University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Student Research Projects, Dissertations, and Theses - Chemistry Department

Chemistry, Department of

8-2012

Experimental and Theoretical Studies in Solid-state Nuclear Magnetic Resonance

Monica N. Kinde
University of Nebraska Lincoln, chemmon1@gmail.com

Follow this and additional works at: https://digitalcommons.unl.edu/chemistrydiss

 Part of the Physical Chemistry Commons

Kinde, Monica N., "Experimental and Theoretical Studies in Solid-state Nuclear Magnetic Resonance" (2012). *Student Research Projects, Dissertations, and Theses - Chemistry Department*. 38.
https://digitalcommons.unl.edu/chemistrydiss/38

This Article is brought to you for free and open access by the Chemistry, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Student Research Projects, Dissertations, and Theses - Chemistry Department by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

EXPERIMENTAL AND THEORETICAL STUDIES IN SOLID-STATE NUCLEAR MAGNETIC RESONANCE

by

Monica N. Kinde

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Chemistry

Under the Supervision of Professor Gerard S. Harbison

Lincoln, Nebraska

August, 2012

EXPERIMENTAL AND THEORETICAL STUDIES IN SOLID-STATE NUCLEAR MAGNETIC RESONANCE

Monica N. Kinde, Ph.D.

University of Nebraska, 2012

Advisor: Gerard S. Harbison

Solid-state nuclear magnetic resonance (SSNMR) has proven to be a powerful tool for probing molecular structure and dynamics. Deuterium SSNMR is particularly useful due to the presence of anisotropic interactions whose motional averaging contributes structural and dynamical insight. The magnitude and type of molecular motion can be determined from analysis of solid echo deuterium lineshapes and/or relaxation studies. This work uses various applications of solid-state NMR as well as *ab initio* and density functional computational methods to study three different areas of physical chemistry: biophysical chemistry, materials science and fundamental concepts of physical chemistry. The first project addressed the dynamics of biomolecules as an aid in understanding protein recognition and binding to damaged DNA by selectively labeling the [2″-^{2}H] furanose ring in two deoxynucleosides. The next project uses ^{1}H, ^{2}H, and ^{13}C SSNMR to gain structural insight of self-assembling organic molecules by assessing experimental and theoretical nuclear magnetic parameters. Finally, variant isotopic labeling and deuterium SSNMR are used to understand fundamental thermodynamic isotope effects in amino acids. Of central importance is the dependence on electronic and magnetic parameters in chemical environment; hydrogen bond strength directly correlates to the magnitude of these parameters as obtained experimentally and theoretically.

This work is dedicated to:

my best friend and fiancé, John Persons,

my mother, Michelle,

my father, Doug,

my three siblings, Rebecca, Derik and Jessica,

for showing me nothing but love and support throughout this process.

Acknowledgements

I would first like to extend the deepest appreciation and gratitude to my advisor and mentor, Professor Gerard Harbison. Your guidance, knowledge and expertise in the field of NMR has provided me learning and other opportunities too numerous to mention. I extend a hearty thank you to those individuals that served on my supervisory committee: Professors Hui Li (reader), Robert Powers (reader), Robert Spreitzer and Xiao Cheng Zeng. I appreciate all the insight you have provided me throughout my tenure at UNL. To my lab mates that have taught me so much over the years, John Persons, Jun Zhou, Xiongjiang Wu, Yali Wang, Adrienne Roehrich, Chunlei Guo; my sincere appreciation.

I have to convey my utmost appreciation for Dr. Joe Dumais for patiently teaching me the ins and outs of experimental NMR. To the other individuals of the UNL NMR Instrumentation Center, Sara Basiaga, Dr. Martha Morton, Dr. Matt Shortridge, John Persons and Jayson Kempinger; thank you so much for your patience, guidance and for the willingness to work around our crazy solid-state voodoo.

To the many colleagues-turned-friends I have met at UNL, thank you. Judy, Mattie, Sara, Rachel, Neil, and Jayson, your friendships and intellectual collaborations mean very much to me. To the amazing and hardworking staff at UNL Chemistry that keeps this department running smoothly, I appreciate all your help throughout the years. To the friends that supported me from afar, thank you for your endless support, love and kind words. To the wonderful water aerobics ladies, thank you for always turning my frustrated days into light-hearted exercise (and socializing).

This journey would have never come to fruition had it not been for the loving support of my biggest fans and motivators, my parents Michelle Burgin and Douglas Kinde. I love and respect you so much and appreciate your willingness to stand by me in times of joy and in the more difficult times. To my beautiful, talented and wonderful siblings Rebecca, Derik and Jessica: it has been such a pleasure watching you three grow into successful and independent young adults. Nothing makes me more proud than to call you three my siblings. To my grandparents Joyce Kinde and Eugene and Betty Laudenslager: thank you for a lifetime of unconditional love and endless support. To my late grandfather Bill Kinde, thank you for passing on that wonderful "Kinde" gene to me. I would like to think I have made you very proud.

To my new "Wisconsin family," Carol, Joe, Steve and Grandma Gray: thank you for accepting me as one of your own. I may never be a Packers fan, but I think I can live with occasionally donning your Cardinal and White if you can muster up the strength to sport the Maize and Blue every once in a while.

Finally, to my one and only, my sun and stars, my best friend, my partner in crime, John Persons: I am so glad we were able to traverse this rocky path together. Without your love and support, this process would have been much more difficult. I am so completely happy that we were able to become such good friends these last several years and can't wait for the new adventures we have waiting for us in Pittsburgh. We have the rest of our lives to make thousands of more wonderful memories together and I look forward to making every single one of them with you.

Table of Contents

List of Tables

List of Figures

Chapter 1

Overview of NMR, Theory, and *ab initio* Calculations

Summary

Since the first experiments in 1945, Nuclear Magnetic Resonance (NMR) has been rapidly developed and is currently one of the most important spectroscopic techniques in the physical sciences. Its ability to probe the chemical environment of spin-active nuclei has proven greatly beneficial in structure determination, though being originally limited to the study of small organic compounds in solution. However, advances in experimental techniques and theory have made NMR amenable to the study of larger biomolecules (proteins, peptides, nucleic acids), powdered solids, single crystals and amorphous materials. This dissertation will describe the various spectroscopic experiments and supporting *ab initio* calculations I have completed while a student in the research group of Professor Gerard Harbison at the University of Nebraska. This chapter, specifically, will provide an overview and brief history of NMR as a technique and its underlying theory, as well as serve as an introduction to *ab initio* and DFT methods in the calculation of nuclear magnetic parameters.

1.1 Introduction

The theoretical basis for Nuclear Magnetic Resonance (NMR) was laid by Pauli in the 1920's.[1] He proposed that nuclei have properties called spin and magnetic moment that, when perturbed with a magnetic field, would cause splitting of their energy levels. I. I. Rabi[2] verified this in 1938 when he laid the theoretical groundwork for what would become the first NMR experiments. Bloch[4] and Purcell[3] completed these first experiments independently in 1945 when they discovered that in a magnetic field, nuclei absorb electromagnetic radiation due to the splitting of their energy levels. It was soon realized that the molecular environment of an NMR active nucleus directly affects the manner in which radio-frequency (r.f.) radiation was absorbed. Therefore, it was inferred that this absorption could correlate to molecular structure.

Early NMR experiments developed as continuous-wave experiments; the frequency was fixed while the magnetic field was swept over a range to study the system's response at each frequency. This method, while an important first step in the development of NMR, was rather inefficient, resulting in a poor signal-to-noise ratio. In the 1970's Fourier transform NMR became available and has since become the dominant technique in NMR. In this method, pulses of strong radio-frequency irradiation excite transitions between spin states. Once the radio-frequency pulse is removed, the spin angular momentum of the perturbed nuclei returns or relaxes back to its equilibrium orientation. This equilibrium position is parallel to the static magnetic field. During this relaxation period, an oscillating signal called a free induction decay (FID) is emitted by the excited nuclei as they relax back to their equilibrium orientations. This time-domain

signal is converted to a frequency-domain signal by the use of Fourier transformation, hence the name Fourier transform NMR.

The first NMR experiments were performed on both liquids and solids; Purcell's first experiments in NMR were performed on paraffin wax while Bloch initially studied water. However, the initial lack of resolution in the NMR experiments of solid materials was problematic, and early advances in NMR were mostly for liquid materials and solutions. Molecular tumbling of liquids and solutes in solution allow orientation-dependent interactions — or anisotropic interactions — to be averaged to zero due to the rapid reorientation of molecules. What results in the spectra of these liquid materials are the sharp, well-resolved lines one might typically associate with an NMR signal. These signals represent the isotropic average of the molecule's magnetic properties with no apparent dependence on the molecule's orientation in space. NMR spectra of solid materials are dominated by these relatively large-scale anisotropic interactions, meaning, rather than seeing well resolved, ~1 Hz wide lineshapes as a result of isotropic interactions, a broadened lineshape on the order of kHz or MHz results.[25] Great strides have been made to improve the resolution of solid-state NMR spectra including the dilution of spins to reduce the broadening resulting from direct dipolar coupling, and physical manipulation of the sample using methods such as as magic and dynamic angle spinning (MAS and DAS, respectively). I will address some of these methods in later sections (see Chapter 3 and 4) as they pertain to my various research topics.

