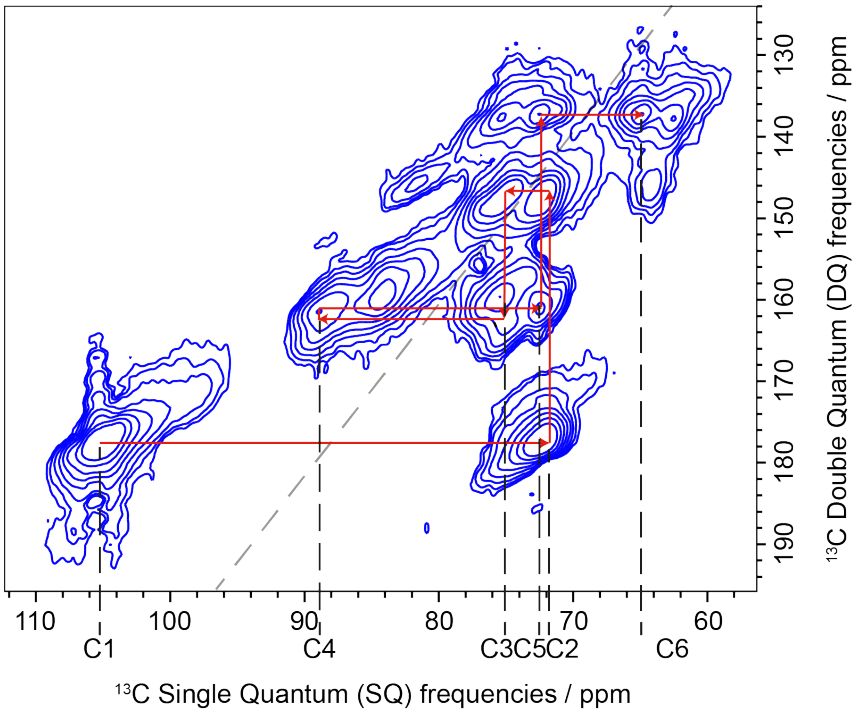


Figure 4.19: $^{13}$C CP MAS 2D refocused INADEQUATE spectrum neutral carbohydrate region with the second main cellulose environment shifts labelled

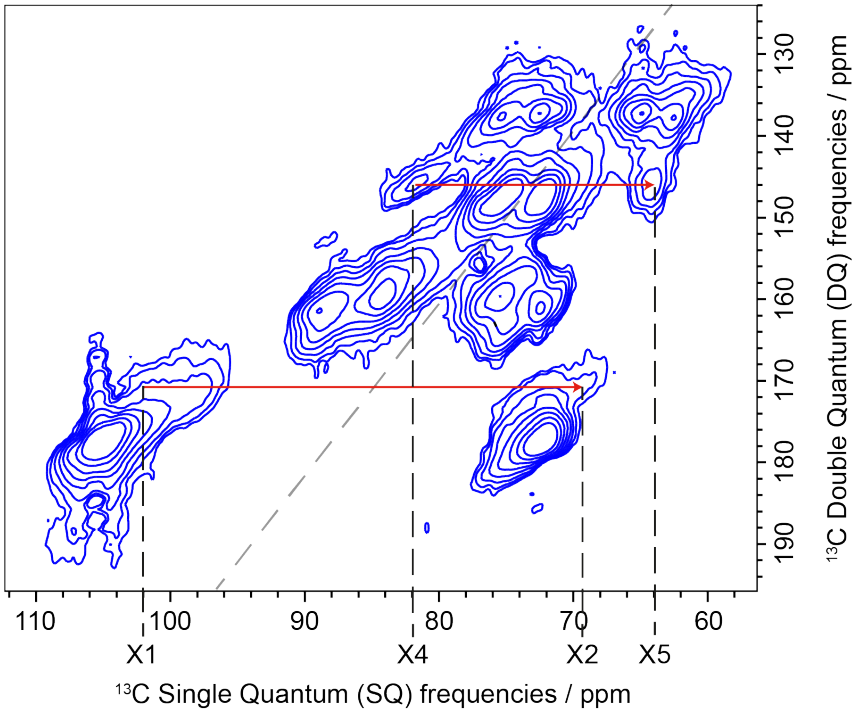


Figure 4.20: $^{13}$C CP MAS 2D refocused INADEQUATE spectrum neutral carbohydrate region with the xylose environment labelled

xylan forms a 2-fold helical structure. This structure allows the xylan to be able to sit along the cellulose microfibril with its chain decorations such as acetate along one side. There is also a small amount of the 3-fold xylan that is not bound to the cellulose in the wheat straw but the predominant form is the 2-fold xylan.[63,69]

### 4.3.1 Conclusions

This work has shown that despite the large overlap of sites it is possible to obtain quantitative analysis of the straw using solid-state NMR. The acquisition of 2D experiments has provided additional resolution enabling assignments of the 2 different cellulose environments as well as that of the xylose. The shifts of the xylose carbons tells us that the majority is in a 2-fold helical structure which would allow it to bind to the large cellulose microfibril.

## 4.4 Using Quantitative and 2D NMR Analysis to Study the Fermentation Process of Wheat Straw by a Fungus, Aspergilus Niger.

The presence of lignin in wheat straw means that the production of biofuels using standard methods is not economically viable. Lignin encases the cellulose in the straw making it hard to extract the fermentable sugars. In order to make the production of biofuels using waste materials, such as wheat straw, possible a pretreatment method is necessary to break down the straw.[124] Fermentation has been identified as a potential pretreatment method as it does not require high energy input and it produces cellulosic enzymes that can be used in the hydrolysis stage to produce the biofuel.

*Aspergilus niger*, also known as black mould, is a fungus that breaks down cellulose and hemicellulose. It is thought that over the fermentation period this fungus could break down some of the cellulose leading to an increase in the surface area which, in turn, will improve the efficiency of the final hydrolysis stage.[3]

The aim of this section is to study the changes that occur in the wheat straw over the fermentation process using quantitative and two dimensional solid-state NMR experiments.

### 4.4.1 Aspergilus Niger NMR Assignments

The wheat straw sample that has been fermented is 97% $^{13}$C labelled whilst the fungus is at natural abundance. This means that initially NMR signals from the fungus are unlikely to be seen. When the fungus degrades and consumes the labelled wheat straw it will grow and become partially

labelled itself. This could mean further in the fermentation process it is likely signal of the fungus could start to appear and interfere with the analysis of the changes in the wheat straw. The main constituent of the cell wall of *Aspergilus niger* is glucan which, like cellulose, is a polymer of glucose. In the fungi cell wall however, the majority of the glucose monomer is connected via a $\beta$- (1-3) linkage rather that the $\beta$-(1-4) linkage seen in cellulose. The second main constituent of the cell wall is a carbohydrate known as chitin. A 1D $^{13}$C CP MAS spectrum has been acquired of *Aspergilus niger*, shown in Figure 4.21, and the main chitin peaks have been assigned where possible from the literature.[9] The carbonyl shift of chitin at 174.8 ppm indicates that it is present in the fungal cell wall in the non-acetylated form. However, there are several polymorphs of chitin present leading to broader peaks for the chitin sites.

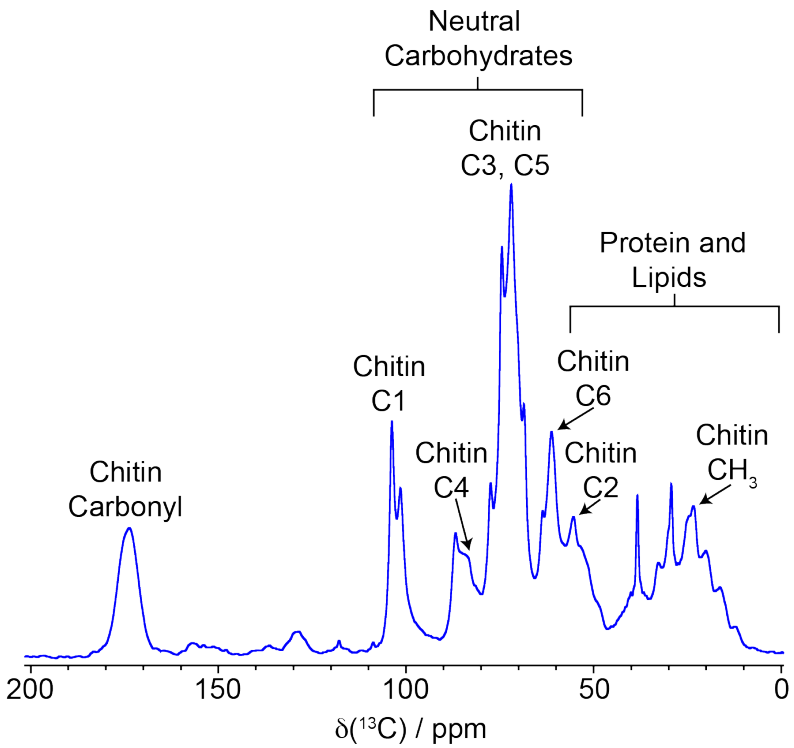


Figure 4.21: $^{13}$C CP MAS 1D spectrum of Aspergillus niger with the key chitin carbon shifts assigned. This spectrum was acquired with 19200 scans and a relaxation delay of 3 s. The $^{1}$H 90° pulse was 2.5 $\mu$s and a CP contact time of 1.5 ms used with SPINAL-64 decoupling during acquisition.[6]

The methyl site of chitin is seen at 22.0 ppm. The neutral carbohydrate region is the most congested as is the case for wheatstraw. It is possible to assign Chitin C1 at 103.6 ppm, C2 at 55.0 ppm, C4 at 83.0 ppm and C6 at 61.0 ppm, whilst the C3 and C5 peaks are expected in the large central peak at 73.0 ppm and 75.7 ppm respectively.[9] Over 80% of the glucan chain in *Aspergillus niger* has a $\beta$-(1-3) linkage which means the NMR shifts are slightly different to those for cellulose. The remaining glucan chains are a mixture of $\beta$-(1-4) and $\beta$-(1-6) linkage. These different linkages have very similar shifts and overlap such that in a 1D experiment it is hard to fully assign, however, the glucan shifts have been comprehensively assigned by Kang *et al.*.[9] The majority of these shifts are close to the cellulose shifts which will make separating these shifts in 1D NMR spectrum near impossible. However in 2D refocused INADEQUATE spectrum they have different double quantum frequencies which will distinguish them from the wheat straw cellulose. The fungus is not $^{13}$C labelled in this work so it is not likely to affect the NMR analysis significantly until the later stages of the fermentation period where it has broken down some of the labelled wheat straw but will be able to judge the growth of the fungus by identifying the narrow chitin shifts in the 2D refocused INADEQUATE spectrum.

### 4.4.2 1D Quantitative Analysis

First, it was necessary to see if it is possible to identify any changes occurring over a 5 day fermentation period using solid-state NMR. This was done by acquiring $^{13}$C MAS spectra at 0, 12, 24, 48, 72 and 120 hours fermentation time. Figure 4.22 shows key regions of the spectrum at 0, 2 and 5 days into the fermentation time where significant changes can be seen in the spectra. Figure 4.22a shows the carbonyl region where a significant change occurs at 169.0 ppm as a new peak grows over the fermentation

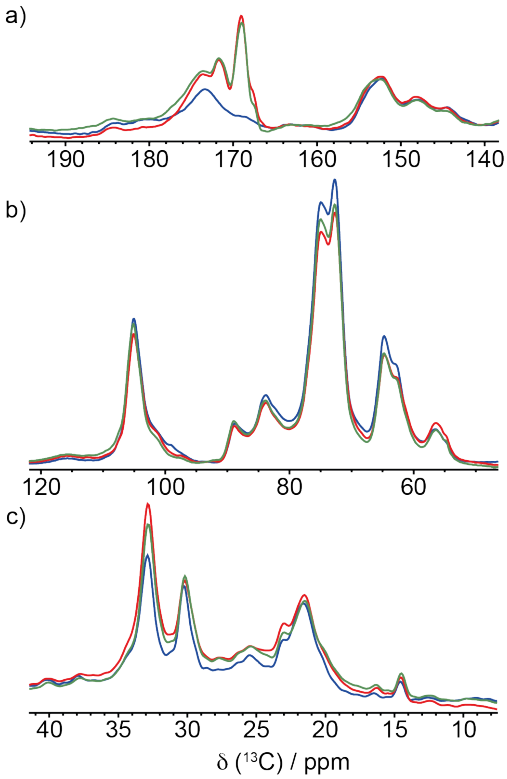


Figure 4.22: $^{13}$C MAS 1D spectra of wheat straw at t = 0 days (blue), t = 2 days (green), t = 5 days (red). Key differences that can be identified in 1D MAS spectra, recorded as a function of fermentation time, a) - carboxylic region, b) - neutral carbohydrate region, c) - aliphatic region.

period. This new peak appears within. the first 48 hours of fermentation and the remains relatively unchanged for the remaining fermentation. It is known in literature that the production of enzymes by the fungus is the most active in the first 48 hours of the fermentation. This suggests that this new peak at 169.0 ppm is likely related to the enzymes produced by the fungus such as the glucose oxidase (GOx or GOD) enzyme. In the neutral carbohydrate region seen in Figure 4.22b it is possible to observe a notable decrease in the cellulose peak intensities over the fermentation time while the protein $\alpha$ carbon peak at 55.0 ppm increases over this same period. Finally in the aliphatic region there seems to be a wide range of changes likely in part due to increased glycoprotein side chains from the growth of the fungus alongside degradation of the wheat straw.

The initial spectra obtained over the fermentation time show changes can be observed over this period by solid-state NMR. The Schmidt-Rohr method was then used to obtain a quantitative analysis of the wheat straw at the different points in the fermentation period. The $^{13}C$ MAS spectra were fitted with 82.0 peaks and the peak areas were corrected for relaxation effects. The large number of peaks were necessary to be able to cover the changes that occur in the spectra during the fermentation time. This allows the same peaks to be used throughout all the spectra acquired over the fermentation period ensuring that the peak regions are comparable.

During fermentation there is a net carbon loss due to the production of $CO_2$. This can be estimated by measuring the change in biomass recovered as shown in Figure 4.23. The carbon loss over the fermentation period has been reflected in the normalisation of the total peak areas in the quantitative analysis.

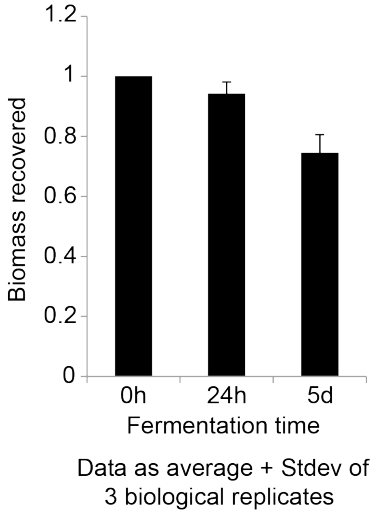


Figure 4.23: The change in the biomass recovered over the fermentation time used as an estimate for the carbon loss throughout the fermentation time in the quantitative analysis.

The quantitative analysis of the peak assigned to the cellulase enzymes at 169.0 ppm shows that the peak grows very quickly in the first 24 hours of fermentation and then remains constant level for the rest of the fer-

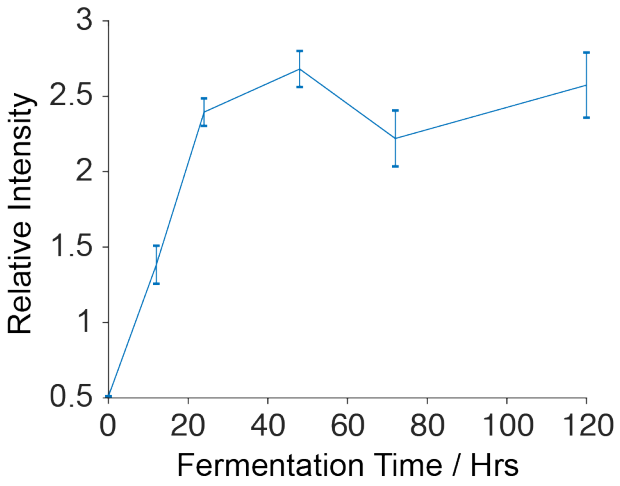


Figure 4.24: Quantitative analysis showing the growth of cellulosic enzymes from the peak at 169.0 ppm.

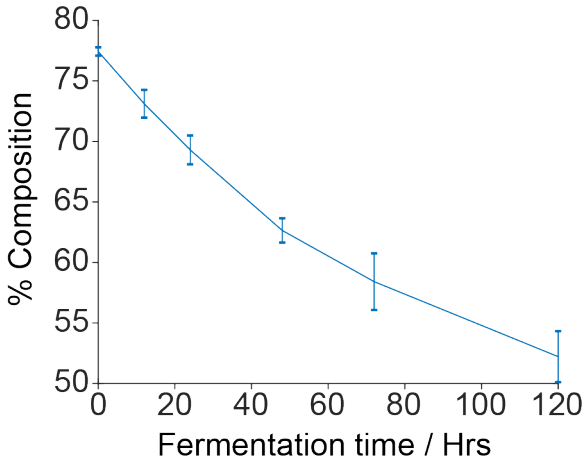


Figure 4.25: Quantitative analysis showing the decrease of cellulose over the fermentation time from peaks between 60-110 ppm.

mentation time. This early production and plateau of the enzymes is well documented and consistent with the literature.[3] The second main change is the decrease of the cellulose peaks over the fermentation time. The quantitative analysis of the whole of the neutral carbohydrate region in Figure 4.25 shows that the cellulose and hemicellulose decrease by around 25% over the 5 day period. It is important to calculate how much the hemicellulose decreases during the fermentation period relative to the cellulose. To gauge this, the changes in the C1 and X1 carbon peaks over

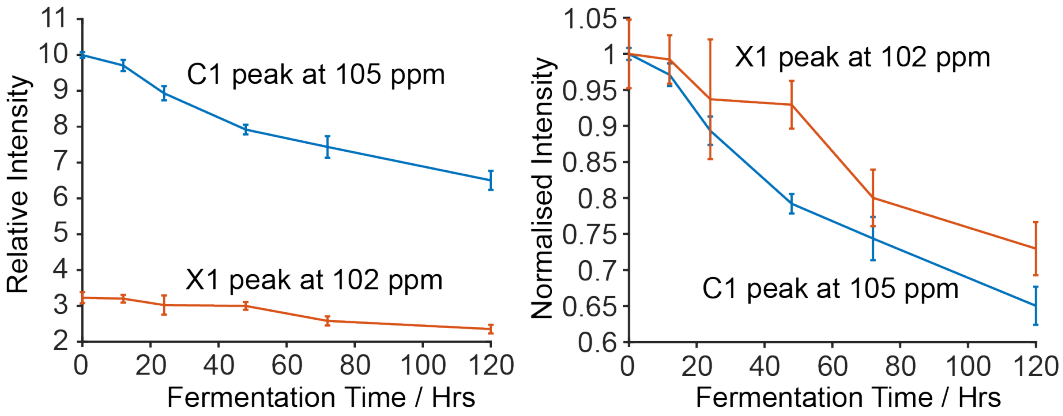


Figure 4.26: Two quantitative analysis plots showing the decrease in cellulose compared to xylan using the C1 at 105.0 ppm (blue) and X1 at 102.0 ppm (red) peaks in the 1D spectrum. Left - relative peak intensity changes over the fermentation time. Right - Comparison of the normalised peak intensity where the peak areas were normalised to the relative area of the fermentation time t = 0.

the fermentation time were compared in Figure 4.26. When comparing the relative changes in the peak intensities there seems to be a more significant decrease in the cellulose during fermentation. However when comparing the normalised decrease in these two polymers the xylan is degraded only slightly less than the cellulose. The decrease in the xylan peak could also be masked by the peaks from the chitin and glucan chains of the fungus as it grows during the fermentation time. The decrease in the neutral carbohydrate region is mostly from the degradation of cellulose whilst there is some degradation of xylan. It was initially thought that all the xylan was degraded first before the cellulose by the fungus but this comparison suggests that this is not the case.