1.2 Nuclear Magnetic Resonance Theory

To fully understand nuclear magnetic resonance as an experimental technique, one must understand a few fundamental quantum chemical topics. I will first discuss these topics in general and how they pertain to the NMR experiment, focusing initially on the classical description of magnetization. I will then discuss various elements of the NMR Hamiltonian and dissect each term of the NMR Hamiltonian as it applies to this dissertation.

1.2.1 Spin magnetization. NMR focuses on nuclei that are magnetically active, i.e. nuclei with a non-zero spin. Nuclei that possess spin will behave differently when placed in a magnetic field due to the total spin angular momentum or I. The magnitude of the spin angular momentum is:[24]

$$\left[I(I+1)\right]^{1/2}\hbar \tag{1.1}$$

where $\hbar$ is reduced Planck's constant and I, or the "spin," may be 0, 1/2, 1, 3/2, and so on.[5] A majority of this work focuses on the NMR of spin-1/2 and spin-1 nuclei. A nucleus of spin I in a magnetic field will have $(2I + 1)$ degenerate energy levels: a $I = 1/2$ nuclei has two states (represented by two quantum numbers, $m_S = \pm 1/2$) while $I = 1$ has three states ($m_S = 0, \pm 1$).[27] When placed in a magnetic field the spins come to an equilibrium orientation under the Zeeman interaction (interaction with the static magnetic field). With respect to the magnetic field, the x and y components of the spin ensemble averages to zero while the z component does not. This results in the bulk magnetic moment of I_Z aligning parallel to the static field. The spin angular momentum causes the nuclei to precess, or rotate, about the static field at a characteristic frequency called the

Larmor frequency (Figure 1.1). This is of fundamental concern in NMR as this form of spectroscopy is concerned with the rotational movement of nuclei: as the nucleus rotates, the inherent charge of the nucleus rotates as well, forming an induced magnetic field called the magnetic moment, $\boldsymbol{\mu}$. The nuclear energy change associated with this process is called the Larmor energy and is given as[26]:

$$\Delta E_l = \hbar \gamma B_z = \hbar \omega_l \tag{1.2}$$

where $\mathbf{B}_0$ is the static magnetic field, ω_l is the Larmor frequency, and γ is known at the gyromagnetic ratio. The gyromagnetic ratio is a proportionality constant that relates the nuclear magnetic dipole moment to the spin angular momentum by:

$$\boldsymbol{\mu} = \gamma \mathbf{I} \tag{1.3}$$

The magnetic moment orientation is quantized, leading to discrete energy levels. Radiofrequency electromagnetic radiation promotes the transition between these energy levels and the energy difference can be measured. The differences in energy levels are directly proportional to the strength of the magnetic field at the nucleus.

Fundamentally, the time-dependent Schrödinger equation involves all of the nuclei and electrons of an atom. Therefore, the full Hamiltonian operator describes all of the interactions between these nuclei and electrons. This full wave function and Hamiltonian is not useful, as no analytical solution barring gross approximations exists.

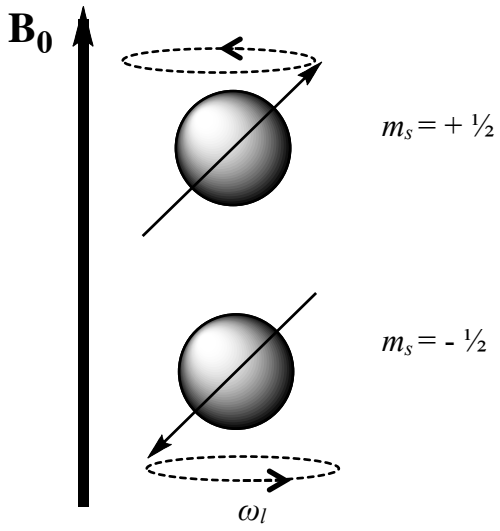

Figure 1.1. Spin-1/2 nuclei in a static magnetic field, $\mathbf{B_0}$. The angular momenta of the nuclei precess about the magnetic field and the orientation with respect to the magnetic field depends on the spin quantum number, m_s. The precession about the static magnetic field occurs as a characteristic frequency of the nuclei called the Larmor frequency, ω_l.

In NMR, it is convenient to truncate the wave function and Hamiltonian so that only the nuclear spin states are considered: [9]

$$\frac{d}{dt}\left|\Psi_{spin}(t)\right\rangle = -i\ \mathcal{H}_{spin}\left|\Psi_{spin}(t)\right\rangle \tag{1.4}$$

In this equation, $\mathcal{H}_{spin}$ is the nuclear spin Hamiltonian and contains terms that describe the orientation dependence of the nuclear energy. This nuclear spin Hamiltonian consists of terms that include both internal and external interactions of the nuclei. The external interactions include the previously mentioned Zeeman interaction (the interaction of the nucleus with the static magnetic field) as well as the interaction of a spin with an oscillating magnetic field perpendicular to the static field. This oscillating field arises from radio-frequency pulses in a coil. The internal interactions are much smaller in magnitude compared to the external interactions and are therefore usually treated as small perturbations of the system. These internal spin interactions are inherently dependent on the chemical environment of the spins and include chemical shielding, dipole-dipole, J-coupling and quadrupole interactions. Therefore, the total nuclear NMR Hamiltonian can be written as[26]:

$$\mathcal{H}_{total} = \mathcal{H}_Z + \mathcal{H}_{RF} + \mathcal{H}_{CS} + \mathcal{H}_D + \mathcal{H}_Q \tag{1.5}$$

In the following sections, each term in the total nuclear Hamiltonian will be discussed in greater detail. I will also discuss how each term affects a typical NMR experiment. First, however, it is useful to adopt a new frame of reference that is rotating at the Larmor frequency about the *z*-axis of the laboratory frame. We can designate the *z*-axis of the

laboratory frame as parallel to the static field. The new coordinates can be transformed out of the lab frame into this new rotational frame of reference by:

$$\begin{aligned} x_{rot} &= x_{lab}\cos(\omega_l t) + y_{lab}\sin(\omega_l t) \\ y_{rot} &= y_{lab}\cos(\omega_l t) - x_{lab}\sin(\omega_l t) \\ z_{rot} &= z_{lab} \end{aligned} \tag{1.6}$$

In these equations, (x_{lab}, y_{lab}, z_{lab}) are the original lab-frame coordinates and (x_{rot}, y_{rot}, z_{rot}) are the new coordinates in the rotating frame of reference. [28] Notice that for spins exactly on resonance, the effective magnetization in the z direction is canceled.

1.2.2 The Zeeman interaction. The Zeeman interaction is the interaction of a nuclear spin with the static magnetic field, $\mathbf{B_0}$. This field is chosen to be along the z direction. The Zeeman interaction is then given by the expression:

$$\mathcal{H}_Z = -\gamma B_z \mathrm{I}_z \tag{1.7}$$

where I_z is the z component of the angular momentum. γ is again the gyromagnetic ratio and is characteristic for the particular nucleus. Again, we see that the magnitude of the Zeeman interaction is dependent on the Larmor frequency, $\omega_l = -\gamma B_z$.

1.2.3 The radio-frequency interaction. Of utmost importance in an NMR experiment is the use of a radio-frequency (r.f.) coil. The NMR sample is placed inside an r.f. coil that is usually oriented perpendicular to the static field (usually defined as the x-axis of the laboratory frame). When subjected to r.f. pulses, a relatively small transverse magnetic field is generated ($B_{\mathbf{RF}}$), which perturbs the equilibrium orientation of the spins away from z The laboratory frame r.f. Hamiltonian is given as:

$$\mathcal{H}_{RF}(t) \cong -\frac{1}{2}\gamma B_{RF}\left[\cos(\omega_{ref}t + \phi_p)I_x + \sin(\omega_{ref}t + \phi_p)I_y\right] \tag{1.8}$$

In Equation 1.8, ω_{ref} is the spectrometer's reference frequency and ϕ_p is the phase of the applied r.f. pulse. As one might imagine, this interaction is quite large. However, this interaction is zero at all times except during the actual r.f. pulse. Therefore, we can assume that this interaction is zero throughout the duration of the NMR experiment and data acquisition.

1.2.4 The chemical shielding interaction. We now shift our attention from external interactions to the internal interactions present in an NMR experiment. An externally applied magnetic field induces a current in the electron cloud of an atom. The circulating current then generates an induced magnetic field. This induced field can interact with the magnetic dipole moment of the nucleus. The induced field is significantly smaller than the static or transverse field, however it is still able to produce a measurable effect on the frequency of spin precession. The total magnetic field felt by the nuclear spin is the sum of the static field and the induced field; $\mathbf{B_{loc}} = \mathbf{B_0} + \mathbf{B_{induced}}$. This induced field can either shield or deshield the external field by one of two mechanisms. The induced field produced by the circulating electrons around the nucleus tends to occur around the center of motion, resulting in a shielded nucleus. This is known as diamagnetic shielding and is relatively constant regardless of the type of atom or the chemical environment. Paramagnetic effects cause the external magnetic field to mix excited electronic states with the ground state, essentially mixing in small paramagnetic effects. What results is a strengthening of the local field at the nucleus, producing what is known as a deshielded effect.

The chemical shielding Hamiltonian is a function of the applied static field ($\mathbf{B_0}$), a second-order chemical shielding tensor ($\boldsymbol{\sigma}$), the gyromagnetic ratio ($\boldsymbol{\gamma}$), and the spin angular momentum ($\boldsymbol{I}$) as shown below:

$$\mathcal{H}_{CS} = -\gamma \mathbf{B_0} \cdot \sigma \cdot \boldsymbol{I} \tag{1.9}$$

We can always choose a frame of reference in which the chemical shielding tensor is diagonal. We call this frame of reference the principal axis frame (PAF). The chemical shielding tensor in the PAF is given as:

$$\sigma^{PAF} = \begin{pmatrix} \sigma_{xx}^{PAF} & 0 & 0 \\ 0 & \sigma_{yy}^{PAF} & 0 \\ 0 & 0 & \sigma_{zz}^{PAF} \end{pmatrix} \tag{1.10}$$

The principal components are chosen so that $|\sigma_{zz}| \geq |\sigma_{xx}| \geq |\sigma_{yy}|$. In an isotropic liquid, the chemical shielding Hamiltonian simplifies to:

$$\mathcal{H}_{CS} = -\gamma \boldsymbol{B}_0 \sigma_{iso} I_z \tag{1.11}$$

where σ_{iso} is called the isotropic chemical shift and is related to the elements of the chemical shift tensor in the principal axis frame by:

$$\sigma_{iso} = \frac{1}{3}\left(\sigma_{xx}^{PAF} + \sigma_{yy}^{PAF} + \sigma_{zz}^{PAF}\right) \tag{1.12}$$

Combining this with the previously mentioned Larmor frequency, we can get a new chemical shift Larmor frequency of the form:

$$\omega_{CS} = -\gamma B_z \left(1 - \sigma_{iso}\right) \tag{1.13}$$

As the name suggests, the isotropic chemical shift does not depend on the orientation of the molecule with respect to the magnetic field. However, there are additional considerations for an anisotropic material, adding an inherent dependence on the orientation of nuclear spins with respect to the magnetic field. The magnitude of the orientation-dependent part of the chemical shift is called the chemical shift anisotropy (*CSA* or $\Delta\sigma$) and is given by the equation:

$$\Delta\sigma = \sigma_{zz}^{PAF} - \sigma_{iso} = \left(\sigma_{zz}^{PAF} - \frac{\sigma_{xx}^{PAF} + \sigma_{yy}^{PAF}}{2} \right) \tag{1.14}$$

Another important parameter in solid-state NMR is the asymmetry parameter or η. This essentially measures the deviation of a solid-state NMR lineshape from a perfectly symmetric Pake doublet (see Chapter 3) and can range from 0 (perfectly symmetric) to 1 (completely asymmetric). There are two asymmetry parameters: the chemical shielding asymmetry parameter and the quadrupolar asymmetry parameter (see section 1.2.7). The chemical shielding asymmetry parameter can be defined using the principal values of the chemical shielding tensor or by the principal components of the electric field gradient.[25]

$$\eta_{cs} = \frac{\sigma_{xx}^{PAF} - \sigma_{yy}^{PAF}}{\Delta\sigma} \tag{1.15}$$

1.2.5 The dipole-dipole interaction. The precession of a nucleus' magnetic moment about the static field produces smaller, localized magnetic fields. These localized magnetic fields can interact with adjacent nuclei through the process called dipolar coupling.

Homonuclear and heteronuclear dipolar coupling has been studied extensively in solid state NMR as the presence of dipolar coupling can lead to the reduction in resolution due to broadening. This broadening is not problematic in solution, because dipolar coupling averages to zero in solution due to rapid molecular motions. The Hamiltonian describing dipolar coupling between two spin, I_j and I_k is[26]:

$$\mathcal{H}_D = -\frac{\mu_0 \gamma_j \gamma_k \hbar}{4\pi r^3}\left[3\left(\mathbf{I}_j \cdot \mathbf{r}_{jk}\right)\left(\mathbf{I}_k \cdot \mathbf{r}_{jk}\right) - \left(\mathbf{I}_j \cdot \mathbf{I}_k\right)\right] \tag{1.16}$$

In this equation, r_{jk} is the unit vector along the internuclear axis, parallel to the line joining the centers of the two spins, γ are the two gyromagnetic ratios corresponding to the j and k nuclei, r is the distance in meters between the two spins, and μ_0 is the permittivity of the vacuum (4π x 10^{-7} H m^{-1}). While the collection of constants in the front of the dipolar coupling Hamiltonian, often called the dipolar-coupling constant, is independent of orientation of the nuclear spins, it is the unit vector in the full Hamiltonian that provides the inherent dependence on nuclear orientation. Like other anisotropic interactions we will see, the dipolar coupling Hamiltonian can be approximated using secular approximation.

$$\mathcal{H}_D = -\frac{\mu_0 \gamma_j \gamma_k \hbar}{8\pi r^3}\left(3\cos^2\theta - 1\right)\left[3\left(I_{zj} \cdot I_{zk}\right) - \left(\mathbf{I}_j \cdot \mathbf{I}_k\right)\right] \tag{1.17}$$

In this approximated Hamiltonian, a term arises that is dependent on the orientation of the unit vector between the two spins and the static magnetic field. This term, ($3\cos^2\theta$ -1) in which the angle theta describes the orientation between $\mathbf{r}_{jk}$ and the static magnetic field, approaches the limit of zero as $\theta \rightarrow 54.74°$. As we shall see in Chapter 3, this is used

extensively in magic angle spinning NMR (MAS NMR) so that the anisotropic dipolar coupling can be averaged to zero or nearly-zero if a sample is place at an angle of 54.74° to the magnetic field and rapidly rotated. Studying dipolar coupling of compounds not suitable for diffraction studies is a good way to gain structural insight. As referenced in the Hamiltonian (eq. 1.16), the magnitude of the dipolar coupling depends on the r^{-3} distance between the nuclear spins. Therefore, if one were able to successfully measure the dipolar coupling between two spins, this would directly correlate to the internuclear distance between the two spins.

1.2.6 The J-coupling interaction. J-coupling, also known as indirect dipolar coupling, can be present in both solids and liquids. However, it is usually of such a small magnitude that it is completely masked in solid-state spectra by larger anisotropic influences. In solution state, however, it can produce a substantial effect as nuclei can couple together due to the participation of electrons. It is dipolar coupling in that it describes the coupling between nuclear dipole moments, but rather than direct coupling of the nuclei through spatial proximity, it occurs through the participation of local electrons. It is for this reason that many consider J-coupling to be a direct spectral representation of the chemical bond: nuclei that are separated by one or a few chemical bonds will exhibit dipolar coupling. Like the chemical shift, the J-coupling Hamiltonian is a second-rank tensor and depended on molecular orientation. It can be described in tensor form (eq. 1.18) or in scalar form (eq. 1.19):

$$\mathcal{H}_J = 2\pi \mathbf{I}_j \cdot \begin{pmatrix} J_{xx} & J_{xy} & J_{xz} \\ J_{yx} & J_{yy} & J_{yz} \\ J_{zx} & J_{zy} & J_{zz} \end{pmatrix} \cdot \mathbf{I}_k \qquad (1.18)$$

$$\mathcal{H}_J = 2\pi \frac{J_{xx} + J_{yy} + J_{zz}}{3} \mathbf{I}_j \cdot \mathbf{I}_k \qquad (1.19)$$

As previously mentioned, the J-coupling is much smaller compared to the dipolar coupling, the quadrupolar coupling and chemical shielding effects. For this reason, its effects are often ignored in solid-state NMR.

1.2.7 The quadrupole coupling interaction. In addition to the magnetic effects previously described, there exists an electric component of the total Hamiltonian. This Hamiltonian describes the interaction between the electrical quadrupole moment of a nucleus with an electric field gradient. Spherical nuclei (i.e. $I = ½$) have an over-all charge distribution of zero and have no nuclear electronic interaction with an electric field gradient. This means that while the nuclear magnetic (dipole) interaction with the magnetic field is non-zero, there is no quadrupolar interaction. However, non-spherical nuclei (i.e. $I > ½$) have a non-zero overall charge distribution and therefore a non-zero electric quadrupole moment. For completeness, the distribution of electric charge in a nucleus is most effectively denoted as a sum of electric multipoles. The zeroth-order multipole represents the charge, the first-order would be the electic dipole, which is always zero, and the second-order the electric quadrupole. These are functionally similar to the spherical harmonics and resemble the orbital shapes of the hydrogen atom that arise from spherical harmonics.