In the previous 1D and 2D assignments of wheat straw it is possible to distinguish between the two cellulose environments in the straw. To study the changes in the different types of cellulose during the fermentation period the two C4 peaks and C6 peaks changes were compared as shown in Figures 4.27 and 4.28. The relative intensity changes of cellulose sites show that there is one cellulose site that has a larger apparent decrease than

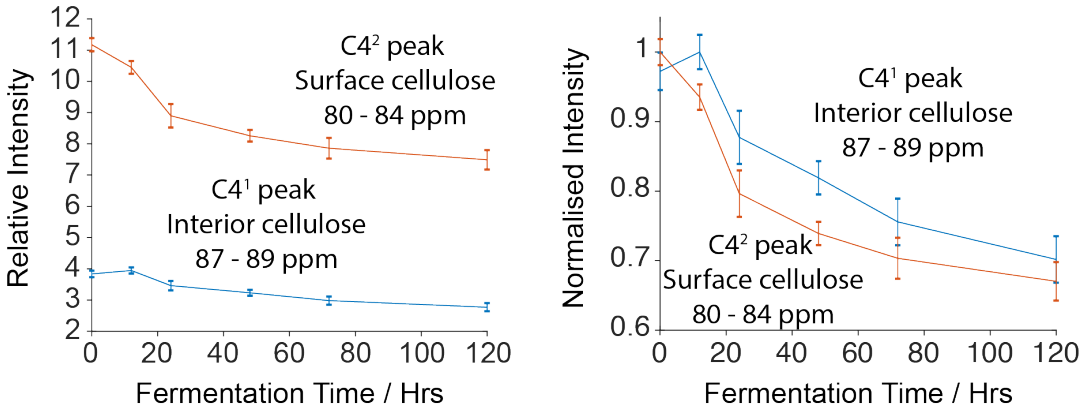


Figure 4.27: Two quantitative analysis plots showing the decrease in the 2 cellulose environments seen in the C4 shifts - $C4^1$ at 88.0 ppm (interior cellulose - blue) and $C4^2$ at 83.0 ppm (surface cellulose - red) in the 1D spectrum. Left - relative peak intensity changes over the fermentation time. Right - Comparison of the normalised peak intensity where the peak areas were normalised to the relative area of the fermentation time t = 0.

the other. This can be explained as the larger decrease is likely due to there being a xylan peak underneath the $C4^2$ and $C6^1$ signals. This means their relative decreases include xylan degradation too. The proportional change over the fermentation time shown on the right in both Figures 4.27 and 4.28 are more interesting. The proportional decrease of the two C4 cellulose environments indicate they decrease by an approximately equal proportion. There is an initial delay in the decrease in one of C4 and C6 regions but it is not the same cellulose environment. This initial delay of 12 hours can be explained as the peak that decreases straight away is the site with the xylan peaks underneath them. This suggests that some of the xylan is degraded initially and as the cellulose is exposed it is then degraded by the fungus.

The quantitative analysis gives a lot of insight into what is happening during the fermentation period. However, 1D spectra have a huge number of overlapping shifts and so it is likely that the 2D NMR may show changes that would otherwise be hidden underneath larger peaks.

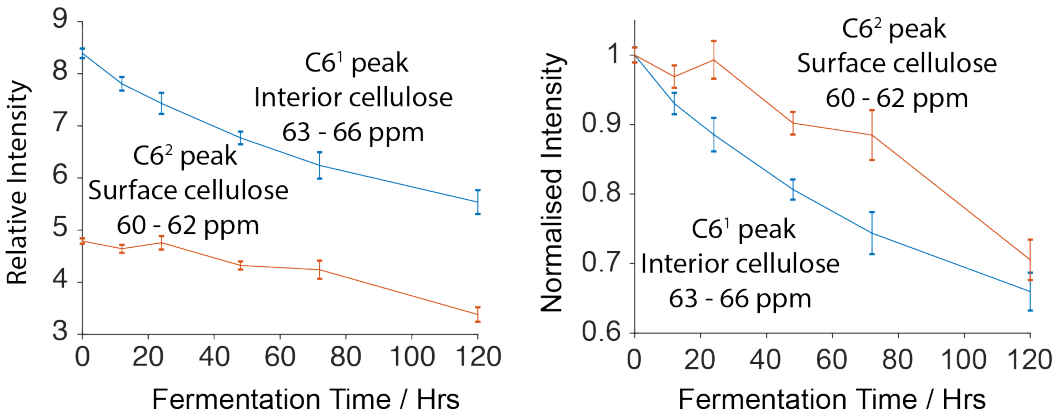


Figure 4.28: Two quantitative analysis plots showing the decrease in the 2 cellulose environments seen in the C6 shifts - $C6^1$ at 65.0 ppm (interior cellulose - blue) and $C6^2$ at 62.0 ppm (surface cellulose - red) in the 1D spectrum. Left - relative peak intensity changes over the fermentation time. Right - Comparison of the normalised peak intensity where the peak areas were normalised to the relative area of the fermentation time t = 0.

### 4.4.3 2D Refocused INADEQUATE Analysis of the Fermentation Process

In the first day of fermentation the quantitative analysis shows that the cellulase enzymes have been produced and reached an optimal concentration. This is a key point in the fermentation process and the changes in the wheat straw structure have been studied further by acquiring a 2D refocused INADEQUATE spectrum at 24 hours into the fermentation time. An additional 2D refocused INADEQUATE was also acquired at the end of the 5 day fermentation period to study the final state of wheat straw structure.

Figure 4.29 shows overlaid 2D spectra at fermentation times t = 0 (red) and t = 24 hours (blue) showing the changes within the straw by the first day of the fermentation. Initially, there is a new cross peak associated with the cellulase enzymes produced by the fungus at 169.0 ppm. The aliphatic region, however, does not show any changes in this 24 hour period. In the neutral carbohydrate region there have been more significant changes

within first day of fermentation to the cellulose and xylose environments as seen in Figure 4.30.

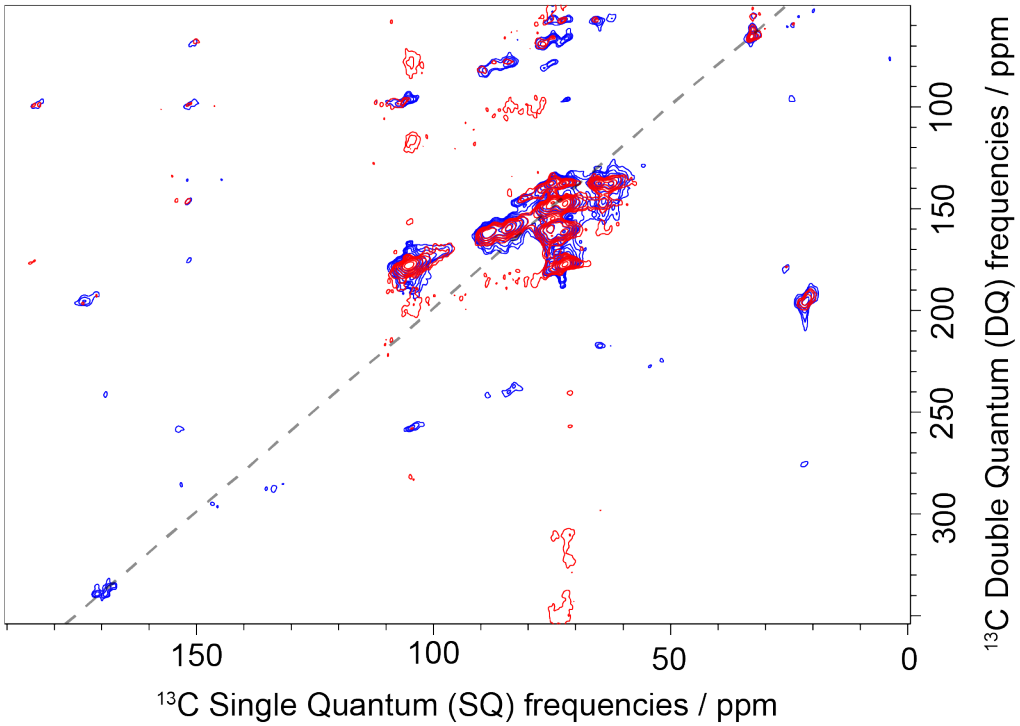


Figure 4.29: $^{13}$C-$^{13}$C 2D MAS refocused INADEQUATE spectra for t = 0 days (red) and t = 1 day (blue)

The comparison of the neutral carbohydrate region shows that all of the cellulose and xylose peaks have become significantly broader within the first 24 hours of fermentation. This could indicate that the fungus has begun to disrupt the structure of the cellulose and hemicellulose leading to more amorphous structures. Some of the cellulose broadening could also be due to the glucan polymers found in the fungus cell wall. The glucan peaks are likely to be very small this early stage in the fermentation process but their NMR shifts are very similar to cellulose so there maybe some contribution from the fungus.

There is a new cross peak at 71.1 ppm with a double quantum frequency of 154.4 ppm, meaning this new signal correlates with a shoulder of the broadened C4 peak at 83.5 ppm. This peak does not match any of the

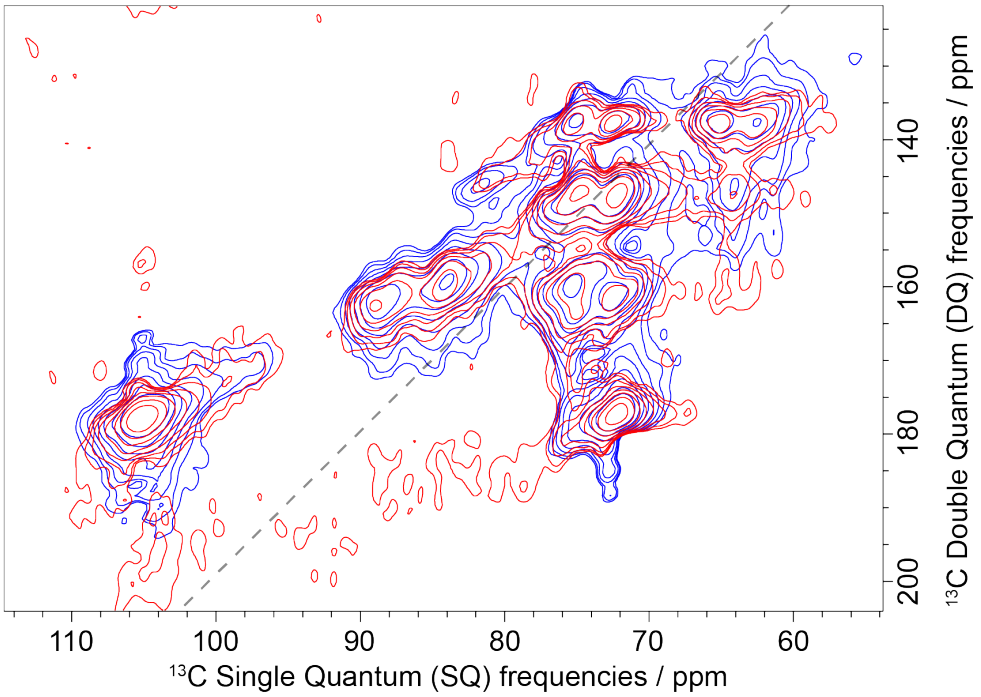


Figure 4.30: $^{13}$C-$^{13}$C 2D MAS refocused INADEQUATE spectra of the neutral carbohydrate region for t = 0 days (red) and t = 1 day (blue)

glucan or chitin peaks of the fungus given by Kang *et al.* and so this probably suggests a change that is occurring in the wheat straw itself. This is likely a change in the conformation of the C6 glycosidic bond by the disruption of the cellulose microfibril structure by the fungus, creating a shift in the C4 and C5 cellulose shifts.

There is evidence that the xylose environment has been affected in the first 24 hours of fermentation as the X4 peak has shifted downfield from 81.1 ppm to 83.0 ppm. This could indicate that some of the xylose is undergoing changes in its environment along side the changes in the cellulose microfibril. If the cellulose structure has been perturbed this could limit how effectively the xylan can lay across the microfibril leading to differences in the interactions with the cellulose. If the cellulose microfibril is more disordered then the binding with the xylan could be slightly different at different parts in the sample which could explain the significant broadening of the X4 peak. There is a small new peak at 77 ppm that correlates with

a peak at 65.0 ppm which has been assigned in literature to 3-fold helix xylan.[63] The presence of the 3-fold xylan suggests that the perturbations of the cellulose structure has lead to some of the xylose to become unattached from the cellulose. This results in the formation of a solution-like 3-fold helical xylan structure in the absence of cellulose.

The final refocused INADEQUATE spectrum that has been acquired allows the comparison of the changes in the cellulose structure over the whole 5 day fermentation period to be studied. The comparison of the wheat straw 2D spectra at t = 0 and t = 5 days is shown in Figure 4.31. At 169.0 ppm, as in the t = 1 day refocused INADEQUATE spectrum, a cross peak can be assigned to the degradative enzymes that have been produced by the fungus. By the end of the fermentation period there is a significant increase in the cross peaks seen in the aliphatic region. This increase is likely due to an increase in lipids and proteins produced as the fungus has grown. If it is possible to see some of the aliphatic correlations associated with the fungus it is likely other more abundant components will also be visible in the 2D spectrum, therefore care has to be taken when assigning new peaks. There is a slight shift in the acetate methyl peak which could be due to the chitin methyl and carbonyl correlation which is expected in the same region.

The neutral carbohydrate region is the main focus of the changes that have occurred over the fermentation period. The comparison of this region with the initial wheat straw has been shown in Figure 4.32. There is a new pair of cross peaks at 55.0 ppm and 104.3 ppm which have been assigned to the C1 and C2 carbons of chitin in the fungus. It is also possible to correlate this C2 peak to a peak at 75.0 ppm which is the C3 of chitin. These assignments of the chitin are illustrated in Figure 4.33. The cellulose peaks, as observed in the t = 1 day spectrum, are significantly

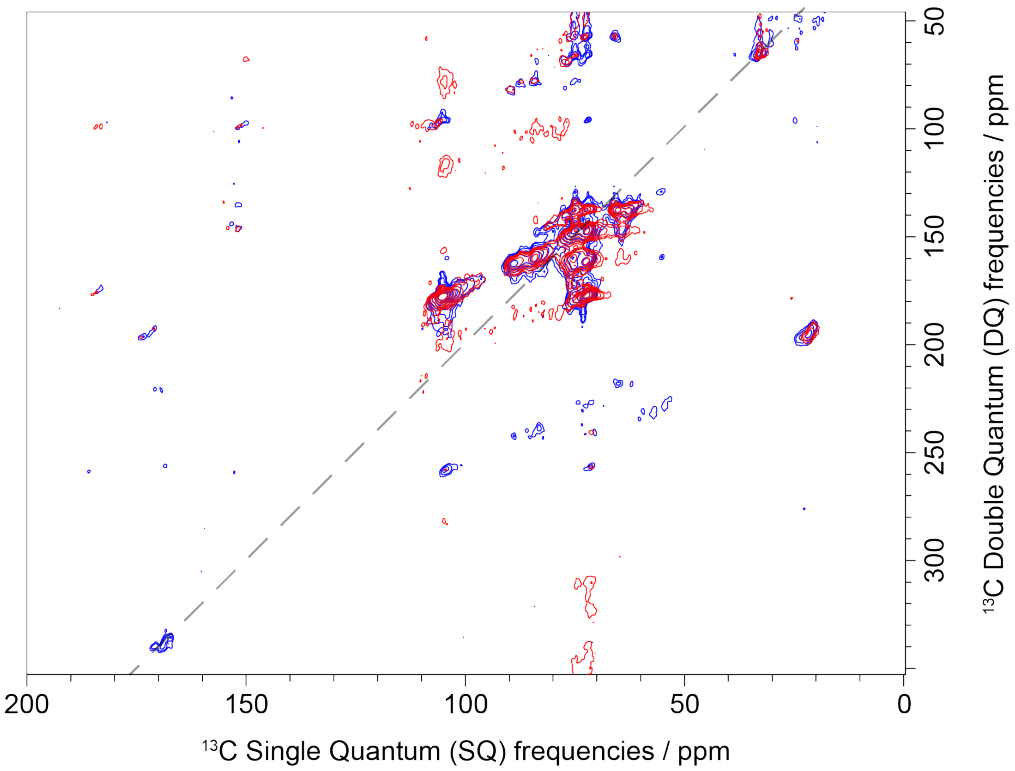


Figure 4.31: $^{13}$C-$^{13}$C 2D MAS refocused INADEQUATE spectra for t = 0 days (red) and t = 5 days (blue)

broadened which is partially due to the glucan in the fungus but there are additional small correlations that cannot be assigned to the known glucan shifts. There are many low intensity peaks surrounding the cellulose peaks which could be signs of the cellulose attempting to reorder to maintain its more favourable microfibril structure. The significant broadening of the cellulose peaks alongside a rise of small pockets of different cellulose environments suggest that some of the microfibril structure has begun to break down within the straw.