The principal axis system is chosen so that the electric field gradient (**V**) is a traceless, second-order tensor of the form:

$$\mathbf{V} = \begin{pmatrix} V_{xx} & 0 & 0 \\ 0 & V_{yy} & 0 \\ 0 & 0 & V_{zz} \end{pmatrix} \tag{1.20}$$

As previously stated, the principal values are assigned so that $|V_{zz}| > |V_{xx}| > |V_{yy}|$. At the nucleus, the nuclear quadrupole moment couples to the electric field gradient, resulting in what is called quadrupolar coupling. The full form of the quadrupolar coupling Hamiltonian describes the coupling of spin **I** with the electric field gradient **V** and is of the form:

$$\mathcal{H}_Q = \frac{eQ}{2I(2I-1)\hbar}\mathbf{I}\cdot\mathbf{V}\cdot\mathbf{I} \tag{1.21}$$

In this expression **I** is the vector of the nuclear spin, **V** is the field gradient tensor as described by equation 1.20 , and eQ is the nuclear quadrupole moment of the nucleus. The Hamiltonian can also be described in the principal axis frame as:

$$\mathcal{H}_Q = \frac{e^2qQ}{4I(2I-1)\hbar}\left[3I_{z,PAF}^2 - \mathbf{I}^2 + \frac{1}{2}\eta\left(I_{x,PAF}^2 - I_{y,PAF}^2\right)\right] \tag{1.22}$$

In this expression, $e^2qQ/\hbar$ is known as the quadrupole coupling constant. It is measured in Hz and usually notated as QCC or C_Q. In this expression eq is equal to the V_{zz} component of the electric field gradient tensor in the principal axis system. For this reason it can be seen how the magnitude of the quadrupole coupling constant depends on the field gradient. Furthermore, the shape or, specifically, the splitting in a ^{2}H powder pattern is proportional to the magnitude of the quadrupole coupling constant (Figure 1.2). Because the quadrupolar interaction is significantly larger than the

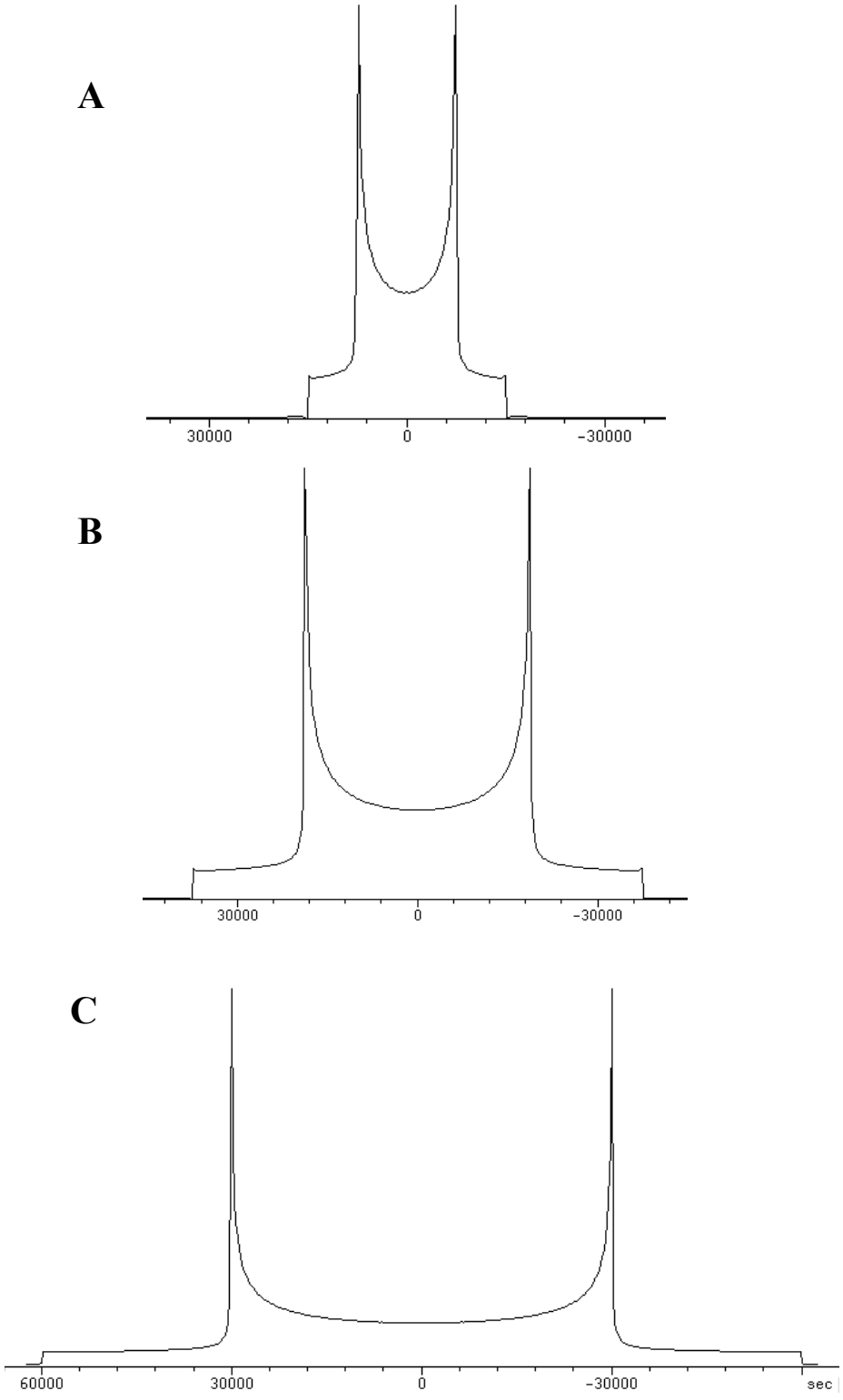

Figure 1.2. Simulated quadrupole echo ^{2}H SSNMR lineshapes for A) C_Q = 20 kHz, B) C_Q = 50 kHz and C) C_Q = 80 kHz. Simulated with SIMPSON 3.1.[23]

chemical shielding anisotropy, the quadrupolar asymmetry parameter can substantially distort the quadrupolar lineshape (figure 1.3). The quadrupolar asymmetry parameter can be defined using the principal components of the electric field gradient.:

$$\eta = \frac{V_{xx} - V_{yy}}{V_{zz}} \tag{1.23}$$

As mentioned, the secular approximation of the quadrupolar Hamiltonian can only be used when the quadrupolar interaction is much smaller than the Zeeman interaction. Since the Zeeman interaction is described primarily in the laboratory frame, it is useful to convert the quadrupolar Hamiltonian from the principal axis frame to the laboratory frame (i.e. $\left(I_{x,PAF}, I_{y,PAF}, I_{z,PAF}\right) \longrightarrow \left(I_x, I_y, I_z\right)$). This is most often achieved by the use of a Wigner rotation matrix, which is used to rotate an object with tensor properties between various frames of reference.[28] In the laboratory frame quadrupolar Hamiltonian, as with the other anisotropic Hamiltonian terms, there arises a dependence on the term ($3\cos^2\theta$ - 1). However, the quadrupolar interaction in anisotropic solids is often the largest internal interaction and cannot be completely averaged out of the spectral lineshape, even with rapid magic angle spinning. It can, however, be partially averaged out of the spectrum, resulting in significantly narrowed spectral frequencies.

1.3 Theoretical Calculations

The theoretical calculation of NMR parameters is an excellent tool to assist in the analysis and comparison of experimental results. In particular, *ab initio* methods

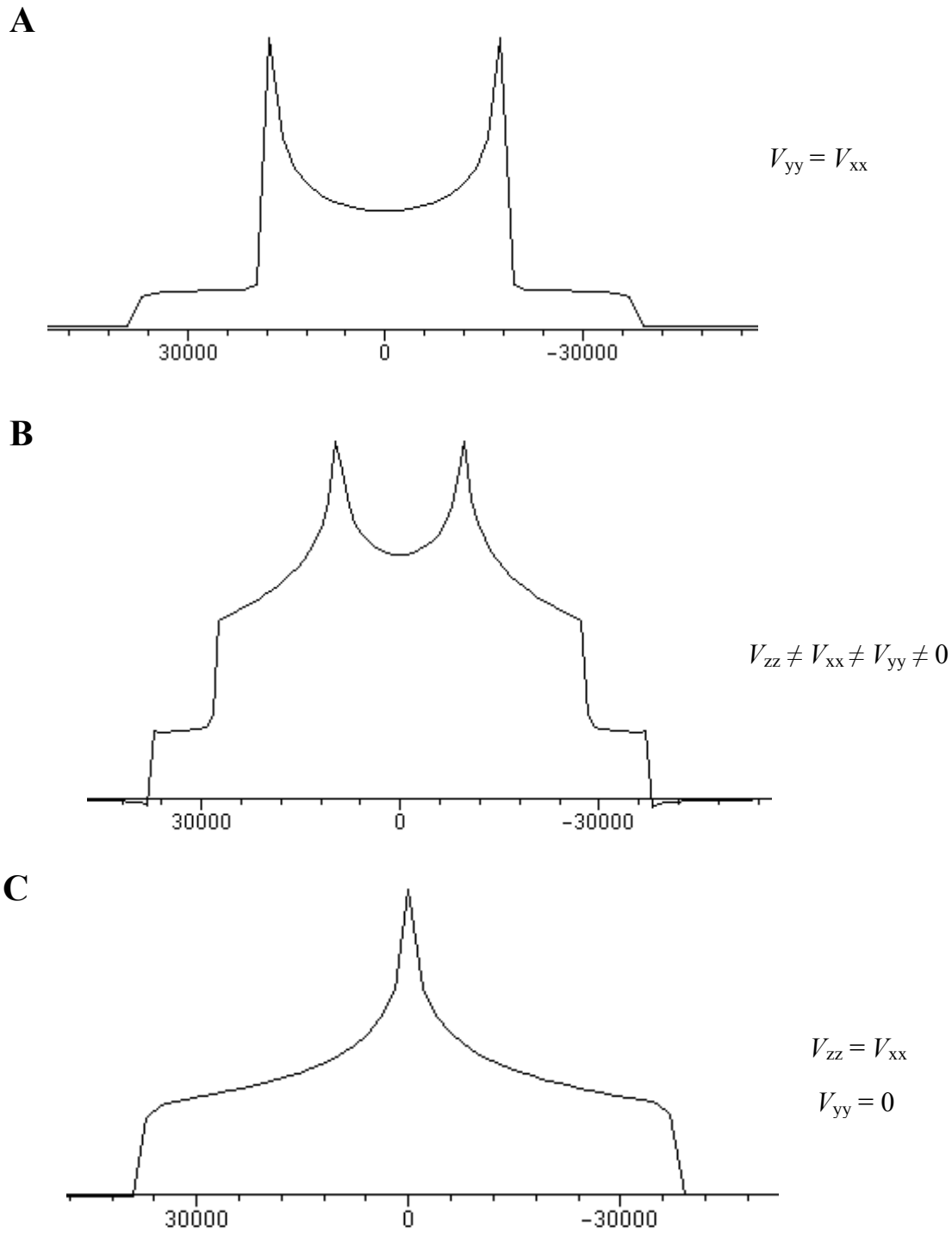

Figure 1.3. Simulated quadrupole echo ^{2}H SSNMR lineshapes for A) $\eta = 0$, B) $\eta = 0.5$ and C) $\eta = 1$. Simulated with SIMPSON 3.1.[23]

are used because they are derived directly from first order principles, without the need of any empirical or experimental information. Electronic and magnetic properties can be easily interpreted empirically for some systems, while other more complicated systems require the prediction of these properties based on computational means. Many modern spectroscopists consider it routine to include magnetic and electronic parameters from *ab initio* calculations as well as empirical results. This will only become more routine as computational techniques are improved. The general computational method is to solve the ground-state wave function and energy using a variety of approximation methods. It has previously been stated that NMR interactions are treated as perturbations of the system. For this reason, magnetic and electronic properties can be calculated as the response of the approximated wave function to the introduction of these perturbations. In this case, the perturbation is the operator associated with the particular NMR interaction.[6-9]

I will use this section of my thesis to describe the various computational methods I used throughout the duration of my studies, including *ab initio* methods and density functional theory (DFT). In particular, I will discuss the various approximation methods as well as the process of calculating the electric field gradient tensor and the chemical shielding tensor.

1.3.1 Hartree-Fock method. The exact solution of the electronic Schrödinger equation for anything other than exceedingly simplified systems is impossible. For this reason, computational methodologies must rely on a variety of approximations to simplify the system. Of central importance is the Born-Oppenheimer approximation in which the coupling between the nuclear and electronic motion is neglected. The Born-Oppenheimer

method is an iterative method in which the electronic wave function is solved independently of the fixed nuclear coordinates. The electronic energy eigenvalue is a function of the chosen nuclear coordinates. A variety of nuclear coordinates are chosen and the electronic Schrödinger equation solved in order to form a potential energy surface that is a function of the nuclear coordinates. This function serves as a potential for the nuclear kinetic energy and the Schrödinger equation for the nuclear motion can be solved. The Schrödinger equation under the Born-Oppenheimer approximation is given by:

$$\left(T_n + E_j(R)\right)\Psi_{nj}(R) = E_{total}\Psi_{nj}(R) \tag{1.24}$$

In this equation, T_n gives the nuclear kinetic energy. The electronic potential energy as a function of nuclear coordinates is given as $E_j(R)$. Solution of this equation for the wave function leads to the molecular energy and the energy levels for molecular vibrations and rotations.[9]

Expanding on the Born-Oppenheimer approximation, D. J. Hartree introduced a new iterative *ab initio* method in 1927 called self-consistent field (SCF).[10] The Hartree-Fock method is an approximation method, necessary because the full electronic wave function is only exactly solvable for the H_2^+ molecule (or other one electron models). Other more complicated systems require approximation methods. The HF method uses the variational theory to provide approximate solutions of the electronic wave function. The first step in the HF method is to make an initial guess of the wave function. The energy associated with the approximate wave function can be calculated by the expectation value of the Hamiltonian.

$$E_e = \frac{\langle \Psi | \mathcal{H}_e | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{1.25}$$

These approximate wave functions are often given the notation ***Φ.*** The total electronic wave function must be antisymmetric with respect to the exchange of any two electrons. The HF method utilizes Slater determinants to construct electronic wave functions to satisfy this antisymmetry condition. The Slater determinants are of the form

$$\Phi_{SD} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_N(1) \\ \phi_1(2) & \phi_1(2) & \dots & \phi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(N) & \phi_2(N) & \cdots & \phi_N(N) \end{vmatrix} \tag{1.26}$$

In this determinant, the columns serve as single electron wave function, or the molecular orbitals, while the rows represent the electron coordinates. We select the trial wave function so that it includes only one Slater determinant. Variational theory can then be used to derive the appropriate HF equations by minimizing the expectation value of the Hamiltonian. If the expectation value does not represent a minimum value of the energy, we choose a new wave function and the process starts again.

The electronic Hamiltonian is the sum of the electronic kinetic energy (***T_e***), the nucleus-electron attraction potential (***V_ne***), the electron-electron repulsion potential (***V_ee***), and the nucleus-nucleus repulsion potential (***V_nn***). The total form of the Hamiltonian is provided below, as well as the expressions for the individual energies.

$$\boldsymbol{H}_e = \boldsymbol{T}_e + \boldsymbol{V}_{ne} + \boldsymbol{V}_{ee} + \boldsymbol{V}_{nn}$$

$$\boldsymbol{T}_e = -\sum_{i}^{N_{elec}} \frac{1}{2} \nabla_i^2$$

$$\boldsymbol{V}_{ne} = -\sum_{a}^{N_{nuc}} \sum_{i}^{N_{elec}} \frac{Z_a}{|R_a - r_i|}$$

$$\boldsymbol{V}_{ee} = \sum_{i}^{N_{elec}} \sum_{j>i}^{N_{elec}} \frac{1}{|r_i - r_j|}$$
$$\boldsymbol{V}_{nn} = -\sum_{a}^{N_{nuc}} \sum_{b>a}^{N_{nuc}} \frac{Z_a Z_b}{|R_a - R_b|} \tag{1.27}$$

The nuclear repulsion term ($\boldsymbol{V}_{nn}$) is independent of the electron coordinates and may be integrated over the trial wave function to yield a constant.

$$\langle \Phi | \boldsymbol{V}_{nn} | \Phi \rangle = V_{nn} \tag{1.28}$$

The one-electron operators ($\boldsymbol{T}_e$ and $\boldsymbol{V}_{ne}$) provides a non-zero contribution to the Slater determinant only for the identity operator. For coordinate (1) of the Slater determinant, this yields a matrix element for orbital (1) only, as shown below.

$$\left\langle \phi_1(1)\phi_2(2)\ldots\phi_N(N) \left| -\frac{1}{2}\nabla_i^2 - \sum_{a}^{N_{nuc}} \frac{Z_a}{|R_a - r_i|} \right| \phi_1(1)\phi_2(2)\ldots\phi_N(N) \right\rangle$$
$$= \left\langle \phi_1(1) \left| -\frac{1}{2}\nabla_1^2 - \sum_{a}^{N_{nuc}} \frac{Z_a}{|R_a - r_1|} \right| \phi_1(1) \right\rangle$$
$$= -\frac{1}{2}\nabla_1^2 - \sum_{a}^{N_{nuc}} \frac{Z_a}{|R_a - r_1|} \tag{1.29}$$

Any element that involves the permutation of electrons with the one-electron operators is zero, as follows:

$$\left\langle \phi_1(1)\phi_2(2)\ldots\phi_N(N) \left| -\frac{1}{2}\nabla_i^2 - \sum_{a}^{N_{nuc}} \frac{Z_a}{|R_a - r_i|} \right| \phi_2(1)\phi_1(2)\ldots\phi_N(N) \right\rangle = 0 \tag{1.30}$$