There have also been further changes in the xylan environment over the fermentation period. There is a shift downfield of a significant proportion of the xylan X4 to 83.6 ppm. There is also a small peak besides the original X4 that is shifted upfield slightly to 80.0 ppm. The shift of some of the X4 to 83.6 and 80.0 ppm shows there is a distribution of xylan environments

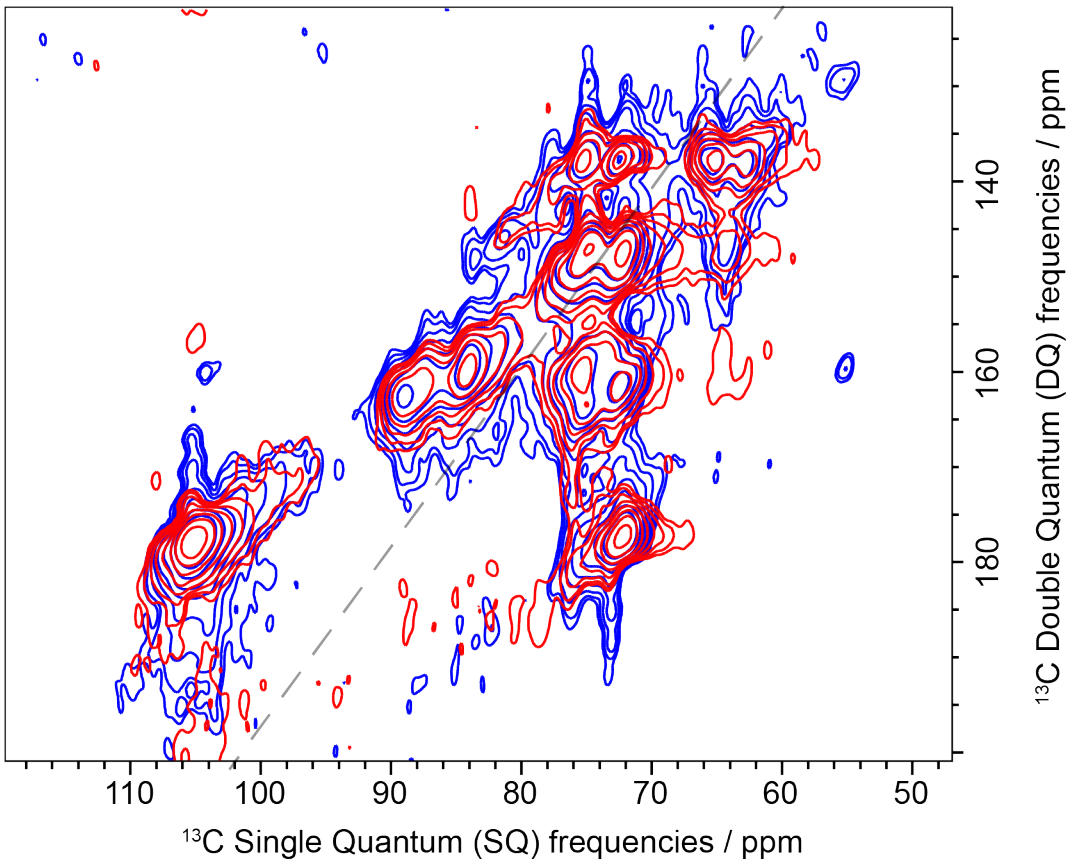


Figure 4.32: $^{13}$C-$^{13}$C 2D MAS refocused INADEQUATE spectra of the neutral carbohydrate region of wheat straw for t = 0 days (red) and t = 5 days (blue)

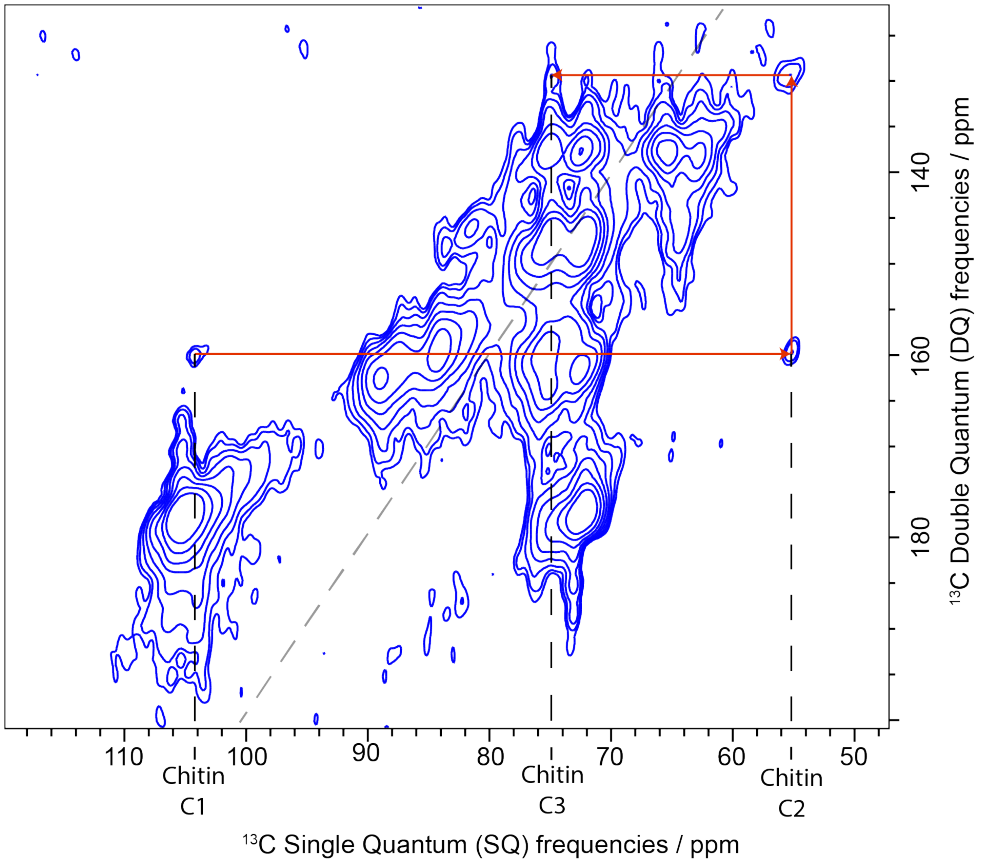


Figure 4.33: $^{13}$C-$^{13}$C 2D MAS refocused INADEQUATE spectra of the neutral carbohydrate region of wheat straw at t = 5 days fermentation time with the chitin carbon correlations identified that is present in the fungus.

which could be showing intermediate conformations of xylan moving from 2-fold xylan to 3-fold. The 3-fold xylan peak identified in the t = 1 day refocused INADEQUATE spectrum at 77.0 ppm can still be seen showing that there is still some xylan that is unbounded and in its 3 fold helical structure. There is also an increase in this peak's intensity indicating that the degradation of the cellulose forces more of the xylan away from the microfibril.

## 4.5 Conclusion

The Schmidt-Rohr method works well for very broad and overlapping humic acid spectra. The quantitative analysis of the two humic acid samples provided an insight into the composition of the humic acid and the reactivity from calculating the $sp^2/sp^3$ ratio. The first humic acid sample was found to have an $sp^2/sp^3$ ratio of 0.9 ± 0.09 and its composition is comparable with Leeheer's model.[75] Whilst the second humic acid sample had a $sp^2/sp^3$ ratio of 1.4 ± 0.05 and its composition comparable with the Dragonov model.[75,76]

The Schmidt-Rohr quantitative method was used to obtain quantitative analysis of the wheat straw and was compared with literature compositions. The Schmidt-Rohr method gave reasonable values for cellulose and hemicellulose content but the lignin content was smaller than expected at 7.8% compared to 8-20% in literature. This difference could be due to the larger error associated with the lignin peak or it could also be a young sample of wheat straw which does not have as high lignin content. The Schmidt-Rohr method was used to quantify changes in the wheat straw during the 5 day fermentation period with *Aspergilus niger*.

Solid state NMR is known to be extremely sensitive to changes in en-

vironments. Due to this, it makes it ideal to identify changes in the wheat straw structure over the fermentation time. The quantitative analysis, has shown key changes such as the production of enzymes and the decrease in the cellulose and hemicellulose by around 25% over the 5 day period. It has also shown that some of the xylan is degraded initially before the cellulose in the first 12 hours of fermentation after which the cellulose is degraded at a similar rate to xylan. The interior cellulose appears to degrade faster than the surface cellulose but this could be due to a reordering of the cellulose as it is exposed during the fermentation period. The 2D analysis showed that there was an increase in the presence of 3 fold helical xylan indicating that the hemicellulose is starting to disconnect from the cellulose microfibril. There are additional shifts for the 2 fold helical xylan suggesting that the way it lies across the cellulose microfibril is changing as the sample becomes more disordered. The cellulose environment changes significantly with many new shifts identified indicating a potential opening up of the microfibril structure. This suggests the pretreatment method increases the surface area of the cellulose leading to a more efficient hydrolysis stage.

# Chapter 5

# Using DNP-Enhanced Solid-State NMR to Study Natural Abundance $^{13}C$ Pharmaceuticals as Illustrated by Naloxone

Carbon-13 NMR has a very low sensitivity due to the $^{13}C$ nucleus' low natural abundance and small gyromagnetic ratio. One method to improve the sensitivity of $^{13}C$ NMR is to use samples that are partially or fully $^{13}C$ labelled. In the previous chapter, wheat straw samples were fully carbon-13 labelled which enabled 2D analysis of the straw. An alternative method for enhancing the sensitivity of NMR is Dynamic Nuclear Polarisation (DNP) NMR. This method was originally theorised in 1953 by Overhauser[53] and demonstrated by Slichter[54] but with technological advances in high powered microwave sources, such as gyrotrons, it has made high frequency DNP possible.

These samples were provided by GSK as an initial sample to investigate how DNP NMR could be used to better inform about the structure and analysis of pharmaceuticals. Naloxone is a well known pharmaceutical whose unit cell structure has been characterised by X-ray diffraction and NMR studies. In this chapter, DNP NMR is used enhance the sensitivity of $^{13}$C NMR to study two different forms of a pharmaceutical, naloxone. The aim is to show how DNP NMR can be used to cut down experimental time and provide enough sensitivity that natural abundance samples can be studied with 2D analysis in a relatively short period of time.

## 5.1 Naloxone

Opioids, such as morphine, are agonists which means that they bind to receptors that tigger a response in the brain.[7] Opioids bind to receptors and cause a reduction in perceived pain alongside euphoric effects, but when abused can cause death due to respiratory arrest.[125] Naloxone is an antagonist that is up to 1000 times more effective at binding to the opioid receptors than morphine and blocks the analgesic and euphoric effects of the opioid. Naloxone is used to counter the affects of drug overdoses and is now given to habitual drug users in case of an overdose situation.[7] Naloxone commonly comes in the form of naloxone hydrochloride dihydrate due to its increased solubility in water.[7,125]

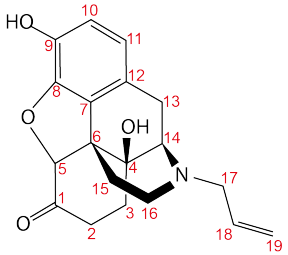


Figure 5.1: The structure of naloxone.

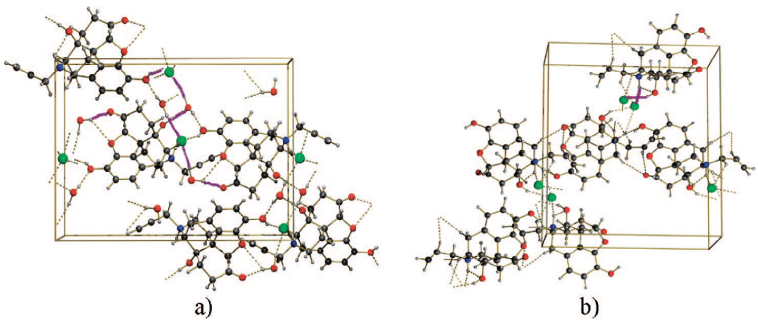


Figure 5.2: The 3D crystalline packing of a) naloxone hydrochloride dihydrate and b) anhydrous naloxone hydrochloride with the hydrogen bonding indicated by purple bonds taken from Guguta *et al.*.[7]

A number of crystal structures have been measured and refined for the hydrated and anhydrous forms of both alkaloid and salt of naloxone.[7,125–127] All hydrated and anhydrous forms of naloxone hydrochloride crystallise in an orthorhombic structure with space group $P2_12_12_1$. The hydrochloride salt forms a large three-dimensional hydrogen-bonded network, as shown in Figure 5.2, with one molecule per unit cell.[7] The carbonyl, OH groups and Cl anion readily form hydrogen bonds with the water molecules, while the nitrogen in naloxone only forms hydrogen bonds with the Cl anion in hydrochloride. In the absence of water the HCl maintains the three dimensional structure by forming many hydrogen bonds between the Cl anion and the -OH groups and nitrogen of the naloxone molecule.[127]

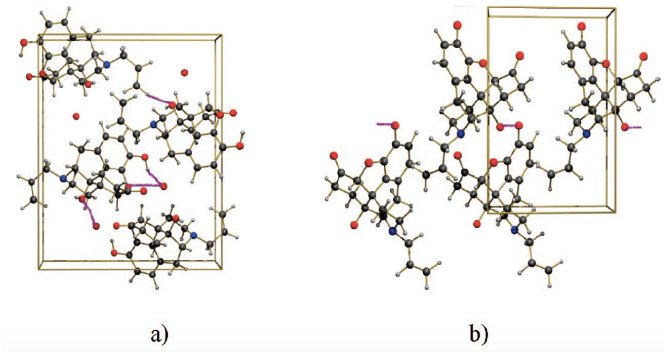


Figure 5.3: The 3D crystalline packing of a) naloxone monohydrate and b) naloxone anhydrate with the hydrogen bonding indicated by purple bonds taken from Guguta *et al.*.[7]

Hydrous and anhydrous forms of the alkaloid of naloxone have a different structure due to the lack of additional intermolecular bonding that the HCl provides. Naloxone monohydrate and anhydrate form one-dimensional chains of hydrogen-bonded molecules, as shown in Figure 5.3. In the monohydrate there is hydrogen bonding between -OH groups and water molecules. Whereas in the anhydrate, a chain of naloxone molecules is created by hydrogen bonds between -OH groups and the allyic end of the molecule. These free base naloxone forms have a monoclinic structure with a space group of $P2_1$.[7]

The $^{13}C$ NMR shifts are affected by the changes in the hydrogen-bonding structure. The carbonyl in particular is known to shift significantly from 208.1 ppm and 211.7 ppm when the anhydrous naloxone hydrochloride is hydrated to its dihydrate form. Other shifts also move in the NMR spectrum which can be partially assigned as they are the carbons which are more involved and affected by the changes in the hydrogen bonding. However, the majority of the carbons are unassignable through one-dimensional NMR spectroscopy.[7,127]

## 5.2 Experimental Details

Initial $^{13}C$ CP MAS experiments on the naloxone hydrochloride dihydrate were acquired using a 600 MHz Bruker Avance III solid-state NMR spectrometer operating at a magnetic field of 14.1 T corresponding to Larmor frequencies of 600.13 MHz and 150.90 MHz for $^{1}H$ and $^{13}C$ respectively. The naloxone hydrochloride dihydrate powder was packed in a 4 mm zirconia rotor. An $E^{free}$ 4 mm triple resonance MAS probe with a MAS rate of 8 kHz $\pm$ 5 Hz was used. Adamantane was used as an external reference with the downfield shift set to 38.4 ppm in compliance with the IUPAC

convention.[114] The CP experiment had an initial 90° $^{1}$H pulse of 3 $\mu$s and a CP contact time of 1 ms, 680 scans and a relaxation delay of 17 s. The acquisition time was 24 ms and SPINAL-64 proton decoupling was applied during acquisition.[6]

All DNP NMR experiments were acquired using the DNP MAS NMR facility at University of Nottingham. The DNP NMR setup comprises of a Bruker Avance III HD 600 MHz spectrometer operating at a magnetic field of 14.1 T corresponding to $^{1}$H and $^{13}$ Larmor frequencies of 600.13 MHz and 150.90 MHz respectively. The high power microwaves (10 W at the centre of the wave guide) for DNP excitation are provided by a gyrotron that operates continuously at a frequency of 395 GHz. The samples were wetted with the solvent-radical combination of 10 mM AMUPol in $CD_3OD/D_2O/H_2O$ (40:40:20) v/v/v with $CaCl_2$ to impregnate the radical into the sample. The wetted sample was then packed into a 3.2 mm sapphire rotor sealed with a Teflon plug and a zirconia cap. The 3.2 mm low-temperature triple-resonance MAS probe was used with MAS rates between 8 and 10 kHz and sample temperature of 100 ± 3 K. All CP DNP NMR experiments were acquired with 8 scans, relaxation delay of 5.2 s and an acquisition time of 24 ms. The 90° $^{1}$H pulse was set to 3 $\mu$s and the CP contact pulse of 2 ms with SPINAL-64 decoupling during acquisition. They both were Fourier transformed with 50 Hz of EM line broadening. The saturation recovery build up curves for both samples were acquired with a MAS rate of 8 kHz, a train of 100 saturation pulses separated by 1 ms, 8 scans and recovery delays ranging from 0.1 s to 125 s.

The DNP NMR $^{13}$C CP refocused INADEQUATE MAS spectra of naloxone alkaloid was acquired with a MAS rate of 8 kHz, 32 scans and a relaxation time of 5.2 s. It has been acquired with the States-TPPI method with 320 points in the $t_1$ dimension and 2048 points in the $t_2$ dimension.

The 1/4J delay was set to 3.12 ms and the spectral width was 424 ppm in F1 and 300 ppm in F2. The 90° $^{1}$H pulse was 3 $\mu$s and the $^{13}$C 90° and 180° pulse were 2.9 $\mu$s and 5.8 $\mu$s respectively. At the start of each scan there was a train of 100 saturation pulses separated by 1 ms. This was then Fourier transformed in both dimensions with zero filling points to 1024 points in F1 and 8192 points in F2. Line broadening was applied in both dimensions with QSINE of two in F1 and EM line broadening of 50 Hz in F2.

## 5.3 Results

### 5.3.1 DNP NMR of Naloxone Alkaloid

The naloxone molecule has several carbons in similar environments making it difficult to fully assign the shifts in the 1D CPMAS spectrum. It is possible to identify the carbonyl shift at 213.0 ppm as it is distinct from all other sites. The majority however, can only be assigned generally as aliphatic in the region 0 - 50 ppm and aromatic in the region 100-150 ppm. In order to make a complete assignment of the naloxone shifts a two-dimensional refocused INADEQUATE is necessary. In order to obtain a 2D spectrum in a reasonable time frame, sensitivity enhancement by DNP NMR is required as the cross peaks in the INADEQUATE spectrum are 1% of the intensity of a 1D CP MAS spectrum.[46]

The signal enhancement of DNP depends on both the sample preparation and relaxation times. The naloxone alkaloid sample was ground into a very fine powder to ensure a large surface area for the impregnation with the radical solution. The solvent-radical combination was used to wet the naloxone alkaloid sample to introduce the radical for DNP enhancement.

To get the best DNP enhancement, the DNP build up times have to

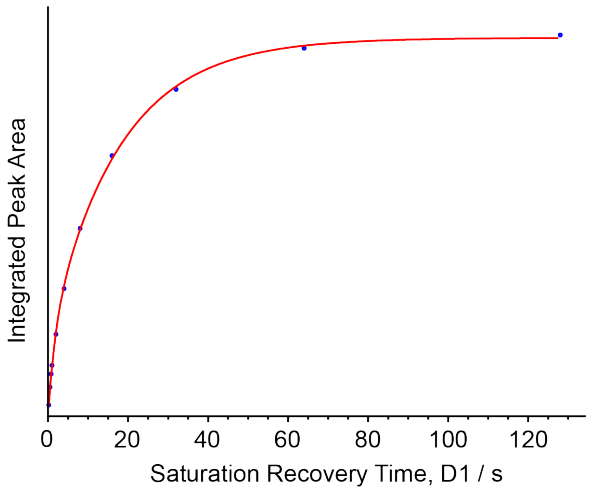


Figure 5.4: A saturation recovery DNP build up curve for the Naloxone alkaloid sample with 10 mM AMUPol in $CD_3OD/D_2O/H_2O$ (40:40:20) v/v/v and $CaCl_2$. Acquired with a MAS rate of 8 kHz, a train of 100 saturation pulses separated by 1 ms with a recovery delays ranging from 0.1 s to 125 s. The two component exponential fitting shown in red has $T_1$ of 2 s and 17 s with weightings of 0.6 and 0.4 respectively.

be measured for each sample as it depends on the relaxation of the nuclei in the sample, relaxation of the electron polarised in DNP and the spin diffusion of magnetisation to the protons in the sample. A saturation recovery experiment was used to measure the DNP build-up time and the peak areas were plotted as a function of recovery time, as shown in Figure 5.4. Fitting of the build up curve resulted in a best fit to a two component exponential with build up times of $T_1$ = 2 s and 17 s with a weighting of 0.6 and 0.4 respectively.

The relaxation time was set to 5.2 seconds as a compromise between the long and short $T_1$s. The enhancement was then measured by acquiring two 1D CP MAS spectra with the microwaves turned on and off. The comparison of these two spectra, shown in Figure 5.5, reveal that this DNP setup gave an enhancement of 18 for this sample. This enhancement, whilst significantly less than the DNP theoretical maximum still results in a 324 fold reduction in the experimental time. The microwave on/off spectrum

comparison was also compared for a relaxation delay of 34 s to incorporate the longer component seen in the build up curve, but the enhancement only increased to 19 so the shorter relaxation time was used in further experiments as it gives the largest signal in the same time.

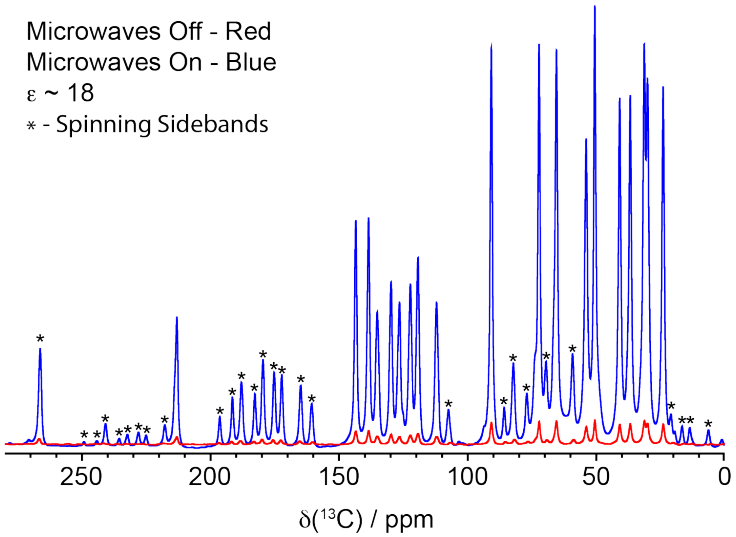


Figure 5.5: The $^{13}$C CP MAS spectra of naloxone alkaloid with 10 mM AMUPol in $CD_3OD/D_2O/H_2O$ (40:40:20) v/v/v and $CaCl_2$. The spectra were acquired with MAS rate of 8 kHz with microwaves on (blue) and off (red) showing a DNP enhancement of 18. Both spectra have been acquired with 8 scans, relaxation delay of 5.2 s and an acquisition time of 24 ms. The 90° $^1$H pulse was set to 3 $\mu$s and the CP contact pulse of 2 ms with SPINAL64 decoupling during acquisition. They both were Fourier transformed with 50 Hz of EM linebroadening.

Literature states that crystal structures of all the different forms of naloxone have only one molecule in its unit cell, meaning there should be 19 carbon shifts in the NMR spectrum. Within the CP $^{13}$C NMR spectrum there are many more peaks present as there are many spinning sidebands in the spectrum. These are distinguised by acquiring two spectra at different spinning speeds, 8 and 10 kHz, shown in Figure 5.6.

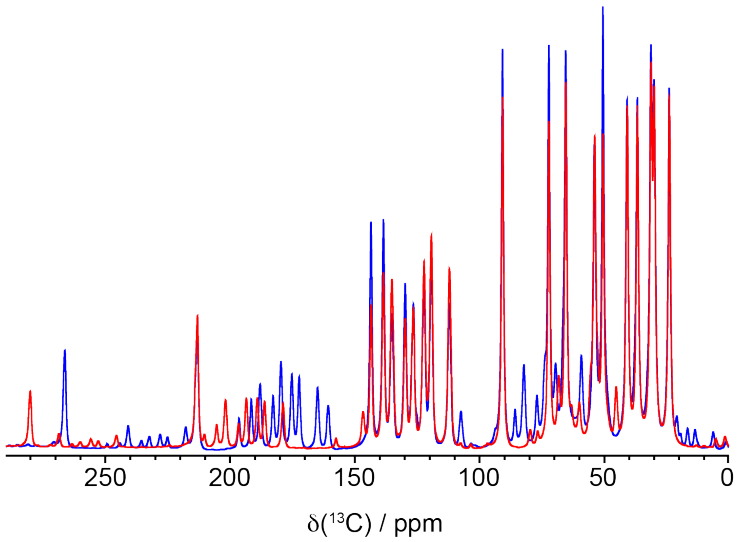


Figure 5.6: The $^{13}C$ CP MAS DNP spectra of naloxone alkaloid with 10 mM AMUPol in CD3OD/D2O/H2O (40:40:20) v/v/v and $CaCl_2$ acquired at 2 different MAS rates, 8 kHz (blue) and 10 kHz (red). Both spectra have been acquired with 8 scans, relaxation delay of 5.2 s and an acquisition time of 24 ms. The 90° $^1H$ pulse was set to 3 $\mu$s and the CP contact pulse of 2 ms with SPINAL64 decoupling during acquisition. They both were Fourier transformed with 50 Hz of EM line broadening.

The naloxone aromatic carbons at 100-150 ppm have significant spinning sidebands which is expected as their CSA tends to be larger than the aliphatic environments at 0-50 ppm. This is due to as the $sp^2$ carbon environment being less symmetric in its local electronic environment than $sp^3$ environments. The assignment of the spinning sidebands reveals the 19 carbon shifts the crystal structure predicts in the naloxone alkaloid spectrum, shown in Figure 5.7. The positions of the spinning sidebands in the spectrum with an MAS rate of 8 kHz, are more conveniently placed as they overlap less with the isotropic carbon shifts and so this MAS rate is used in further experiments on this sample.

The refocused INADEQUATE spectrum was chosen to help assign the naloxone carbon shifts as it provides correlations between directly bonded carbon nuclei. This experiment was chosen over short range dipolar coupling driven experiments as it leaves little ambiguity of which carbons are

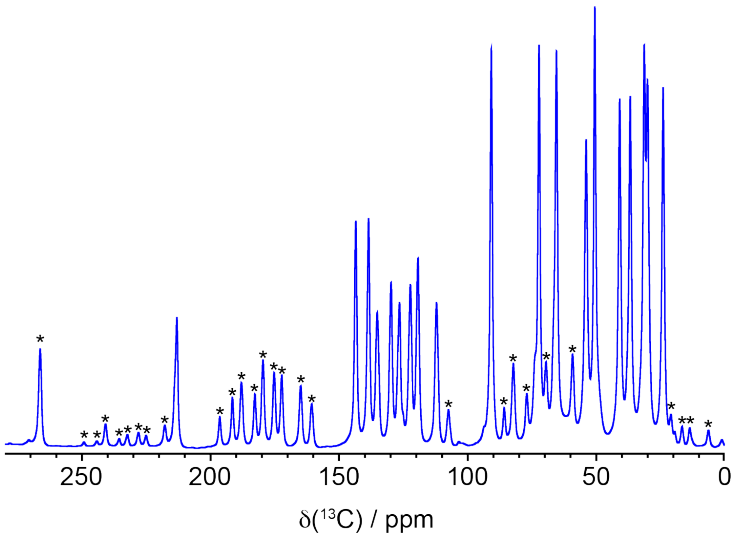


Figure 5.7: The $^{13}C$ CP MAS DNP spectrum of naloxone alkaloid with 10 mM AMUPol in $CD_3OD/D_2O/H_2O$ (40:40:20) v/v/v and $CaCl_2$. The spectrum was acquired with a MAS rate of 8 kHz, 8 scans, a relaxation delay of 5.2 s and an acquisition time of 24 ms. The 90° $^1H$ pulse was set to 3 $\mu$s and the CP contact pulse of 2 ms with SPINAL-64 decoupling during acquisition. They both were Fourier transformed with 50 Hz of EM line broadening and the spinning sidebands are labelled with an asterisk (*)

bonded together. The refocused INADEQUATE spectrum shown in Figure 5.8, was acquired over a 16 hour period with a recycle delay of 5.2 s, 32 scans per slice and 304 slices in the $t_1$ dimension. If the spectrum was acquired using standard NMR to have a similar signal to noise ratio and resolution the experiment would have taken more than 216 days. This 2D spectrum has high resolution in both dimensions and a good signal to noise ratio allowing each of the carbon shifts to be assigned easily.

Figure 5.9 illustrates the way the assignments were made by walking through the spectrum of carbon correlations of the benzene ring in naloxone. The complete assignment of each of the carbon shifts is given in Table 5.1 alongside the labelled carbon sites in the naloxone molecule.

| Carbon no. | Assigned chemical shift / ppm |
|---|---|
| 1 | 213 |
| 2 | 37 |
| 3 | 30/ 31 |
| 4 | 72 |
| 5 | 91 |
| 6 | 51 |
| 7 | 130 |
| 8 | 143 |
| 9 | 138 |
| 10 | 119 |
| 11 | 122 |
| 12 | 126 |
| 13 | 24 |
| 14 | 66 |
| 15 | 30/ 31 |
| 16 | 41 |
| 17 | 54 |
| 18 | 135 |
| 19 | 112 |

Table 5.1: The chemical shift assignments for all the naloxone alkaloid taken from the correlations made in the 2D refocused INADEQUATE spectrum.

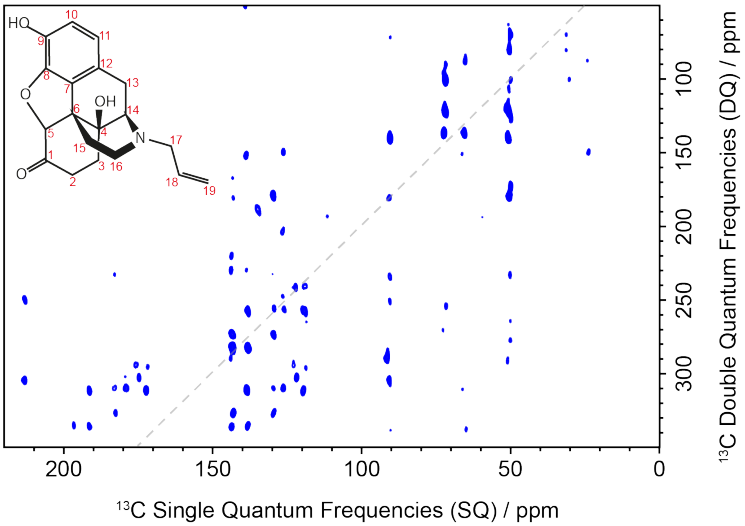


Figure 5.8: The DNP NMR $^{13}$C CP refocused INADEQUATE MAS spectra of naloxone alkaloid with 10 mM AMUPol in $CD_3OD/D_2O/H_2O$ (40:40:20) v/v/v and $CaCl_2$. This spectrum was acquired with a MAS rate of 8 kHz, 32 scans and a relaxation time of 5.2 s. It has been acquired with the States-TPPI method with 320 points in the $t_1$ dimension and 2048 points in the $t_2$ dimension. The 1/4J delay was set to 3.12 ms and the spectral width was 424 ppm in F1 and 300 ppm in F2. The 90° $^1$H pulse was 3 $\mu$s and the $^{13}$C 90° and 180° pulse were 2.9 $\mu$s and 5.8 $\mu$s respectively. At the start of each scan there was a train of 100 saturation pulses separated by 1 ms.

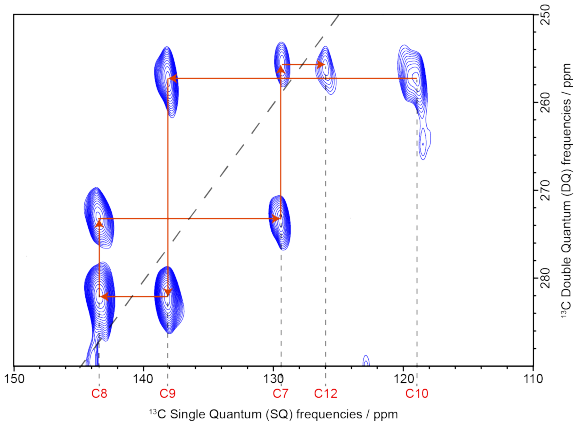


Figure 5.9: The $^{13}$C CP refocused INADEQUATE MAS spectrum of naloxone alkaloid with 10 mM AMUPol in $CD_3OD/D_2O/H_2O$ (40:40:20) v/v/v and $CaCl_2$ with parameters given in Figure 5.8 caption. This shows a small region of the spectrum to illustrate the assignment of the naloxone sites using the 2D spectrum. Here it follows around the carbon shifts of the majority of the benzene ring in the naloxone molecule.

### 5.3.2 DNP NMR of Naloxone hydrochloride

The Naloxone hydrochloride dihydrate sample was prepared in the same way as the naloxone alkaloid sample with the addition of 10 mM AMUPol in $CD_3OD/D_2O/H_2O$ (40:40:20) v/v/v with $CaCl_2$ to impregnate the sample with DNP radicals. The DNP build up time was found to be much longer than the previous sample with the main component having a relaxation time of 75 s with a weighting of 0.96.

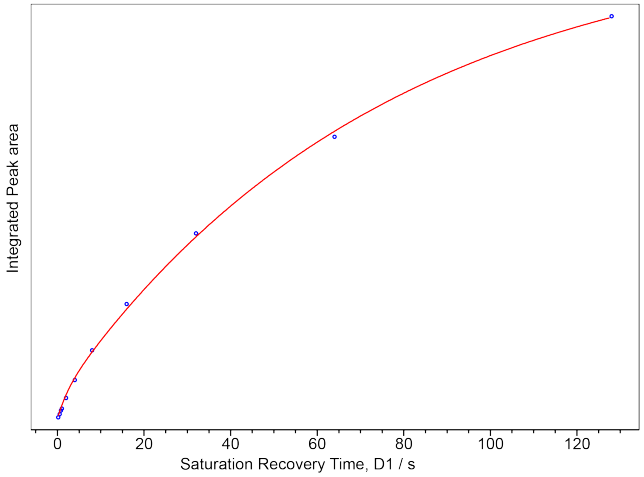


Figure 5.10: A saturation recovery DNP build up curve for the Naloxone hydrochloride dihydrate sample with 10 mM AMUPol in $CD_3OD/D_2O/H_2O$ (40:40:20) v/v/v and $CaCl_2$. Acquired with a MAS rate of 8 kHz, a train of 100 saturation pulses separated by 1 ms with a recovery delays ranging from 0.1 s to 125 s. The two component exponential fitting shown in red has $T_1$ of 75 s and 2 s with weightings of 0.96 and 0.04 respectively.

The long relaxation time means, that whilst it may provide a large signal enhancement, there was not enough time to acquire an INADEQUATE spectrum for this sample so was limited to a CP MAS spectrum.

The addition of radicals to the sample, and the low temperature can also cause a broadening of the spectrum which can impact the resolution of the spectrum significantly.[128] The 1D CP MAS spectrum acquired on a normal solid-state NMR machine of the same field without the DNP radicals has been compared with the 1D CP DNP MAS spectrum to gauge the extent

of the paramagnetic broadening of the line shapes and any movement in the $^{13}C$ isotropic shifts. Figure 5.11 shows that there are some small shifts in the $^{13}C$ peaks in the DNP spectrum particularly in the carbonyl shift which moves from 211.7 ppm in the CP MAS spectrum to 212.5 ppm in the DNP CP MAS spectrum. There is also a significant shift of the peak at 53.8 ppm to 52.9 ppm and a shift in 127.6 ppm peak to 128.1 ppm. The changes in the shift could arise from the change in the temperature as the DNP is acquired at very low temperatures (100 $\pm$ 3 K) where the water inside the sample is frozen fixing hydrogen bonds in place. The carbonyl site is known to be very involved in hydrogen bonding and the other two sites that show significant change are likely to be the carbon nuclei close to hydrogen bonding of the 2 -OH groups. The line shapes of the DNP spectrum are slightly broader than the standard NMR spectrum, however, this does not cause much of an issue as the majority of the isotropic shifts remain resolvable. The only shift the broadening really affects is the peak at 127.6 ppm which shifts to 128.1 ppm and the broadening means the 2 peaks at 128.1 and 128.8 ppm are not fully resolved. These changes in the spectrum are of significantly less importance than the large reduction in total experiment time.

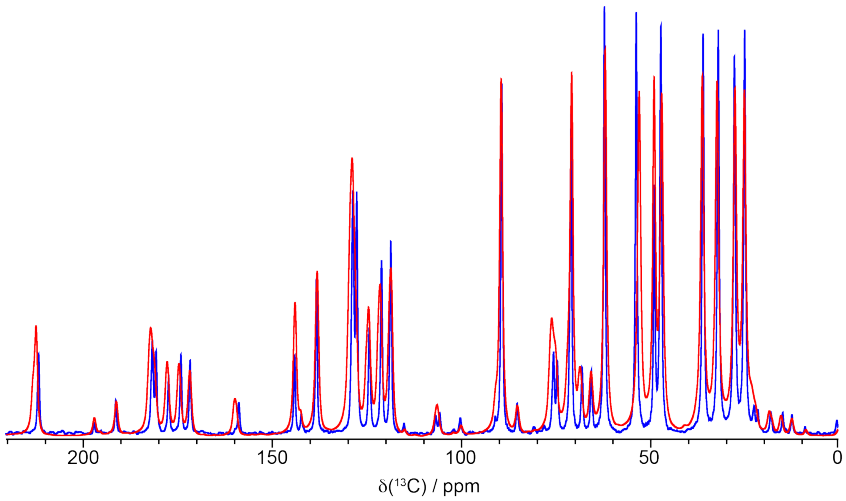


Figure 5.11: Comparison of the 1D CP MAS spectrum (blue) and the 1D CP DNP MAS spectrum (red) of naloxone hydrochloride dihydrate. The 1D CP NMR experiment had an initial 90° $^{1}$H pulse of 3 $\mu$s and a CP contact time of 1 ms, 680 scans and a relaxation delay of 17 s. Whilst the 1D CP DNP MAS had an initial 90° $^{1}$H pulse of 3 $\mu$s and a CP contact time of 2 ms, 8 scans and a relaxation delay of 125 s. They both were acquired with a MAS rate of 8 kHz and SPINAL-64 proton decoupling during acquisition. They have also both been Fourier transformed with 50 Hz of EM line broadening.

### 5.3.3 Naloxone Alkaloid and Naloxone Hydrochloride Comparison

Naloxone is usually administered in the form of Naloxone HCl dihydrate but addition of the salt and water changes the packing and the hydrogen bonding of the unit cell. These differences are illustrated in the comparison of the 1D carbon spectra of naloxone alkaloid and naloxone hydrochloride dihydrate seen in Figure 5.12. It shows that whilst there is only one naloxone molecule per unit cell for both samples, there is a shift of at least 1 ppm for the the majority of carbon peaks.