The two-electron operator ($\mathbf{V}_{ee}$) will be zero for any operator involving the exchange of

three or more electrons. Only the identity operator and the operator responsible for the permutation of two electrons are allowed and will provide a non-zero result. Below I show the two-electron operator with the identity operator:

$$\begin{aligned}
&\left\langle \phi_1(1)\phi_2(2)\ldots\phi_N(N)\left|\frac{1}{r_1-r_2}\right|\phi_1(1)\phi_2(2)\ldots\phi_N(N)\right\rangle \\
&=\left\langle \phi_1(1)\phi_2(2)\left|\frac{1}{r_1-r_2}\right|\phi_1(1)\phi_2(2)\right\rangle\ldots\left\langle \phi_N(N)\middle|\phi_N(N)\right\rangle \\
&=\left\langle \phi_1(1)\phi_2(2)\left|\frac{1}{r_1-r_2}\right|\phi_1(1)\phi_2(2)\right\rangle = J_{12}
\end{aligned} \tag{1.31}$$

This new matrix element, J_{ij}, is known as the Coulomb integral and represents the repulsion between two like charges. The result of the two-electron operator with the permutation operator provides a new matrix element called the exchange integral, K_{ij}.

$$\begin{aligned}
&\left\langle \phi_1(1)\phi_2(2)\ldots\phi_N(N)\left|\frac{1}{r_1-r_2}\right|\phi_2(1)\phi_1(2)\ldots\phi_N(N)\right\rangle \\
&=\left\langle \phi_1(1)\phi_2(2)\left|\frac{1}{r_1-r_2}\right|\phi_2(1)\phi_1(2)\right\rangle\ldots\left\langle \phi_N(N)\middle|\phi_N(N)\right\rangle \\
&=\left\langle \phi_1(1)\phi_2(2)\left|\frac{1}{r_1-r_2}\right|\phi_2(1)\phi_1(2)\right\rangle = K_{12}
\end{aligned} \tag{1.32}$$

Combining equations 1.28–32, we can now write the total energy as a sum of these terms:

$$E=\sum_i^{N_{elec}}\left(-\frac{1}{2}\nabla_1^2-\sum_a^{N_{nuc}}\frac{Z_a}{\left|R_a-r_1\right|}\right)+\sum_{i=1}^{N_{elec}}\sum_{j>i}^{N_{elec}}\left(J_{ij}-K_{ij}\right)+V_{nn} \tag{1.33}$$

Some of the terms in this expression for the energy can be combined to give a new one-electron energy operator, the Fock operator ($\mathbf{F}_i$).

$$\boldsymbol{F}_i=-\frac{1}{2}\nabla_1^2-\sum_a^{N_{nuc}}\frac{Z_a}{\left|R_a-r_1\right|}+\sum_j^{N_{elec}}\left(\boldsymbol{J}_j-\boldsymbol{K}_j\right) \tag{1.34}$$

Using the operator form of the exchange and Coulomb integrals, the Fock operator describes the repulsion between electrons. The first term describes the kinetic energy of the electron while the second term describes the nuclear-electron attraction. The Fock operator may be used with a set of canonical molecular orbitals ($\phi_i^{'}$) to provide a set of eigenvalue equations, according to:

$$\begin{aligned} \boldsymbol{F}_i\phi_i^{'} &= \varepsilon_i\phi_i^{'} \\ \left\langle \phi_i^{'} \middle| \boldsymbol{F} \middle| \phi_i^{'} \right\rangle &= \varepsilon_i \left\langle \phi_i^{'} \middle| \phi_i^{'} \right\rangle = \varepsilon_i \end{aligned} \tag{1.35}$$

In this equation ε_i is the orbital energy of the i-th electron. The set of functions that solve equation 1.35 is known as the self-consistent field (SCF) orbitals.

The exact functional form of the molecular orbitals is difficult to determine. For this reason, theoretical studies use the basis set approximation to express these unknown molecular orbital functions as a set of known functions. The only limit to the type of function chosen is that the form of the known function should fit the physical problem at hand: in the case of molecular and atomic orbitals, the functions should go to zero as the distance between the electron and the nucleus is maximized. Gaussian functions are easy to integrate and their functional form fits the physical nature of the problem. A molecular orbital can be described as a linear combination of atomic orbitals, where ϕ_i represents the molecular orbital, N_{basis} is the set of basis functions, $c_{\alpha i}$ are the coefficients of the wave function and χ_α represents the set of basis functions.

$$\phi_i = \sum_{\alpha}^{N_{basis}} c_{\alpha i}\chi_\alpha \tag{1.36}$$

Using this new molecular orbital as the linear combination of atomic orbitals, the Hartree-Fock equation (Equation 1.35) can now be rewritten as:

$$\mathbf{F}_i \sum_{\alpha}^{N_{basis}} c_{\alpha i}\chi_\alpha = \varepsilon_i \sum_{\alpha}^{N_{basis}} c_{\alpha i}\chi_\alpha \tag{1.37}$$

Equation 1.36 can be collected into a set of matrices where **F** is the Fock matrix, **S** is the overlap matrix containing the overlap elements between the basis functions, and the **C** matrix contains the coefficients, such that:

$$\mathbf{FC} = \mathbf{SC}\varepsilon$$

$$F_{\alpha\beta} = \langle \chi_\alpha | \mathbf{F} | \chi_\beta \rangle$$

$$S_{\alpha\beta} = \langle \chi_\alpha | \chi_\beta \rangle \tag{1.38}$$

Each element in the Fock matrix ($F_{\alpha\beta}$) contains the one-electron operator (eq. 1.29) as well as a term representing the sum over all of the coefficients of the occupied molecular orbitals multiplied by the two-electron repulsion integral. The later term is most commonly written as:

$$\sum_{j>i}^{N_{elec}} \int \frac{|\Psi_j^2|}{|r_i - r_j|} \tag{1.39}$$

where the $|\Psi_j^2|$ term is known as the density matrix. As mentioned, the SCF method is iterative in nature. This process starts by calculating the one and two electron integrals. Next, an initial guess of the wave functions is made to obtain a density matrix and a guess of the molecular orbital coefficients. These initial guesses are used to form and diagonalize the Fock matrix whose eigenvectors are the new set of molecular orbital coefficients. This new set of coefficients is then used to calculate a new density matrix. This iterative process continues until the density matrix formed is sufficiently similar to the original guess of the wave function.[9,11,28]

1.3.2 Møller-Plesset 2nd order method (MP2). The Møller-Plesset[12] method was developed in 1934 as a perturbational method of the Hartree-Fock method. By adding in the effects of electronic correlation, it serves to improve upon the Hartree-Fock method. MP2 method uses the original HF Hamiltonian (Equation 1.29) as the zeroth-order, unperturbed Hamiltonian.

$$\mathcal{H}_0 = \sum_i^{N_{elec}} \mathbf{F}_i = \sum_i^{N_{elec}} \left(-\frac{1}{2}\nabla_i^2 + \sum_j^{N_{elec}} \left(\boldsymbol{J}_j - \boldsymbol{K}_j \right) \right) \tag{1.40}$$

The unperturbed, zeroth-order wave functions are therefore the eigenfunctions of the unperturbed Hamiltonian. The corresponding zeroth-order energy (E_0) is the sum of all the molecular orbital energies, so that:

$$\mathcal{H}_0 \Psi_0 = \left(\sum_i^{N_{elec}} \varepsilon_i \right) \Psi_0 \tag{1.41}$$

Summing over the Fock operators as shown in equation 1.41 counts the electron-electron repulsion term twice, so that perturbation ($\mathcal{H}'$) to the unperturbed Hamiltonian becomes:

$$\mathcal{H}' = \mathcal{H} - \mathcal{H}_0 = \sum_i^{N_{elec}} \sum_{j>i}^{N_{elec}} \frac{1}{|r_i - r_j|} - 2\left\langle \sum_i^{N_{elec}} \sum_{j>i}^{N_{elec}} \frac{1}{|r_i - r_j|} \right\rangle \tag{1.42}$$

The first-order correction to the ground-state energy ($E_0^{(1)}$) is then,

$$\begin{aligned} E_0^{(1)} &= \left\langle \Psi_0 \left| \mathcal{H}' \right| \Psi_0 \right\rangle \\ &= \left\langle \Psi_0 \left| \sum_{j>i}^{N_{elec}} \frac{1}{|r_i - r_j|} \right| \Psi_0 \right\rangle - \left\langle \Psi_0 \left| \sum_{j\neq i}^{N_{elec}} \left(\boldsymbol{J}_j(i) - \boldsymbol{K}_j(i) \right) \right| \Psi_0 \right\rangle \\ &= -\frac{1}{2} \sum_{ij}^{N_{elec}} \langle ij || kl \rangle \end{aligned} \tag{1.43}$$

The double bar notation is:

$$\langle ij||kl\rangle = \langle ij|kl\rangle - \langle ij|lk\rangle \tag{1.44}$$

where

$$\langle ij|kl\rangle = \left\langle \Psi_i(1)\Psi_j(2)\left|\frac{1}{|r_i - r_j|}\right|\Psi_k(1)\Psi_l(2)\right\rangle \tag{1.45}$$

The total HF energy is now the sum of the zeroth-order and the first-order correction

$$E_{HF} = E_0^{(0)} + E_0^{(1)} = \sum_i^{N_{elec}} \varepsilon_i - \frac{1}{2}\sum_{ij}^{N_{elec}} \langle ij||kl\rangle \tag{1.46}$$