The only carbons that appear to show no shift between samples are the C(13) peak at 24 ppm and one of the peaks at the 30/31 ppm which was assigned to C(3) or C(15). These carbons are unaffected as they are not involved in hydrogen bonding and so the different packing does not affect

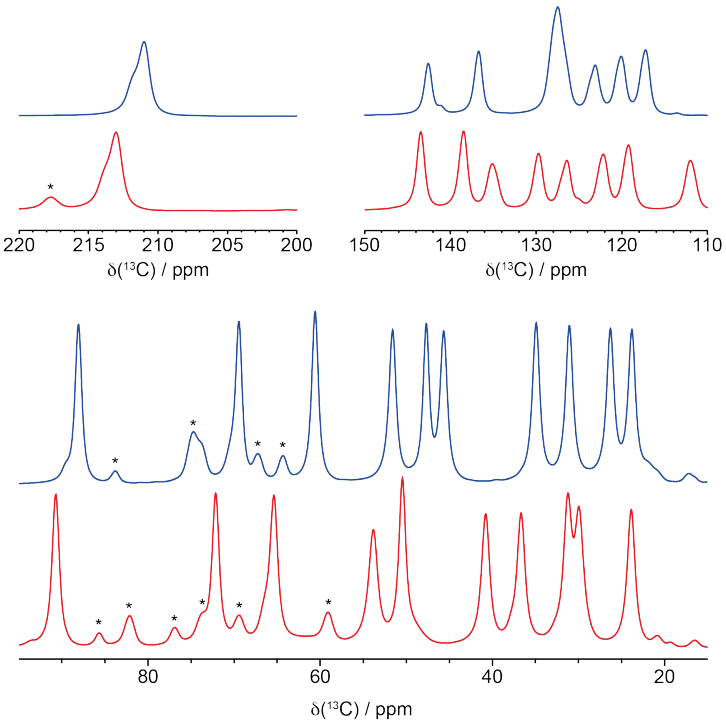


Figure 5.12: A comparison of the $^{13}$C CP MAS DNP spectra of Naloxone HCl (blue) and Naloxone alkaloid (red) with spinning sidebands indicated by an asterix (*).

these sites.

The new hydrogen bond between the nitrogen in naloxone and the chloride anion cause the largest shift of of the carbon peaks, C(14), C(16) and C(17) in the naloxone hydrochloride spectrum. Compared to the naloxone alkaloid spectrum, C(14) and C(17) shift upfield 4.8 and 2.4 ppm respectively whilst C(16) shifts downfield by 4.7 ppm. Further hydrogen bonding with water and the chloride anion to C(1)-O, C(4)-OH and C(9)-OH lead to a shift upfield of 2, 2.8 and 2 ppm respectively. The shift of the rest of the carbons can be attributed to the proximity to hydrogen bonds and the change of the dimensionality of the hydrogen bonding structure from 1D in the alkaloid sample to 3D in the HCl dihydrate sample.

## 5.4 Conclusions

Understanding differences in the unit cell structure of alkaloids, salts and hydrates of pharmaceuticals is extremely important as it can affect the mechanism and the efficiency it is absorbed into the body. Solid-state MAS NMR is a high resolution technique that allows these differences in the unit cell to be seen by shifts in the spectrum.

Naloxone alkaloid has been fully assigned with the aid of the DNP NMR technique that gave a sensitivity enhancement of 18 that meant a 2D refocused INADEQUATE spectrum could be acquired in 16 hours rather than 216 days using standard solid state NMR. However, the enhancement that DNP NMR provides varies significantly between samples and the DNP build up times can be very long, limiting the effective sensitivity gain. The naloxone hydrochloride dihydrate sample was shown to have a very long build up time of 75 s which meant that the time to acquire a suitable 2D refocused INADEQUATE was still unfeasible.

Comparisons of 1D CP MAS spectra of the naloxone hydrochloride dihydrate acquired using standard solid-state NMR and DNP NMR showed that there were some small shifts in some of the chemical environments. This is particularly observable the carbonyl site which is likely due to the low temperatures that DNP NMR is acquired at. There is also some broadening of all the naloxone HCl dihydrate line shapes due to the addition of the biradicals for DNP but this only causes an issue in the resolution of two isotropic shifts.

This chapter shows that the sensitivity that DNP can provide to samples could eliminate the need for labelling of pharmaceuticals for structural analysis in the solid-state. However, the approach it is not without drawbacks and care is required to ensure significant changes in the structure

are not occurring due to addition of radicals and lower temperatures. If DNP NMR becomes more reliable in achieving large enhancements then this technique could be invaluable to the pharmaceutical industry. Further experimentation could be used for rapid 2D solid-state NMR experiments or for looking into whole pill analysis of the drugs allowing for further understanding into how the excipients in pills, such as glucose, interact with the active drug.

# Chapter 6

# Measuring Fluorine Chemical Shift Anisotropies

Magic angle spinning and decoupling removes the majority of the anisotropic interactions leaving relatively narrow peaks at the isotropic shifts. These techniques mean that a lot of information is lost by removing the anisotropic component of the NMR signal. Therefore there has been a lot of work on the development of two-dimensional anisotropic-isotropic experiments in order to extract this information. R-sequences are a set symmetry-based pulse sequences that can be used in the first dimension of a two-dimensional experiment reintroduce or 'recouples' selected anisotropic interactions.[48] This enables the anisotropic interactions to be measured without compromising the resolution of the isotropic spectrum.

The sample for this project was provided by Johnson Matthey to use NMR to study the two polymorphs in the sample. The clearly resolvable fluorine environments provided an opportunity to improve the anisotropic-isotropic experiment used to measure $^{19}$F CSA previously developed by Martini *et al.*.[32] The two polymorphs are inseparable and have very similar physical properties and so it was also interesting to see if there were any

significant differences in the $^{19}$F CSA parameters.

R-sequences have previously been used to measure the proton and fluorine CSAs of crystalline polymers and simple hydrogen bonded molecules.[31,32] This chapter aims to improve a anisotropic-isotropic experiment that uses the R-sequence, $R12_5^4$ to recouple the CSA of $^{19}$F sites. The R-sequence used in the original experiment was significantly affected by the $B_1$ inhomogeneity of the probe. The basic R element of the R-sequence originally used simple 180° rf pulse. This 180° pulse has been replaced by the composite pulse $270°_{0}90°_{180}$ which could make the R-sequence more robust to $B_1$ inhomogeneity of the probe. The $^{19}$F CSA parameters of 2 polymorphs of casopitant have then been measured using the two-dimensional anisotropic-isotropic experiment that employs the R-sequence with the composite pulse as the basic R element.

## 6.1 Measuring Chemical Shift Parameters of Fluorine Sites

Fluorine plays an important role in the properties and structures of many pharmaceuticals, polymers and biological systems.[129,130] Fluorine-19 is a 100% abundant NMR active nuclei and has the third largest gyromagnetic ratio, $\gamma$, which makes it the most NMR sensitive stable nucleus after proton. The fluorine nucleus has a large chemical shift range and is very responsive to changes in its local environment making it ideal for studying different structures and polymorphs.[129] Unfortunately, the large gyromagnetic ratio and abundance of $^{19}$F means that the NMR spectrum is affected by large homonuclear and heteronuclear dipolar coupling.[32] This coupling leads to significant broadening of the NMR line shape and reduced resolution of the isotropic shifts. Magic angle spinning and decoupling removes the majority

of the broadening due to anisotropic interactions enabling the acquisition of high resolution fluorine spectra.[22]

The CSA provides a better characterisation of the local electronic environment of the the nucleus than just the isotropic shift alone as it allows the whole chemical shift tensor to be extracted. In a static spectrum of a powdered sample the CSA can be extracted from the powder pattern line shape.[129] When there are several environments, or the dipolar coupling involved is large, these powder pattern can overlap and broaden significantly causing issues in extracting the CSA parameters from the static spectrum. One commonly used method of extracting the CSA of fluorine sites is using spinning side band analysis that was developed by Herzfeld and Berger.[131] This method allows the CSA to be extracted by analysing the intensity of only a few spinning sidebands. This method is relatively easy to implement and is regularly used in samples where there is minimal overlap with the isotropic shifts of other sites. However, spectra with more than a couple of sites can quickly get very complicated and difficult to interpret.[32,129,132] Therefore, for more complicated samples the two-dimensional anisotropic-isotropic methods were developed.

The first methods for measuring CSA using a two-dimensional anisotropic-isotropic experiments were developed for measuring the CSA of protons as it was found that the anisotropic parameters relate to the strength and length of hydrogen bonds.[133,134] The measurement of the CSA parameters of proton and fluorine sites is not trivial as they are both affected by large homonuclear dipolar couplings and, for protons, this is compounded further as the CSA is very small.[31] In 2007, Brouwer and Ripmeester introduced using the R-sequence, $R18_2^5$, to recouple the $^1$H CSA into the indirect dimension using a moderate spinning speed of just over 16 kHz. This experiment was acquired with PMLG-5 homonuclear decoupling in the direct

dimension resulting in a high resolution proton spectrum. The $^1$H CSA line shapes were then fitted by SIMPSON to extract the CSA parameters.[134] A similar result was also achieved by Duma *et al.* where the $^1$H CSA was reintroduced using a rotary resonance experiment with a MAS rate of 10 kHz.[135] Hou *et al.* measured the the $^1$H CSA of different amide sites in proteins by transferring the recoupled magnetisation to neighbouring nitrogen sites.[136]

The commercial availability of smaller rotors that have to capability to spin at ultrafast MAS rates ($\nu_r > 50$ kHz ) meant that high resolution proton spectra were possible without the need for strong decoupling sequences. Miah *et al.* developed an R-sequence, $R16_3^2$ which recouples the $^1$H CSA of hydrogen bonded molecules. This R-sequence had a large scaling factor, $\kappa$ allowing the very small $^1$H CSAs to be measured. This sequence also had low enough rf amplitude requirements allowing the sample to be spun at MAS rates $\geq$ 60 kHz.[31] This experiment worked very well on a range of samples and was used to optimise the crystal structures of ascorbic acid by fitting a simulated 2D spectrum using the GIPAW approach in CASTEP.[133] The $R16_3^2$ sequence was quite sensitive to $B_1$ inhomogeneity which showed in the spectrum as a central spike in the CSA line shape. The intensity of the spike gives an indication of how affected the sequence is to $B_1$ inhomogeneity. It was found that a large central spike resulted in decreased intensity in the CSA line shape and a loss in sensitivity and resolution. Changes to the CSA line shape caused by the $B_1$ inhomogeneity means that it can affect the CSA parameters extracted during fitting of the spectrum.[51]

Pandey *et al.* later improved this ultrafast MAS regime for measuring $^1$H CSAs by developing R-sequences that had large scaling factors as well as being significantly more robust to $B_1$ inhomogeneity.[51] They showed

that replacing the 180° pulse, that makes up the train of pulses in the R-sequence, with a composite pulse can make the experiment more robust to rf inhomogeneity. Due to rf amplitude limitations of the probe, the R-sequence had to be carefully chosen such that even with the addition of the composite pulse the experiment could still be acquired with ultrafast MAS rates. They found that $R18_8^7$ and $R20_9^8$ sequences with the composite pulse $270°_0 90°_{180}$ as the basic R-element had a large scaling factor and were significantly more robust to $B_1$ inhomogeneity.[51] Meanwhile Zhang *et al.* recoupled the CSA in a 3D DQ/CSA/SQ experiment which allows the CSA of different proton sites to be resolved even when the resolution in the direct dimension is low.[137]

Martini *et al.* was the first to use an anisotropic-isotropic experiment to measure the CSA of fluorine sites.[32] The same R-sequence for measuring $^1$H CSAs could not be used for $^{19}$F CSAs as they are much larger than these of protons. The original $R16_3^2$ sequence does not have a big enough spectral width for the $^{19}$F CSA line shape to fit into without folding in the spectrum.[32] Therefore, a new sequence which had a longer cycle time for the R-sequence, and so a larger spectral width, was needed that could accommodate the $^{19}$F CSA line shapes. The $R12_5^4$ sequence could be made to have a sufficiently large spectral width and was found to recouple the $^{19}$F CSA reintroducing the homonuclear dipolar coupling. The two-dimensional anisotropic-isotropic experiment developed by Martini *et al.* used the $R12_5^4$ sequence where the data points were acquired every 4 basic R-elements in a method referred to as oversampling to increase the spectral width. This experiment was used to extract the $^{19}$F CSA for the different polymorphs of PTFE. However, the error associated with the asymmetry parameter was very large due to the effects of the $B_1$ inhomogeneity. In the same paper it was also shown that strong heteronuclear ($^1$H-$^{19}$F) dipolar coupling of

the order 5-25 kHz could be observed in the $^{1}$H CSA line shape and were extracted using the $R16_3^2$ sequence on a zinc phosphonate sample.

## 6.2 Casopitant

The pharmaceutical, casopitant, shown in Figure 6.1, is an antagonist that was developed by GSK to treat nausea and vomiting induced by chemotherapy. Analysis of its crystalline structure found that there were two stable polymorphs of casopitant referred to as form 1 and form 3.[138] Polymorphism can occur in any crystalline solid and one of the most widely accepted definitions comes from McCrone who states "A polymorph is a solid crystalline phase of a given compound resulting from the possibility of at least two different arrangements of the molecules of that compound in the solid state."[139] The study of polymorphism in different compounds is of particular interest in materials, polymers and pharmaceuticals as different polymorphs have different physical and chemical properties such as melting point and solubility.[138,139]

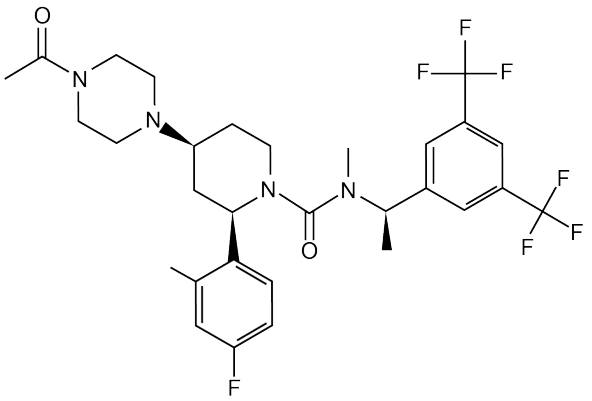


Figure 6.1: Casopitant molecule

Pharmaceutical polymorphs can have significantly different important properties such as the mechanism and the absorption rate of the drug into the body. A significant proportion of drug development is based on under-

standing the physical properties of the different polymorphs and generating methods to produce a stable polymorph with the ideal physical properties. In order to achieve this, the different polymorphs are usually separated to study their individual physical properties. However, this is not always achieved and despite many attempts form 1 and form 3 of casopitant were found to be inseparable. It was only possible to change the major polymorph present in the sample where one method contained 91% form 1 and another contained 78% of form 3.[138] These two extremes were used to study the physical and chemical properties of the polymorphs. It was found that form 1 and form 3 have very similar physical properties and are only quantitatively separated by transition XRPD and $^{19}$F solid-state NMR.[138]

In the $CF_3$ region of the 1D $^{19}$F MAS spectrum the polymorphs were not distinguishable with only a single peak at -63 ppm. In the aryl fluoride region, three different $^{19}$F shifts were resolvable. The $^{19}$F shift at -115.7 ppm was found to be a mixture of both polymorphs. The other two $^{19}$F shifts in the region at -116.6 ppm and -118.9 ppm were assigned to form 3 and form 1 respectively.[138] The two different shifts for each polymorph indicates that there are two molecules in a unit cell.

## 6.3 Experimental

All experiments in this chapter were acquired on a casopitant sample with 90% form 1 and 10% form 3 polymorph mixture using a 600 MHz Bruker Avance III solid-state NMR spectrometer operating at a magnetic field of 14.1 T corresponding to Larmor frequency of 564.7 MHz for $^{19}$F with a 1.3 mm double resonance probe and MAS rates between 50-60 kHz.

The 1D $^{19}$F MAS spectrum was acquired using an experiment with a

very short echo time of 40 $\mu$s to remove any broad background signals from the probe. The spectrum was run with an MAS rate of 60 kHz, 16 scans and a relaxation delay of 17 s. The spectral width was set to 200 kHz and the $\pi/2$ pulse length was 2.3 $\mu$s.

Simulations of the resulting CSA patterns for a $R12_5^4$ sequence with and without the $270_0°90_{180}°$composite pulse were run with a defined CSA of 10 ppm. The distance between the 2 peaks in the CSA lineshape is related to the size of the CSA and so was used to calculated the scale factor for each of the R-sequences used in this work.

The R-sequence used was the same as that presented by Martini *et al.*, the $R12_5^4$, which when the single R-element is a simple 180° pulse has a scaling factor, $\kappa$, of 0.155. When the R-element is replaced with the composite pulse $270_0°90_{180}°$the $\kappa$ increases to 0.163. The standard R-sequence was acquired with an MAS rate of 60.038 kHz which requires a $^{19}$F rf frequency of 72.0 kHz. The composite pulse version of the R-sequence was acquired with a MAS rate of 50.813 kHz which gave a rf amplitude requirement of 121.9 kHz. Both experiments were acquired with 32 $t_1$ increments with data points sampled every 4 R-elements leading to a dwell time in the $t_1$ dimension of 27.8 $\mu$s. In $t_2$ 4096 complex points were acquired with a spectral width of 200 kHz. At the start of each scan a train of 200 saturation $\pi/2$ pulses separated by 10 ms and a recovery delay of 5 s was used.

Numerical simulations of the $^{19}$F CSA line shapes were produced using SIMPSON using 615 ($\alpha$, $\beta$) orientations and 5 uniformiy distributed $\gamma$ angles to achieve powder averaging. In all simulations the $B_1$ (rf) inhomogeneity was incorporated by taking a set of 17 points along the experimentally measured $B_1$ inhomogeneity distribution of the probe using the experiment by Odedra and Wimperis shown in Figure 6.2.[8]

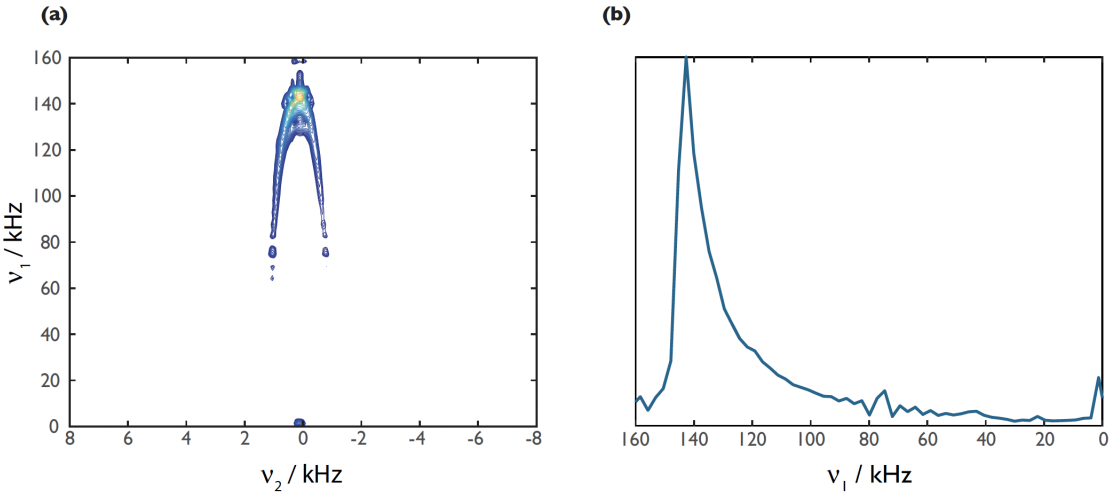


Figure 6.2: The $B_1$ inhomogeneity profile of the 1.3 mm probe given using the experiment of Odedra and Wimperis with a nominal field of 135 kHz.[8] a) Illustrates the nominal rf field felt across the rotor and shows that the top and bottom of the rotor have more rf inhomogeneity associated with it than the centre of the rotor. b) A projection parallel to $\nu_1$ of the $B_1$ inhomogeneity over the whole rotor. This shows that the majority of the sample feels the nominal rf field but there is a portion of the sample that feels a range rf fields around the nominal field.

The CSA parameters were extracted from the experimental slices using an array of simulated line shapes with varying chemical shift anisotropy, $\zeta$, and asymmetry, $\eta$, values. A $\chi^2$ parameter for each simulated line shape was then plotted as a function of the $\zeta$ and $\eta$ producing an error surface that shows the optimal values for the $^{19}$F CSA and asymmetry. Assuming this error surface is properly normalised using the standard equation given by:

$$\chi^2 = \frac{1}{\nu}\sum_{i=1}^{n}\frac{(S_i - E_i)^2}{e^2} \tag{6.1}$$

Where $S_i$ and $E_i$ are the simulated and experimental spectrum, e is the standard deviation due to the noise in the experiment spectrum, n is the total number of date points and $\nu$ is the number of degree of freedom = n-p-1 with p being the number of fitting parameters. The 95% confidence limits for this error surface is then given by two times the $\chi^2_{min}$.

## 6.4 Measuring the $^{19}$F CSA of Casopitant polymorphs

### 6.4.1 Identifying the polymorphs of Casopitant

An initial 1D $^{19}$F MAS spectrum of casopitant in Figure 6.3 shows that 3 peaks are resolved in the aryl fluoride region (-110 to -125 ppm) with form 1 and form 3 at -118.2 ppm and -116.6 ppm respectively and an overlapping site for both polymorphs at -115.7 ppm as seen in literature.[138] There are two isotropic shifts for both polymorphs indicating that both have two molecules in their unit cells. The ratio of the aryl peaks indicates that the casopitant sample is made up of around 90% form 1. The $CF_3$ region also has three sites. However, these have not been assigned previously in literature as they were not resolvable. Assuming the ratio of the intensities is maintained the small peak at -61.8 ppm is associated with form 3 which only makes up 10% of the sample.

### 6.4.2 Comparison of the R12$_5^4$ sequences with and without composite pulse R-element

The two-dimensional anisotropic-isotropic experiment used the R12$_5^4$ sequence to recouple the $^{19}$F CSA in the indirect dimension. In this section the original R12$_5^4$ sequence that was developed by Martini *et al.* has been compared with the R12$_5^4$ sequence which has had the basic R-element replaced with the composite pulse, $270°_0 90°_{180}$. Their spectra, CSA lineshapes and error surfaces are compared to ascertain if the composite pulse makes the experiment more robust to $B_1$ inhomogeneity. The original experiment was known to be affected by the $B_1$ inhomogeneity as it broadened the $^{19}$F CSA line shapes of PTFE such that some asymmetry parameters

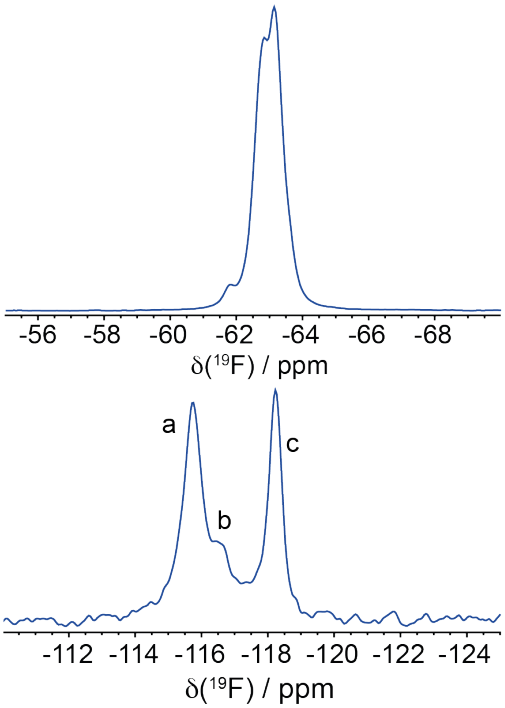


Figure 6.3: $^{19}$F MAS spectra of casopitant acquired at a MAS rate of 60 kHz. Top: the $CF_3$ region bottom: the aryl fluoride region. The aryl peaks have been assigned to the two different polymorphs of casopitant where peak "a" at -115.7 ppm arises from both polymorphs. Peak "b" at -116.6 ppm and peak "c" at -118.2 ppm are assigned to form 3 and form 1, respectively.

has significantly large errors.

Figure 6.4 compares the $^{19}$F isotropic-anisotropic spectra recorded with both experiments on the same sample of casopitant and have been processed in the same way. It shows that there is a large central spike in CSA line shapes of the original $R12^4_5$ sequence experiment whilst the central spike is reduced in the $R12^4_5$ sequence with the composite pulse spectrum. The presence and intensity of the central spike was used by Pandey *et al.* as a measure of the effect of the $B_1$ inhomogeneity.[51] The reduced central spike suggest that the composite pulse experiment is more robust to the $B_1$ inhomogeneity of the probe. The $CF_3$ region in the original experiment was significantly more affected by the central spike as it has a much smaller CSA and so causes significant reduction in resolution of the CSA line shape.

The composite pulse experiment has a slightly larger scaling factor and a reduced central spike means that the CSA line shapes of the $CF_3$ region can be resolved.

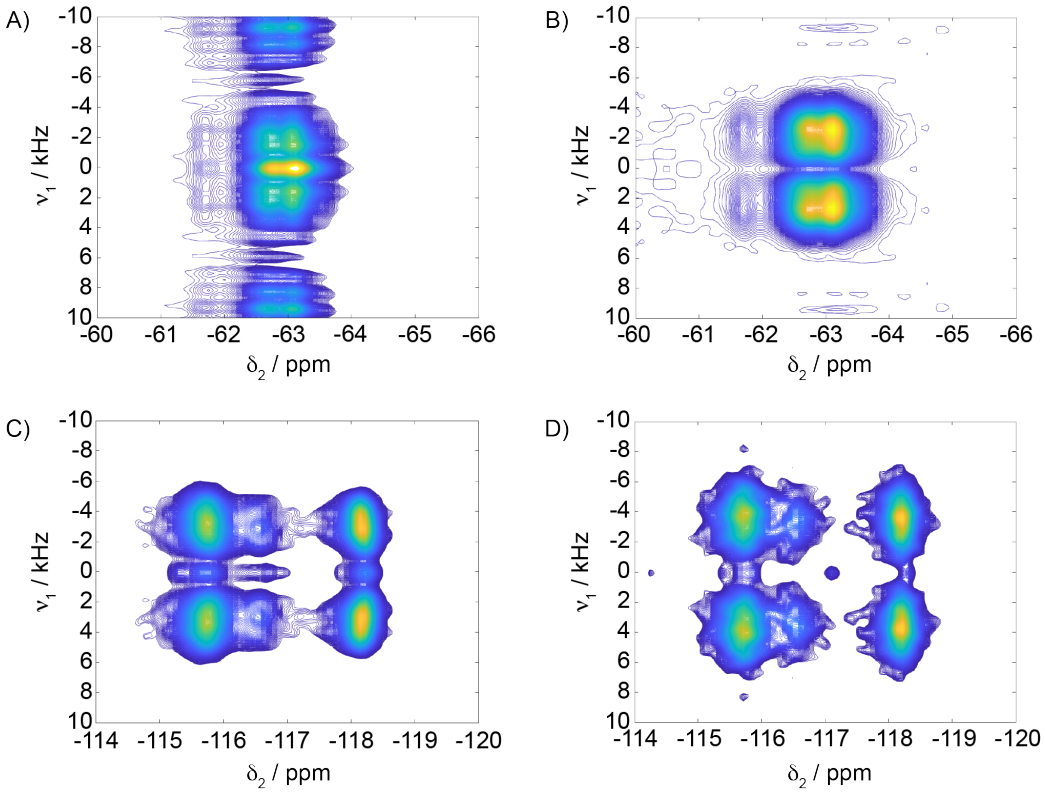


Figure 6.4: A) and B) shows the $CF_3$ region of the $^{19}F$ anisotropic-isotropic correlation spectra of casopitant whilst C) and D) shows aryl fluoride region of the $^{19}F$ anisotropic-isotropic correlation spectra of casopitant. A) and C) were acquired using the $R12_5^4$ sequence with the basic R-element a single 180° pulse, an MAS rate of 60.038 kHz and corresponding rf amplitude 72.04 kHz. B) and D) were acquired using the $R12_5^4$ sequence with the basic R-element a $270°_0 90°_{180}$ composite pulse, an MAS rate of 50.813 kHz and corresponding rf amplitude 121.9 kHz.

The comparison of the $^{19}$F CSA line shapes in Figure 6.5 shows more clearly the effect of $B_1$ inhomogeneity on the resolution of the $CF_3$ sites in the original R-sequence experiment which become unresolvable. The $^{19}$F CSA line shapes of the three aryl fluoride sites which have a significantly larger CSA and so are resolved in both experiments. The composite pulse R-sequence $^{19}$F CSA line shapes of the aryl fluoride sites have detail in their line shapes that can be fitted to extract the asymmetry parameter more accurately than the original R-sequence experiment. The $^{19}$F CSA line shapes of the $CF_3$ environments, whilst now resolvable in the composite pulse experiment, still shows some evidence of being affected by the $B_1$ inhomogeneity. For instance it has a less structured CSA line shape than the aryl fluoride sites.

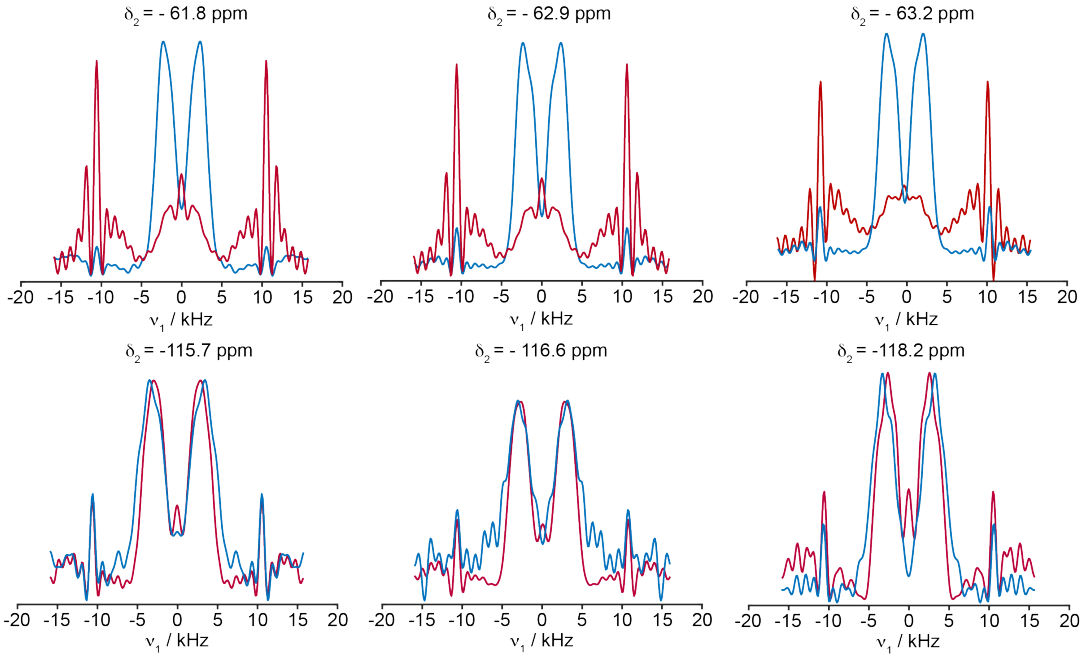


Figure 6.5: A comparison of the cross-sections taken at $\delta_2$ frequencies showing the recoupled CSA line shapes of the 6 different isotropic $^{19}$F shifts for the $R12_5^4$ sequence with the R-element a 180° pulse shown in red and the $R12_5^4$ sequence with the R-element a $270°_0 90°_{180}$ composite pulse shown in blue.

The resolvable CSA line shapes have been fitted using simulated line

| Assignment | Isotropic Shift $\delta$ ppm | CSA $\zeta$ ppm | Asymmetry $\eta$ |
|---|---|---|---|
| Original R-sequence | | | |
| Form 1+3 | -115.7 | 107± 10 | 0.8 ±0.4 |
| Form 3 | -116.6 | 96± 15 | 0.8±0.4 |
| Form 1 | -118.2 | 107± 12 | 0.8± 0.35 |
| Composite pulse R-sequence | | | |
| Form 1+3 | -115.7 | 91.5±10 | 0.85±0.15 |
| Form 3 | -116.6 | 81±10 | 0.80± 0.1 |
| Form 1 | -118.2 | 83±8 | 0.95±0.2 |

Table 6.1: A table comparing the $^{19}$F chemical shift parameters of casopitant with 95% confidence limits measured using the original R-sequence by Martini *et al.* and the R -sequences with the $270^\circ_0 90^\circ_{180}$ composite pulse as the basic R-element.

shapes generated using SIMPSON by the method described in the experimental section. The extracted chemical shift parameters are summarised in Table 6.4.2 and the corresponding error surfaces are shown in Figure 6.6. They show that the asymmetry is significantly better defined in the composite pulse experiment compared to the original. The extracted CSA of the original experiment is larger than the CSA extracted using the composite pulse experiment. The original experiment CSA values are also much larger than literature values of similar aryl fluoride sites whose values tend to be between 55-75 ppm.[129] Therefore, the composite pulse experiment gives CSA values much closer to literature values as well as significantly reduced error in the asymmetry.

The comparison of the R-sequence with and without the composite pulse $270^\circ_0 90^\circ_{180}$ has conclusively shown that the addition of the composite pulse does make the R-sequence significantly more robust to $B_1$ inhomogeneity. The composite pulse experiment gives more well defined line shapes enabling extraction of the CSA parameters with significantly reduced errors. The reduction of the central spike means that smaller $^{19}$F CSA line shapes can be resolved using this experiment too.

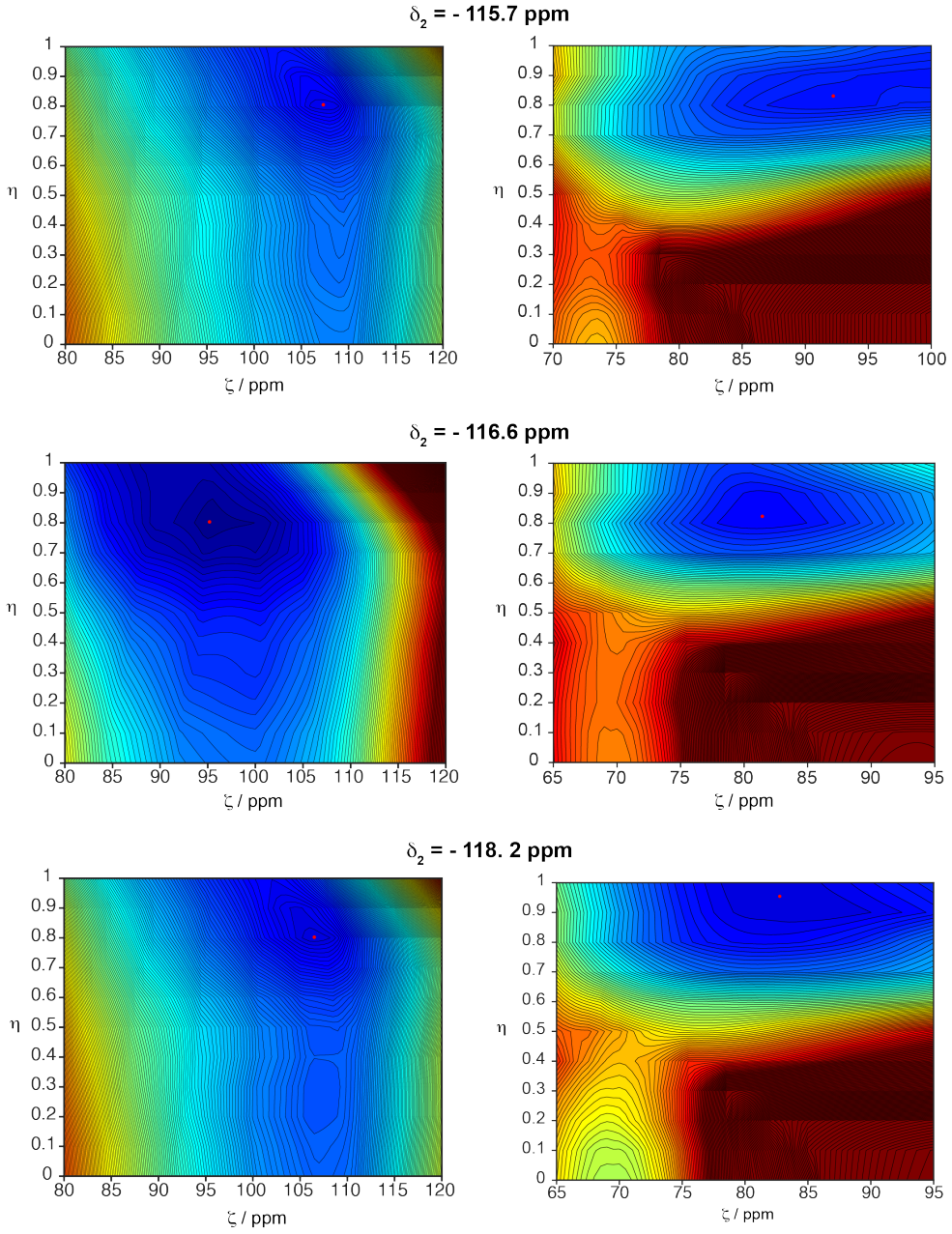


Figure 6.6: A comparison of the 3 error surfaces for the $^{19}$F shift parameters for the aryl fluoride sites of casopitant produced during the fitting process with the best fit parameters indicated by the red dot. On the left are error surfaces relating to the $R12^4_5$ sequence with the r-element a single 180°pulse whilst on the right are those relating to the $R12^4_5$ sequence with the r-element a $270°_0 90°_{180}$ composite pulse.

### 6.4.3 Measuring the $^{19}$F CSA of Different Polymorphs of Casopitant

Due to the additional resolution of the 1D $^{19}$F MAS spectrum in comparison to literature there are six different fluorine sites of casopitant whose CSA

can be measured. However only four of these sites relate to the two separate polymorphs whilst the other two peaks are a mixture. The composite pulse $R12_5^4$ experiment, shown to be more robust to $B_1$ inhomogeneity, has now been used to to extract the CSA parameters for the $CF_3$ isotropic shifts as well as the aryl fluorine environments.

Initial simulations of the $R12_5^4$ sequence with the $270°_0 90°_{180}$ composite pulse have been run using SIMPSON NMR simulation program to observe the line shape changes with varied $\zeta$ and $\eta$ shown in Figure 6.7. The simulations show that the CSA line shape is sensitive to changes in $\zeta$, since the peak moves when $\zeta$ is varied, as shown in Figure 6.7a. The line shape is also sensitive to changes in the asymmetry, $\eta$ which causes changes in the overall line shape and not a shift in the peak positions.