The second-order correction to the HF energy can now be determined by:

$$\begin{aligned} E_0^{(2)} &= \sum_{n\neq 0} \frac{\left|\left\langle \Psi_0 \left| \mathcal{H}' \right| \Psi_n^{(0)} \right\rangle\right|^2}{E_0^{(0)} - E_n^{(0)}} \\ &= \sum_{j<i}\sum_{s>r} \frac{\left|\left\langle \Psi_0 \left| \sum_{j>i}^{N_{elec}} \frac{1}{|r_i - r_j|} - \sum_{j\neq i}^{N_{elec}} \left(\boldsymbol{J}_j(i) - \boldsymbol{K}_j(i)\right) \right| \Psi_{ij}^{rs} \right\rangle\right|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_r - \varepsilon_s} \\ &= \sum_{j<i}\sum_{s>r} \frac{\left|\langle ij||rs\rangle\right|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_r - \varepsilon_s} \end{aligned} \tag{1.47}$$

so that the total MP2 molecular energy is the sum of the zeroth-order energy and the first- and second-order corrections:

$$E_{MP2} = E_0^{(0)} + E_0^{(1)} + E_0^{(2)} \tag{1.48}$$

1.3.3 Density functional theory (DFT). Density function theory (DFT) is a pseudo-empirical alternative to *ab initio* methods. It was developed from an idea originally proposed by Hohenberg and Kohn[7] that suggested that the ground state electronic energy, wave function, and molecular properties can be determined completely from the electron density (ρ) rather than the wave function. The complexity of a wave function increases exponentially as the system's complexity increases (i.e. increase in number of electrons). The number of variables needed to describe the electron density is independent of the number of electrons. One major drawback in DFT is that each density produces a different ground state energy, and the functionals connecting the density to the energy are unknown. For this reason, a fundamental goal in DFT method is to design functionals that coordinate the electron density to the ground state energy. Using this notation, a wave function or the electron density can be considered functions while the energy is a functional of the electron density. Functions are dependent on variables and denoted as $f(x)$. Functionals are dependent on functions and are denoted as $F[f]$ so that

$$E = E[\rho] \tag{1.49}$$

Early efforts in DFT were fruitless until 1965 when Kohn and Sham[8] suggested the use of an idealized reference system of orbitals with non-interacting electrons to calculate the electron kinetic energy. This system of orbitals represents the total approximate electron density by:

$$\rho_{aprrox} = \sum_{i}^{N_{elec}} |\phi_i|^2 \tag{1.50}$$

The DFT method is similar to the HF method both conceptually and computationally. Similar to the HF method, the total electronic energy functional will be a function of the

kinetic energy ($T[\rho]$), the nuclear-electron attraction ($E_{ne}[\rho]$), and the repulsion between any two electrons ($E_{ee}[\rho]$). As with the HF method, the electron-electron repulsion term can be expressed as a function of the exchange ($K[\rho]$) and Coulomb ($J[\rho]$) functionals. Fundamental to the KS theory is that the electrons are non-interacting so that the total kinetic energy is now simplified to:

$$T_{KS} = -\sum_{i}^{N_{elec}} \left\langle \phi_i \left| \frac{1}{2} \nabla_i^2 \right| \varphi_i \right\rangle \tag{1.51}$$

Physically, however, we know that the electrons are interacting. So while this equation does not provide the exact kinetic energy, the difference between the approximated kinetic energy and the real kinetic energy is small. This small difference is accounted for by the use of an exchange-correlation term ($E_{xc}[\rho]$).

$$E_{xc} = \left(T_{exact}[\rho] - T_{KS}[\rho]\right) + \left(E_{ee}[\rho] - J[\rho]\right) \tag{1.52}$$

The total DFT energy can now be written as:

$$E_{DFT}[\rho] = T_{KS}[\rho] + E_{ne}[\rho] + E_{xc}[\rho] + J[\rho] \tag{1.53}$$

However, the exact $E_{xc}[\rho]$ is unknown and the DFT method seeks to determine a decent approximation of E_{xc}. There are many different exchange-correlation functionals used in DFT. The most common and the one used exclusively in this research is the B3LYP functional. The total exchange-correlation functional can be described by the sum of a separate exchange and correlation functional.

$$E_{xc}[\rho] = E_x[\rho] + E_c[\rho] \tag{1.54}$$

Hybrid exchange-correlation functionals are commonly used.[13-18] These functionals are combinations of various exchange ($E_x[\rho]$) and correlation ($E_c[\rho]$) functionals. B3LYP is a hybrid exchange-correlation functional, which combines the Local Spin Density Approximation exchange and correlation functionals (LSDA), the exact Hartree-Fock exchange functional, Becke's[17] 1988 exchange functional (B88), and the Lee-Yang-Parr[13] correlation functional (LYP) by:

$$E_{xc}^{B3LYP} = (1-a)E_x^{LSDA} + aE_x^{HF} + b\Delta E_x^{B88} + (1-c)E_c^{LSDA} + cE_c^{LYP} \qquad (1.55)$$

In this equation the a, b, and c parameters are empirical parameters that have been determined from a fit to the experimental data. Typically, $a \sim 0.2$, $b \sim 0.7$, and $c \sim 0.8$.

1.3.4 Calculation of NMR parameters: Gauge-included atomic orbitals (GIAO). Modern NMR spectroscopists will use computational methods to calculate chemical shielding tensors and electric field gradients. These can be manipulated to produce theoretical chemical shifts and quadrupole coupling constants to coincide with those parameters obtained experimentally. Essentially, the calculation of these magnetic parameters is the quantum theory of an atom in a magnetic field. The calculation of magnetic properties is complicated by the presence of a gauge origin in the operators, which suggests that the corresponding expectation value might be origin dependent. Solution of the complete Schrödinger equation from an exact wave function will produce origin-independent results. However, the gauge origin becomes an issue when using a finite basis set of approximate wave functions. For basis functions that are atom-centered

in nature, the use of gauge-including atomic orbitals (GIAO) has become a preferred solution.[6]

For an electromagnetic field, the magnetic field can be given by equation 1.56

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{1.56}$$

and the electric field as

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \tag{1.57}$$

A in these equations represents a vector potential with an arbitrary origin and ϕ the scalar electric potential. The vector potential (**A**) is often chosen as

$$\mathbf{A} = \frac{1}{2}\mathbf{B} \times \left(\mathbf{r} - \mathbf{R}_{ORG}\right) \tag{1.58}$$

In this equation, $\mathbf{R}_{ORG}$ is known as the gauge origin. If we denote the zeroth-order Hamiltonian ($\mathcal{H}^{(0)}$) as the Hamiltonian in the absence of any magnetic field, the magnetic field-dependent Hamiltonian will be a function of the zeroth-order Hamiltonian, the electron momentum operator (equation 1.59), and the vector potential describing the electromagnetic field, as shown in equation 1.60.

$$\boldsymbol{p} = -i\hbar\nabla + e\mathbf{A} \tag{1.59}$$

$$\mathcal{H}_{EMF} = \mathcal{H}^{(0)} + \mathbf{A}\boldsymbol{p} + \frac{1}{2}\mathbf{A}^2 \tag{1.60}$$

Referring back to the arbitrarily chosen equation for the vector potential of the electromagnetic field (**A**), we can obtain the magnetic field-dependent Hamiltonian

$$\mathcal{H}_{EMF} = \mathcal{H}^{(0)} + \frac{1}{2}\left(\left(\mathbf{r} - \mathbf{R}_{ORG}\right) \times \boldsymbol{p}\right)\mathbf{B} + \frac{1}{8}\left(\mathbf{B} \times \left(\mathbf{r} - \mathbf{R}_{ORG}\right)\right)^2 \tag{1.61}$$

As expected, this expression not only depends on the magnitude of the magnetic field but also on our choice of the gauge origin ($\mathbf{R}_{ORG}$). As mentioned, magnetic parameters calculated with an exact wave function are gauge invariant. Therefore, calculations done with very large basis sets can produce results that are considered approximately gauge invariant. Unfortunately, with large electron systems or higher levels of theory, the use of very large basis sets is not a feasible option. For magnetic shielding calculations, Ditchfield[20] proposed the use of local gauge origins to define the vector potentials. In this method, called the gauge invariant/including atomic orbital method (GIAO), a local gauge origin was placed on the center of each atomic orbital. In this method, the gauge-invariant atomic orbitals (χ_v) are defined as:

$$\chi_v = \phi_v \exp\left[-\frac{ie}{\hbar c}\mathbf{A}_i \cdot \mathbf{r}\right] \tag{1.62}$$

Where ϕ_v represents the real atomic orbital and i is the center of each orbital. The gauge factor is then determined by the vector potential that, as denoted by its subscript i, is also at the same center. The molecular orbital can now be written as:

$$\psi' = \psi \exp\left[-\frac{ie}{\hbar c}\sum_i \mathbf{A}_i \cdot \mathbf{r}\right] \tag{1.63}$$

and the Schrödinger equation changes to

$$\mathcal{H}\psi' = E\psi' \tag{1.64}$$

In this new expression, *E* is independent of the gauge origin. It therefore follows that all other expectation values would also be independent of gauge origin, including magnetic properties and other observable properties. The GIAO method has since been expanded to other computational methods by applying gauge factors to molecular orbitals rather than each atomic orbital. These include Schindler and Kutzelnigg's individuals gauge localized orbital (IGLO)[29] and localized orbital/local origin (LORG)[30] method of Hansen and Bouman.