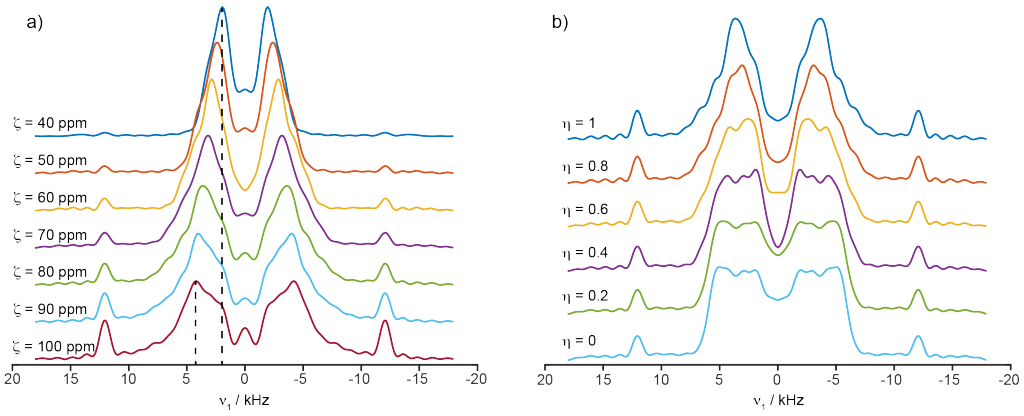


Figure 6.7: SIMPSON simulations of the $^{19}F$ CSA line shapes recoupled using the $R12_5^4$ sequence with the basic element a $270°_0 90°_{180}$ composite pulse and including the effects of $B_1$ inhomogeneity. In a) $\eta = 1$ whilst $\zeta$ varies showing that the line shape is sensitive to differences in the $^{19}F$ CSA while in b) $\zeta = 80$ ppm and $\eta$ is varied showing that the line shape is also sensitive to the asymmetry.

The CSA line shapes of the composite pulse experiment have been fitted by running several simulations with varying $\zeta$ and $\eta$ parameters to produce a $\chi^2$ error surface. The isotropic-anisotropic spectra of the $CF_3$ and aryl fluoride regions with the corresponding $\nu_1$ slices through the isotropic shifts are given in Figures 6.8 and 6.9 respectively. The red dashed lines show the

CSA line shape simulated by SIMPSON with the best fitting parameters extracted from the error surfaces for each recoupled CSA line shape. The extracted CSA parameters and 95% confidence intervals are summarised in Table 6.4.3.

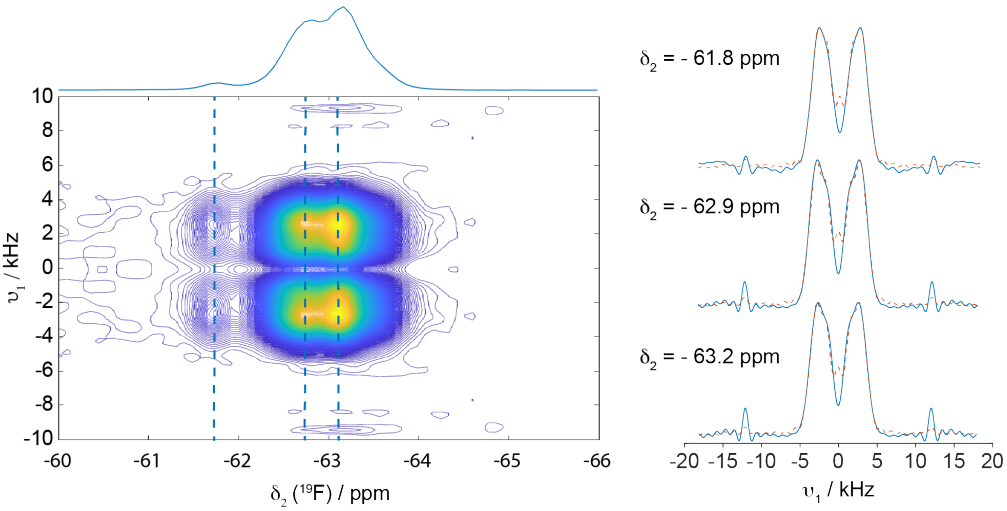


Figure 6.8: Left shows the $CF_3$ region of the $^{19}F$ anisotropic-isotropic correlation spectra of casopitant which was acquired using the $R12_5^4$ sequence with the basic R-element a $270°_0 90°_{180}$ composite pulse. A MAS rate of 50.813 kHz was used and rf amplitude 121.9 kHz corresponding to a $\pi$ pulse duration of 4.1 $\mu$s. Right shows a comparison of the cross-sections taken at $\delta_2$ frequencies showing the recoupled experimental CSA line shapes of the three different isotropic $^{19}F$ shifts in the $CF_3$ region of casopitant (blue) and the best fit parameters SIMPSON simulations of the CSA line shapes (red dashed).

The physical and chemical parameters of the two polymorphs of casopitant are known to be very similar and so it is interesting there is some slight differences between their CSA parameters. The CSA parameters can give an insight in to the differences between the local environment and packing between the two crystalline polymorphs. In the aryl fluoride region the overlapping site at $\delta_2$= -115.7 ppm has the largest $\zeta$ of 91.5 ppm suggesting this fluorine site is more affected by its environment than the second aryl site for the 2 polymorphs. A similar increase of 10 ppm in the CSA was found when the aromatic fluorine site was in close proximity to

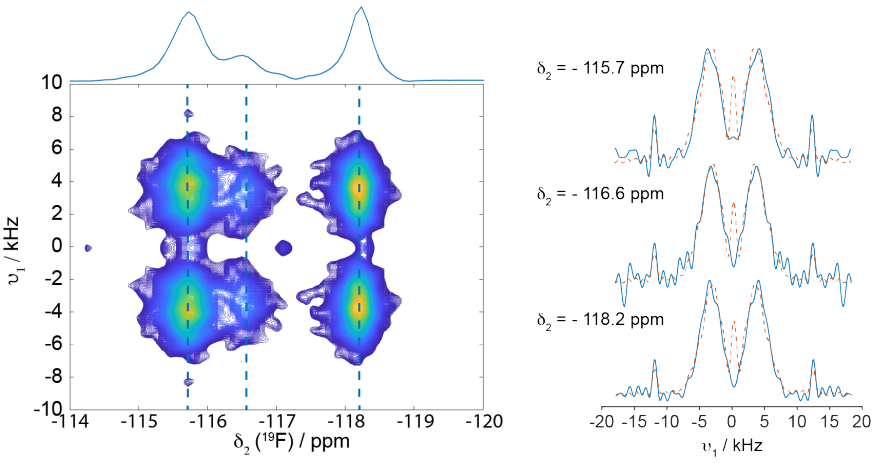


Figure 6.9: Left shows the aryl fluoride region of the $^{19}$F anisotropic-isotropic correlation spectra of casopitant which was acquired using the $R12_5^4$ sequence with the basic R-element a $270°_0 90°_{180}$ composite pulse. A MAS rate of 50.813 kHz was used and rf amplitude 121.9 kHz corresponding to a $\pi$ pulse duration of 4.1 $\mu$s. Right shows a comparison of the cross-sections taken at $\delta_2$ frequencies showing the recoupled experimental CSA line shapes of the three different isotropic $^{19}$F shifts in the aryl fluoride region of casopitant (blue) and the best fit parameters SIMPSON simulations of the CSA line shapes (red dashed).

OH group when studying fluorine substituted amino acids.[129] This increase also accompanied a change in the sign of $\zeta$ which can not be found using this experiment due to the symmetric nature of the recoupled CSA line shape. Even though there is a significant difference in $\zeta$ for the fluorine shift at -155.7 ppm there does not appear to be much difference in the axial asymmetry of the site compared to the other aryl sites but this site is a mixture of the two polymorphs. The aryl fluoride site for Form 1 at -118.2 ppm has a slightly larger $\zeta$ than the aryl site for form 3 at -116.6 ppm and is the most axially asymmetric site of all the fluorine sites in the sample.

The $CF_3$ sites have a much smaller CSA than the aryl fluoride sites which is expected as the $CF_3$ environment is closer to spherical symmetry than the planar C-F bond. The $CF_3$ sites are also significantly more symmetric than the aryl fluorides with $\eta$ of between 0.55-0.6 rather than 0.8-0.95. This increase in axial symmetry is again attributed to the mo-

| Assignment | Isotropic Shift $\delta$ ppm | CSA $\zeta$ ppm | Asymmetry $\eta$ |
|---|---|---|---|
| Form 3 | -61.8 | 48±7 | 0.60±0.4 |
| Form 1 | -62.9 | 49.5±5 | 0.55±0.4 |
| Form 1+3 | -63.2 | 48.5±5 | 0.60± 0.35 |
| Form 1+3 | -115.7 | 91.5±10 | 0.85±0.15 |
| From 3 | -116.6 | 81±10 | 0.80± 0.1 |
| Form 1 | -118.2 | 83±8 | 0.95±0.2 |

Table 6.2: A table of the $^{19}$F chemical shift parameters in casopitant with 95% confidence limits measured in this work from the fitting of simulated spectra from the $R12_5^4$ sequence with the R element a $270°_0 90°_{180}$ composite pulse.

tional averaging of the rotation of the $CF_3$ groups. The measured $^{19}$F CSA values for the $CF_3$ region are close to the literature values which fall between 32-55 ppm. However, in literature the $CF_3$ sites have been assumed to be completely axially symmetric (i.e. $\eta$=0) due to the rotation of the $CF_3$ bonds. In the literature, the powder pattern line shapes were significantly affected by homonuclear dipolar coupling.[132] In this experiment the homonuclear dipolar coupling is not recoupled like the CSA hence the line shapes here are not affected by the same issues as the study of the powder patterns.

There is very little variation in the CSA parameters between the three $CF_3$ sites but the larger error associated with $\eta$ arises from very little defining features on the CSA line shape. This is due to the fact that CSA is so much smaller than the aryl fluoride sites causing the CSA line shape to be more affected by the $B_1$ inhomogeneity in the same way the Martini *et al.* experienced with the original R-sequence.[32]

The error surfaces of all six fluorine sites is given in Figure 6.10 where it shows that the $\eta$ of the aryl groups is much more defined than the $\eta$ for the $CF_3$ sites. This larger error of the $\eta$ could be attributed to $B_1$ inhomogeneity effects.

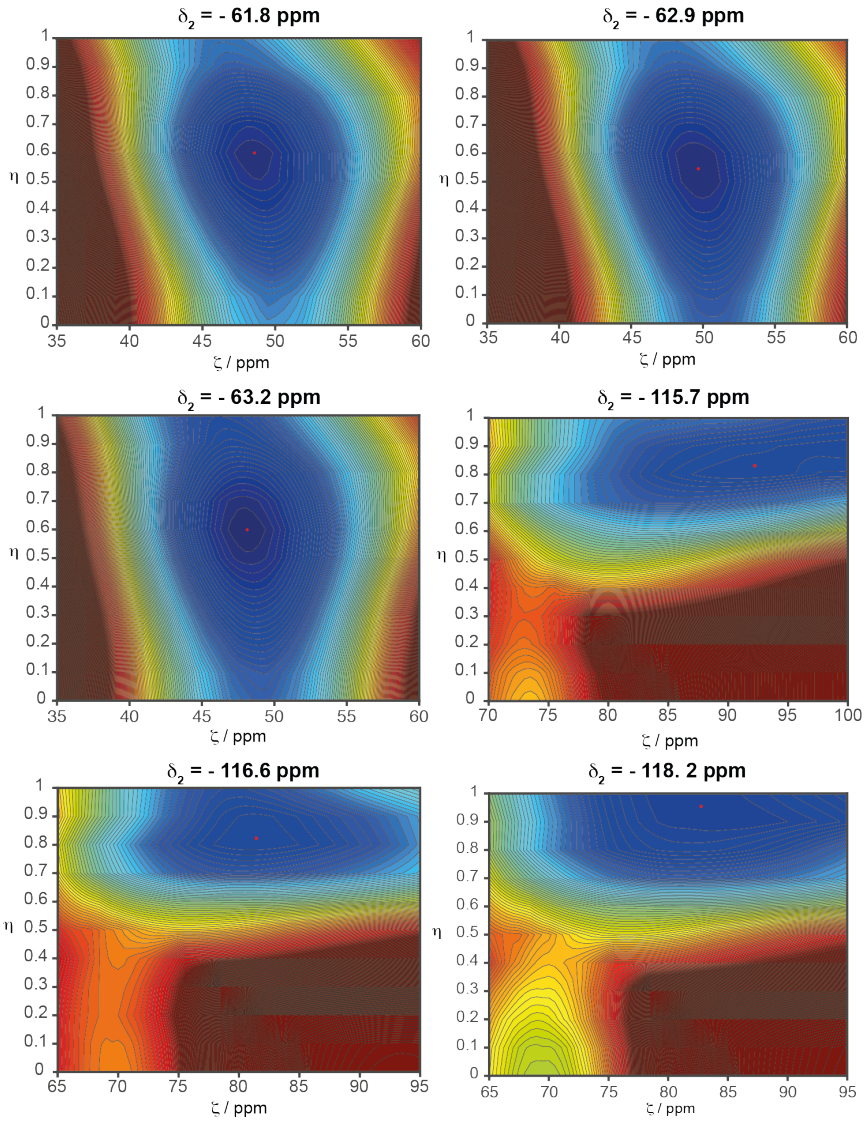


Figure 6.10: The error surfaces for all 6 different isotropic fluorine shifts. The errors have been calculated for each point on the error surface by simulating the CSA lineshape for the whole error surface and compared with the experimental line shapes. The optimum values for the CSA, $\zeta$ and asymmetry, $\eta$ for each fluorine environment is represented as a red dot on the error surface.

## 6.5 Conclusion

In this chapter it has been conclusively shown that the addition of the composite pulse, $270°_{0}90°_{180}$ makes the originally developed $R12^{4}_{5}$ experiment by Martini *et al.* significantly more robust to $B_1$ inhomogeneity. The composite pulse results in a significantly reduced central spike allowing the

recoupled $CF_3$ line shapes to be resolved. The $^{19}F$ CSA line shapes for the aryl fluoride sites with the composite pulse experiment has a more structured line shape enabling the chemical shift parameters to be extracted with less error associated with them.

In the castopitant sample there were six different fluorine sites whose CSA parameters have been extracted using the composite pulse experiment. The aryl fluorine sites for form 1 and 3 at -118.2 ppm and -116.6 ppm respectively had very similar $\zeta$. However, form 1 was found to be more axially asymmetric than form 3. The $CF_3$ environments had a significantly smaller CSA and asymmetry due to having a more spherically symmetric environment and the motional averaging which occurs with the rotation of the $CF_3$ bonds. Overall, it was found that there are only small differences in the chemical shift parameters between the two polymorphs of casopitant.

# Chapter 7

# Summary and Final Conclusions

In this work several different applications and methods of solid state NMR have been demonstrated.

In Chapter 4, solid-state NMR has been shown to be an effective technique to study the the molecular architecture and composition of plants. The majority of techniques used to study the changes in plants involves extracting the different components resulting in the loss of any structural information about the plant. A technique, such as solid-state NMR, that can give a better understanding of how pretreatment methods affect the plant structure could help development of biofuels.

In Chapter 4, the Schmidt-Rohr quantitative method alongside 2D analysis of wheat straw, was used to observe in-situ changes in the straw during a 5 day fermentation with *Aspergilus niger*. The quantitative analysis showed that cellulose and hemicellulose decreased by around 25% in the fermentation time with some of the xylan being degraded before the cellulose in the first day. The production of the cellulase enzymes by the fungus, also seen to occur in the first day and then remain constant for the

remaining fermentation time.

The 2D analysis of the wheat straw during the fermentation time showed that the cellulose peaks get progressively broader indicating the cellulose structure is becoming more amorphous. There is also a shift in the xylan C4 environment during the fermentation time which shows that the xylan is becoming unattached from the cellulose and changing its structure by going from a 2-fold conformation to a 3-fold conformation.

In Chapter 5, DNP-enhanced NMR was used to study the structure of two different forms of Naloxone. A DNP enhancement of 18 was achieved for the Naloxone alkaloid sample and a 2D refocused INADEQUATE spectrum was acquired in 16 hours rather than 216 days using standard NMR. The 2D spectrum was then used to make a complete assignment of the $^{13}$ NMR shifts of the naloxone alkaloid. The low temperatures the DNP NMR is acquired at did broaden the NMR lines compared to standard NMR but it did not affect the resolution of the carbon shifts. The addition of the radical in to the sample did cause some carbon shifts meaning the addition of the radicals did affect the crystal structure. The large sensitivity enhancement that DNP enhanced NMR provided in this case out weighs the difference in the spectrum caused by the temperature and radicals. DNP-enhanced NMR with large enhancements, such as those achieved on the naloxone, could enable whole pill analysis of drugs. This would allow an insight into the structure of the active ingredient inside the pill and whether it interacts with the bulking material the majority of the pill is made from.

Finally in Chapter 6, a two-dimensional anisotropic-isotropic experiment, which uses the $R12_5^4$ sequence to recouple the $^{19}$F CSA in the indirect dimension, was made more robust to $B_1$ inhomogeneity. This was achieved by replacing the simple 180°rf pulse as the R-element in the R-sequence

with the $270^{\circ}_{0}90^{\circ}_{180}$ composite pulse. The $^{19}F$ chemical shift parameters for two inseparable polymorphs of casopitant were measured. The measure parameters of the aryl and $CF_3$ sites of casopitant were close to the literature values of similar sites. The differences between the chemical shift parameters of the two polymorphs of casopitant were found to be very small.

# References


[1] H. Yang, T. Wang, D. Oehme, L. Petridis, M. Hong and J. D. Kubicki, *Cellulose*, 2018, **25**, 23–36.

[2] J. D. Kubiki, H.Yang, D. Sawada, H. O'Neill, D. Oehme and D. Cosgrove, *Sci. Rep. -UK*, 2018, **8**, 13983–13990.

[3] N. Pensupa, *Ph.D. thesis*, University of Nottingham, 2015.

[4] R. Dupree, T. J. Simmons, J. C. Mortimer, D. Patel, D. Iuga, S. P. Brown and P. Dupree, *Biochemistry*, 2015, **54**, 2335–2345.

[5] D. Massiot, F. Fayon, M. Capron, I. King, S. Le Calvé, B. Alonso, J.-O. Durand, B. Bujoli, Z. Gan and G. Hoatson, *Magn. Reson. Chem.*, 2002, **40**, 70–76.

[6] B. Fung, A. Khitrin and K. Ermolaev, *J. Magn. Reson.*, 2000, **142**, 97 – 101.

[7] C. Guguta, T. P. J. Peters and R. de Gelder, *Cryst. Growth Des.*, 2008, **8**, 4150–4158.

[8] S. Odedra and S. Wimperis, *J. Magn. Reson.*, 2013, **231**, 95 – 99.

[9] X. Kang, A. Kirui, A. Muszynski, M. C. D. Widanage, A. Chen, P. Azadi, P. Wang, F. Mentink-Vigier and T. Wang, *Nat. Comms.*, 2018, **9**, 2747–2758.

[10] J. Mao, W.-G. Hu, K. Schmidt-Rohr, G. Davies, E. Ghabbour and X. BS, *Soil Sci. Soc. Am. J.*, 2000, **64**, 873.

[11] A. Felix and B. Diarra, *J. Sustain. Agr.*, 1993, **3**, 37–64.

[12] J. M. Lawther, R. Sun and W. B. Banks, *J. Agr. Food Chem.*, 1995, **43**, 667–675.

[13] M. Tutt, T. Kikas and J. Olt, *Agronomy Research*, 2012, **10**, 209–276.

[14] B. C. Saha, L. B. Iten, A. M. Cotta and Y. V. Wu, *Process Biochem.*, 2005, **40**, 3693–3700.

[15] P. Kaparaju and C. Felby, *Bioresource Technol.*, 2010, **101**, 3175–3181.

[16] N. Curreli, M. B. Fadda, A. Rescigno, A. C. Rinaldi, G. Soddu, F. Sollai, S. Vaccargiu, E. Sanjust and A. Rinaldi, *Process Biochem.*, 1997, **32**, 665–670.

[17] E. M. Purcell, H. Torrey and R. Pound, *Phys. Rev.*, 1946, **69**, 37–38.

[18] F. Bloch, W. Hansen and M. Packard, *Phys. Rev.*, 1946, **70**, 474–485.

[19] W. G. Proctor and F. Yu, *Phys. Rev.*, 1950, **77**, 717.

[20] W. D. Knight, *Phys. Rev.*, 1949, **76**, 1259–1260.

[21] E. L. Hahn, *Phys. Rev.*, 1950, **80**, 580–594.

[22] M. J. Duer, *Introduction to solid-state NMR spectroscopy*, Blackwell Publishing Ltd, Oxford, 2004, pp. 1–150.

[23] E. R. Andrew, A. Bradbury and R. G. Eades, *Nature*, 1958, **182**, 1659.

[24] D. C. Apperley, R. K. Harris and P. Hodgkinson, *Solid-state NMR: Basic Principles and Practice*, Momentum Press, New York, 2012, pp. 1–139.

[25] L. B. Andreas, K. Jaudzems, J. Stanek, D. Lalli, A. Bertarello, T. Le Marchand, D. Cala-De Paepe, S. Kotelovica, I. Akopjana, B. Knott, S. Wegner, F. Engelke, A. Lesage, L. Emsley, K. Tars, T. Herrmann and G. Pintacuda, *P. Natl. Acad. Sci.-Biol*, 2016, **113**, 9187–9192.

[26] A. Pines, M. G. Gibby and J. S. Waugh, *J.Chem.Phys.*, 1972, **56**, 1776–1777.

[27] J. Schaefer, E. O. Stejskal and R. Buchdahl, *Macromolecules*, 1975, **8**, 291–296.

[28] W. P. Aue, E. Bartholdi and R. R. Ernst, *J. Chem. Phys.*, 1976, **64**, 2229–2246.

[29] Y. Garro Linck, A. Chattah, R. Graf, C. Romañuk, M. Olivera, R. Manzo, G. Monti and H. Spiess, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6590–6.

[30] X. Zhao, in *Protein Structure Determination by Solid-State NMR*, ed. G. Zhu, Springer, Berlin, Heidelberg, 2012, ch. 7, pp. 187–213.

[31] H. K. Miah, D. A. Bennett, D. Iuga and J. J. Titman, *J. Magn. Reson.*, 2013, **235**, 1 – 5.

[32] F. Martini, H. K. Miah, D. Iuga, M. Geppi and J. J. Titman, *J. Magn. Reson.*, 2015, **259**, 102 – 107.

[33] M. H. Levitt, *Spin Dynamics*, John Wiley & Sons Ltd., West Sussex, UK, 2001, pp. 1–20.

[34] J. Cavanagh, W. Fairbrother, A. PALMER III, M. Rance and N. Skelton, in *Theoretical description of NMR spectroscopy*, Academic Press, 2007, pp. 29–113.

[35] L. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton Uni. Press, Princeton, New Jersey, 1955, pp. 107–250.

[36] M. Bak, J. T. Rasmussen and N. C. Nielsen, *J. Mag. Res.*, 2000, **147**, 296 – 330.

[37] M. Mais, S. Paul, N. S. Barrow and J. J. Titman, *Johnson Matthey Techno. Rev.*, 2018, **62**, 271–278.

[38] W. T. Dixon, *J.Chem. Phys.*, 1982, **77**, 1800–1809.

[39] A. E. Bennett, C. M. Rienstra, M. Auger, K. V. Lakshmi and R. G. Griffin, *J.Chem.Phys.*, 1995, **103**, 6951–6958.

[40] G. Metz, X. Wu and S. Smith, *J. Mag.Res. A*, 1994, **110**, 219 – 227.

[41] J. Keeler, *Understanding NMR*, John Wiley & Sons Ltd., West Sussex, UK, 2nd edn., 2010, pp. 77–102.

[42] J. Keeler, in *Phase Cycling Procedures in Multiple Pulse NMR Spectroscopy of Liquids*, ed. P. Granger and R. K. Harris, Springer Netherlands, Dordrecht, 1990, pp. 103–129.

[43] R. L. Johnson and K. Schmidt-Rohr, *J. Magn. Reson.*, 2014, **239**, 44 – 49.

[44] W.-G. Hu and K. Schmidt-Rohr, *Polymer*, 2000, **41**, 2979 – 2987.

[45] K. Takegoshi, S. Nakamura and T. Terao, *Chem. Phys. Letts.*, 2001, **344**, 631 – 637.

[46] A. Lesage, M. Bardet and L. Emsley, *J. Am. Chem. Soc.*, 1999, **121**, 10987–10993.

[47] S. Cadars, J. Sein, L. Duma, A. Lesage, T. N. Pham, J. H. Baltisberger, S. P. Brown and L. Emsley, *J. Magn. Reson.*, 2007, **188**, 24 – 34.

[48] M. H. Levitt, in *Symmetry-Based Pulse Sequences in Magic-Angle Spinning Solid-State NMR*, American Cancer Society, 2007, ch. 1, pp. 1–16.

[49] T. Gillion and J. Schaefe, *J. Magn. Reson.*, 1989, **81**, 196–200.

[50] M. H. Levitt, in *Composite Pulses*, American Cancer Society, 2007, ch. 1, pp. 1–16.

[51] M. K. Pandey, M. Malon, A. Ramamoorthy and Y. Nishiyama, *J. Magn. Reson.*, 2015, **250**, 45 – 54.

[52] S. Wimperis, *J. Magn. Reson., Series A*, 1994, **109**, 221 – 231.

[53] A. W. Overhauser, *Phys. Rev.*, 1953, **92**, 411–415.

[54] T. R. Carver and C. P. Slichter, *Phys. Rev.*, 1953, **92**, 212–213.

[55] A. J. Rossini, A. Zagdoun, M. Lelli, A. Lesage, C. Coperet and L. Emsley, *Acc. Chem. Res.*, 2013, **46**, 1942–1951.

[56] B. Corzilius, *Phys. Chem. Chem. Phys.*, 2016, **18**, 27190–27204.

[57] J. F. Saeman, J. L. Bubl and E. E. Harris, *Ind. Eng. Chem.*, 1945, **17**, 35–37.

[58] R. S. Fukushima and R. D. Hatfield, *J. Agr. Food Chem.*, 2001, **49**, 3133–3139.

[59] R. H. Attalla and D. L. Vanderhart, *Science*, 1984, **223**, 283–285.

[60] A. Šturcová, I. His, D. C. Apperley, J. Sugiyama and M. C. Jarvis, *Biomacromolecules*, 2004, **5**, 1333–1339.

[61] Y. Nishiyama, P. Langan and H. Chanzy, *J. Am. Chem. Soc.*, 2002, **124**, 9074–9082.

[62] C. Driemeier and L. H. Francisco, *Cellulose*, 2014, **21**, 3161–3169.

[63] T. J. Simmons, J. C. Mortimer, O. D. Bernardinelli, A. C. Poppler, S. P. Brown, E. R. deAzervedo, R. Dupree and P. Dupree, *Nat. Comms.*, 2016, **7**, 1–9.

[64] R. H. Newman and J. A. Hemmingson, *Cellulose*, 1995, **2**, 95–110.

[65] A. N. Fernandes, L. H. Thomas, C. M. Altaner, P. Callow, V. T. Forsyth, D. C. Apperley, C. J. Kennedy and M. C. Jarvis, *P. Natl. A. Sci.*, 2011, 1195–1203.

[66] M.-A. Ha, D. C. Apperley, B. W. Evans, I. M. Huxham, W. G. Jardine, R. J. Viëtor, D. Reis, B. Vian and M. Jarvis, *The Plant Journal*, 1998, **16**, 183–190.

[67] D. P. Delmer and Y. Amor, *The Plant Cell*, 1995, **7**, 987–1000.

[68] P. Bajpai, *Pretreatment of lignocellulosic biomass for biofuel production*, Spinger Singapore, Singapore, 2016, pp. 17–70.

[69] M. Busse-Wicher, T. C. F. Gomes, T. Tryfona, N. Nikolovski, K. Scott, N. J. Grantham, D. N. Bolam, M. S. Skaf and P. Dupree, *The Plant Journal*, 2014, **79**, 492–506.

[70] R. van Acker, R. Vanholm, V. Storme, J. C. Mortimer, P. Dupree and W. Boerjan, *Biotechnol. Biofuels*, 2013, **6**, 1–17.

[71] R. Vanholme, B. Demedts, K. Morreel, J. ralph and W. Boerjan, *Plant Phys.*, 2010, **153**, 895–905.

[72] N. Takahashi and T. Koshijima, *Wood Sci. Technol.*, 1988, **22**, 231–241.

[73] J. F. Saeman, J. L. Bubl and E. E. Harris, *Ind. Eng. Chem*, 1945, **17**, 35–37.

[74] A. M. J. Kootstra, H. H. Beeftink, E. L. Scott and J. P. M. Sanders, *Biochem. Engin. J.l*, 2009, **46**, 126–131.

[75] K. Jackson, I. R. Jonasson and G. B. Skippen, *Earth-Sci. Rev.*, 1978, **14**, 97–146.

[76] B. A. G. de Melo, F. L. Motta and M. H. A. Santana, *Mat. Sci. Eng. C*, 2016, **62**, 967 – 974.

[77] J. Mao, W. Hu, G. Ding, K. Schmidt-Rohr, G. Davies, E. Ghabbour and B. Xing, *Int. J. Environ. An. Ch.*, 2002, **82**, 183–196.

[78] F. M. Swaine, in *Geochemistry of Humus*, Pergamon Press Limited, 1963, ch. 4, pp. 87–147.

[79] J. A. Leenheer, D. M. McKnight, E. M. Thurman and P. MacCarthy, *U.S Geological survey Water Supply Paper*, 1994, **2373**, 195–213.

[80] W. Soetaert and E. J. Vandamme, *Biofuels*, John Wiley and Sons Ltd., West Sussex, UK, 2009.

[81] L. L. Nass, P. A. A. Pereira and D. Ellis, *Crop Science*, 2007, **47**, 2228 – 2237.

[82] S. Searle and C. Malins, *The White paper*, 2013, **422**, 1–15.

[83] S. Morales-delaRosa, J. Campos-Martin and J. L. G. Fierro, *Cellulose*, 2014, **21**, 2397–2407.

[84] A. Kumar, M. Nakachew, B. Yilkal and Y. Mukesh, *J. Chem. Environ.*, 2018, **22**, 62–67.

[85] J. Rencoret, A. Gutiérrez, L. Nieto, J. Jiménez-Barbero, C. Faulds, H. Kim, J. Ralph, A. T. Martinez and J. del Río, *Plant physiol.*, 2011, **155**, 667–682.

[86] S. Wu, J. Mclaren, R. Madl and D. Wang, *Biofuels from Lignocellulosic Biomass*, Springer Science and Business Media, Germany, 2009, pp. 19–41.

[87] C. Karunanithy and K. Muthukumarappan, *Ind. Crops Prod.*, 2011, **33**, 188–199.

[88] A. Hideno, H. Inoue, K. Tsukahara, S. Fujimoto, T. Minowa, S. Inoue, T. Endo and S. Sawayama, *Bioresource Technol.*, 2009, **100**, 2706–11.

[89] A. S. da Silva, H. Inoue, T. Endo, S. Yano and E. P. Bon, *Bioresource Technol.*, 2010, **101**, 7402 – 7409.

[90] K.-L. Chang, J. Thitikorn-amorn, J.-F. Hsieh, B.-M. Ou, S.-H. Chen, K. Ratanakhanokchai, P.-J. Huang and S.-T. Chen, *Biomass and Bioenergy*, 2011, **35**, 90–95.

[91] M. Harun, A. D. Radiah, Z. Z. Abidin and R. Yunus, *Bioresource Technol.*, 2011, **102**, 5193 – 5199.

[92] B. C. Saha, A. Biswas and M. A. Cotta, *J. Biobased Mat. Bioenergy*, 2008, **2**, 210–217.

[93] S. Y. Yoon, S. H. Han and S. J. Shin, *Energy (Oxford)*, 2014, **77**, 19–24.

[94] K. Janga, M.-B. Hägg and S. Moe, *BioResources*, 2011, **7**, 391 –411.

[95] L. J. Jönsson and C. Martín, *Bioresource Technol.*, 2016, **199**, 103 – 112.

[96] L. Feng, L. Qin, Z.-H. Liu, C.-Y. Dong, B.-Z. Li and Y.-J. Yuan, *BioResources*, 2013, **9**, 24–38.

[97] M. Yang, S. Kuittinen, J. Zhang, M. Keinänen and A. Pappinen, *Bioresource Technol.*, 2013, **146C**, 444–450.

[98] J. R. Monte, D. F. Laurito-Friend, S. I. Mussatto, A. Ferraz and A. M. F. Milagres, *Biotechnol. Prog.*, 2018, **34**, 944–951.

[99] S. Mohapatra, C. Mishra, S. S. Behera and H. Thatoi, *Renew. Sustain. Energy Rev.*, 2017, **78**, 1007 – 1032.

[100] R. B. Nair, M. Kalif, J. A. Ferreira, M. J. Taherzadeh and P. R. Lennartsson, *Bioresource Technol.*, 2017, **245**, 145 – 151.

[101] N. Mosier, C. Wyman, B. Dale, R. Elander, Y. Lee, M. Holtzapple and M. Ladisch, *Bioresource Technol.*, 2005, **96**, 673 – 686.

[102] V. B. Agbor, N. Cicek, R. Sparling, A. Berlin and D. B. Levin, *Biotechnol. Adv.*, 2011, **29**, 675 – 685.

[103] E. Z. Hoşgün, D. Berikten, M. Kıvanç and B. Bozan, *Fuel*, 2017, **196**, 280 – 287.

[104] B. Li, L. Ding, H. Xu, X. Mu and H. Wang, *Resour. Conserv. Recy.*, 2017, **122**, 307 – 318.

[105] J. Xu, J. J. Cheng, R. Sharma-Shivappa and J. Burns, *Energy & Fuels*, 2010, **24**, 2113–2119.

[106] J. N. Himmelsbach, D. Raman, R. P. Anex, R. T. Burns and C. R. Faulhaber, *Trans. Asabe*, 2010, **53**, 1921–1927.

[107] H. Y. Yoo, J. H. Lee, D. S. Kim, J. H. Lee, S. K. Lee, S. J. Lee, C. Park and S. W. Kim, *J. Ind. Eng. Chem.*, 2017, **51**, 303 – 311.

[108] A. Bhalla, N. Bansal, R. J. Stoklosa, M. Fountain, J. Ralph, D. B. Hodge and E. L. Hegg, *Biotechnol Biofuels*, 2016, **9**, 34.

[109] M. Østergaard Petersen, J. Larsen and M. H. Thomsen, *Biomass and Bioenergy*, 2009, **33**, 834 – 840.

[110] R. Sindhu, P. Binod and A. Pandey, *Bioresource Technol.*, 2016, **199**, 76 – 82.

[111] J. Zeng, D. Singh and S. Chen, *Bioresource Technol.*, 2011, **102**, 3206 – 3214.

[112] C. Xu, F. Ma, X. Zhang and S. Chen, *J. Agr. Food Chem.*, 2010, **58**, 10893–10898.

[113] P. Daly, J. M. van Munster, M. J. Blythe, R. Ibbett, M. Kokolski, S. Gaddipati, E. Lindquist, V. R. Singan, K. W. Barry, A. Lipzen, C. Y. Ngan, C. J. Petzold, L. J. G. Chan, S. T. Pullan, S. Delmas, P. R. Waldron, I. V. Grigoriev, G. A. Tucker, B. A. Simmons and D. B. Archer, *Biotechnol. Biofuels*, 2017, **10**, 1–19.

[114] C. R. Morcombe and K. W. Zilm, *J. Magn. Reson.*, 2003, **162**, 479–486.

[115] D. Marion and K. Wüthrich, *Biochem. Bioph. Res. Co.*, 1983, **113**, 967 – 974.

[116] K. M. Holtman, N. Chem, M. A. Choppell, J. F. Kadla, L. Xu and J. Mao, *J. Agr. Food Chem.*, 2010, **58**, 9882–9892.

[117] P. G. Hatcher, *Org. Geochem.*, 1987, **11**, 31–39.

[118] P. T. Larsson, E.-L. Hult, K. Wickholm, E. Pettersson and T. Iversen, *Solid State Nucl. Mag.*, 1999, **15**, 31–40.

[119] M. D. Perez, T. Wang, A. Salazar, O. A. Zabontina and M. Hong, *Magn. Reson. Chem.*, 2012, **50**, 539–550.

[120] R. Atalla and D. VanderHart, *Solid State Nuclear Magnetic Resonance*, 1999, **15**, 1 – 19.

[121] R. J. Vietor, R. Newman, M.-A. Ha, D. C. Apperley and M. C. Jarvis, *The Plant Journal*, 2002, **30**, 721–731.

[122] M. D. Perez, Y. Zhang, J. Hayes, A. Salazar, O. A. Zabotina and M. Hong, *Biochemistry*, 2011, **50**, 989–1000.

[123] L. He and N. Terashima, *Holzforschung*, 1991, **45**, 191–198.

[124] V. Arantes and J. N. Saddler, *Biotechnol. Biofuels*, 2010, **3**, 1–11.

[125] R. L. Sime, R. Forehand and R. J. Sime, *Acta Cryst.*, 1975, **31**, 2326.

[126] I. L. Karle, *Acta Cryst.*, 1974, **30**, 1682–1686.

[127] C. Guguta, E. van Eck and R. de Gelder, *Cryst. Growth Des.*, 2009, **9**, 3384–3395.

[128] A. S. L. Thankamony, J. J. Wittmann, M. Kaushik and B. Corzilius, *Prog. Nucl. Mag. Res. Sp.*, 2017, **102-103**, 120 – 195.

[129] U. H. Dürr, S. L. Grage, R. Witter and A. S. Ulrich, *J. Magn. Reson.*, 2008, **191**, 7 – 15.

[130] G. S. Prakash and F. Wang, *Fine Chemicals*, 2012, **30**, 1.

[131] J. Herzfeld and A. E. Berger, *J. Chem. Phys.*, 1980, **73**, 6021–6030.

[132] S. L. Grage, U. H. Dürr, S. Afonin, P. K. Mikhailiuk, I. V. Komarov and A. S. Ulrich, *J. Magn. Reson.*, 2008, **191**, 16 – 23.

[133] H. K. Miah, R. Cresswell, D. Iuga and J. J. Titman, *Solid State Nucl. Mag.*, 2017, **87**, 67 – 72.

[134] D. H. Brouwer and J. A. Ripmeester, *J. Magn. Reson.*, 2007, **185**, 173 – 178.

[135] L. Duma, D. Abergel, P. Tekely and G. Bodenhausen, *Chem. Commun.*, 2008, 2361–2363.

[136] G. Hou, S. Paramasivam, S. Yan and T. Polenova, *J. Am. Chem. Soc.*, 2013, **135**, 1358–1368.

[137] R. Zhang, K. H. Mroue and A. Ramamoorthy, *J. Chem. Phys.*, 2015, **143**, 144201–144206.

[138] Z. Cimarosti, C. Castagnoli, M. Rossetti, M. Scarati, C. Day, B. Johnson and P. Westerduin, *Organic Process Res. Dev.*, 2010, **14**, 1337–1346.

[139] J.-P. Brog, C.-L. Chanez, A. Crochet and K. M. Fromm, *RSC Adv.*, 2013, **3**, 16905–16931.