References

1. Pauli, W. Scientific *Correspondence with Bohr, Einstein, Heisenberd and others*, Vol. 1: 1919-1929, eds. A. Hermann, K. von Meyenn and V. F. Weisskopf. Springer-Verlag, New York, **1979**.
2. Rabi, I. I.; Zacharias, J.R.; Millman, S.; Kusch, P. *Physical Review*, **1938**, *53*, 318.
3. Purcell, E. M.; Torrey, H. C.; Pound, R. V. *Physical Review*, **1946**, *69*, 37-38.
4. Bloch, F.; Hansen, W. W.; Packard, M. *Physical Review*, **1946**, *69*, 127.
5. Levine, I. N. *Quantum Chemistry*, Prentice Hall, **2000**.
6. *Calculation of NMR and EPR Parameters: Theory and Applications*; Kaupp, M., Bu¨hl, M., Malkin, V. G., Eds.; Wiley-VCH, **2004**.
7. Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136,* B864–B871.
8. Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.
9. Jensen, Frank. Introduction to Computational Chemistry. Wiley, 2nd ed. **2007**.
10. Malkin, V. G.; Malkina, 0. L.; Erikson, L. A.; Salahub, D. R. *Theoretical and Computational Chemistry*. Vol 1. *In Density Functional Calculations*.
10. Hartree, D.R. *Mathematical Proceedings of the Cambridge Philosophical Society*. **1928**, *24*, 426.
11. Roothaan, C.C.J. *Reviews of Modern Physics* **1951**, *23*, 69.
12. Møller, C., Plesset, M.S. *Physical Review* **1934**, *46*, 618.
13. Parr, R.G.; Yang, W. *Density Functional Theory of Atoms and Molecules*. Oxford Univeristy Press, Oxford, New York, **1989**.
14. Ziegler, T. *Chem. Rev*., 1991, *91*, 651.
15. Friedrich, K.; Seifert, G.; Grossmann, G. *Z . Phys. D*. **1990**, *17*, 45.
16. Vosko, S. H.; Wilk, L.; Nusair, M. *Canadian Journal of Physics*. **1980**, *58*, 1200–1211.
17. Becke, A.B. *Physical Review A*. **1988**, *38*, 3098.
18. Lee, C.; Yang, W.; Parr, R.G. *Phys. Rev. B.* **1988**, *37*, 785.
19. Slater, J.C. *Physical Review*. **1951**, *81*, 385.
20. Ditchfield, R. *Molecular Physics*. **1974**, *27*, 789-807.
21. Epstein, Saul. *J. Chem. Phys.* **1965**, *42*, 2897.

22. Schreckenback, G.; Ziegler, T. *J. Phys. Chem.* **1995**, *99*, 606-611. [DFT]

23. Tosner, Z., Andersen, R., Nielsen, N.C. and Thomas Vosegaard. SIMPSON version 3.1, March 25, **2011**.

24. Atkins, P. and Friedman, R. *Molecular Quantum Mechanics*. Oxford Publishing, 4th ed., **2005**.

25. Duer, M. J. *Introduction to Solid-State NMR Spectrscopy*. Blackwell Publishing, **2004**.

26. Levitt, M.H. *Spin Dynamics: Basics of Nuclear Magnetic Resonance*. Wiley Scientific, **2001**.

27. Alam, T.M. and Drobny, G.P. *Chem. Rev.* **1991**, *91*, 1545-1590.

28. Wu, X. *Dissertation*, University of Nebraska, **2011.**

29. Schindler, M and Kutzelnigg, W. *J Chem. Phys*. **1982**, *76*, 1919-1933.

30. Hansen, A.E. and Bouman, T.D. *J. Chem. Phys*. **1985**, *82*, 5035-5047.

Chapter 2

Solid-State Deuterium NMR Analysis of Furanose Ring Dynamics in DNA Containing Uracil

Summary

DNA damage has been implicated in numerous human diseases, particularly cancer and the aging process. Single base lesions, such as uracil, in DNA can be cytotoxic or mutagenic, and are recognized by a DNA glycosylase during the process of base excision repair (BER). Increased dynamic properties in lesion-containing DNAs have been suggested to assist recognition and specificity. Deuterium solid-state nuclear magnetic resonance (SSNMR) has been used to directly observe local dynamics of the furanose ring within a uracil:adenine (U:A) base pair and compare to a normal thymine:adenine (T:A) base pair. Quadrupole echo lineshapes, $<T_{1Z}>$, and $<T_{2e}>$ relaxation data were collected and computer modeling was performed. The results indicate that the relaxation times are identical within the experimental error, the solid lineshapes are essentially indistinguishable above the noise level, and our lineshapes are best fit with a model that does not have significant local motions. Therefore, U:A base pair furanose rings appear to have essentially identical dynamic properties as a normal T:A base pair, and the local dynamics of the furanose ring are unlikely to be the sole arbiter for uracil recognition and specificity in U:A base pairs.

2.1 Introduction

The integrity of DNA structure must be maintained for proper cellular function and viability of the genome. Numerous processes create single base lesions and mismatches in DNA including alkylation, oxidation, deamination and incorporation of improper nucleotides. An estimated 10^4-10^6 DNA damage events occur per cell per day in humans.[1] One particularly common error is the presence of uracil (U) in DNA that arises either due to misincorporation of dUMP/dUTP or deamination of cytosine creating a cytosine to thymine (C → T) transition mutation.[2] Deamination produces the mutagenic U:G base pairs while improper dUMP/dUTP incorporation produces potentially cytotoxic U:A base pairs.[2]

Removal of uracil involves the base excision repair (BER) pathway, a primary repair mechanism of single base lesions in DNA.[1, 3-8] The initial steps of BER, performed in this case by uracil DNA glycosylase (UNG), consist of recognition of the damage site, formation of a specific complex, and removal of the lesion. During the removal step, the lesion-containing nucleotide is rotated out of the DNA helix and stabilized in the binding pocket of the glycosylase in a process known as nucleotide flipping or base flipping. Nucleotide flipping is a common motif in DNA-protein interactions,[9] used by glycosylases, methyltransferases, glucosyltransferases, and photolyases. In particular, UNG has been shown to flip its target site,[10] and deoxyuridine is a substrate in both U:G and U:A base pairing contexts.[2]

Many details of BER have been elucidated, but the exact mechanism of lesion recognition in the presence of a vast excess of normal DNA bases remains elusive. The lesions are often similar structurally to normal DNA bases, suggesting that direct

recognition of the damaged base itself may not play a primary or singular role. Indeed, structural studies of free DNA containing single-base lesions have shown varying degrees of deviation from canonical B-form DNA.[11] Few structural studies comparing undamaged dsDNA to uracil- containing dsDNA have been performed. Delort and co-workers found a modest change in the N-glycosidic bond torsion angle as the only alteration due to the presence of a uracil replacing a thymine.[12] Several NMR studies with uracil-containing hairpins have been performed[11, 13] revealing more substantial local structural changes. Raman spectroscopy has shown that uracil-containing dsDNA and normal dsDNA are essentially indistinguishable.[14] These data seem to suggest that a static structural component may not play a significant role in the recognition of uracil by UNG, given the tremendous specificity that UNG has for its substrate. However, other properties in uracil-containing DNA have been shown to differ, in particular thermodynamic parameters vary significantly in DNA containing uracil in various base pairing contexts.[15]

It has been proposed that a dynamic component or flexibility of the lesion nucleotide plays a significant role in the biomolecular recognition process of DNA lesions,[16] and recognition of the local flexibility of the lesion site is likely the first step to identification by the glycosylase.[17] Solution NMR techniques have been used to probe base-pair opening dynamics from imino exchange.[18-19] Wild-type UNG and a series of mutants with single residue active site mutations were expressed and their effect on the imino exchange rate determined. These studies concluded that mutant UNG increases the lifetime of open T:A base pairs, accelerates the imino exchange of the thymine and suggests that UNG takes advantage of increased base pair opening dynamics in T:A and

U:A base pairs. Additionally, DNA containing artificially constrained sugars have shown to change UNG activity due to non-ideal rigid sugar conformation.[20] A dynamic sugar ring alleviates resistance to conformational rearrangement, facilitating the necessary structure for catalysis.

Deuterium (^{2}H) solid-state NMR (SSNMR) has been used to directly examine the dynamic properties in the furanose ring of DNA containing uracil (U:A base pair), as compared to a normal (T:A) base pair (Figure 1). Previous ^{2}H SSNMR investigations have studied the sequence-dependent dynamics within the binding site of the *HhaI* methyltransferase,[21-25] an enzyme which also uses nucleotide flipping during its methylation process.[26] These SSNMR results indicate the furanose ring and backbone methylene groups of the target deoxycytidine residue have the greatest dynamic properties (ring puckering amplitude, rate, and lowest energetic barriers to conformational rearrangement) and it was suggested that these properties contribute to either recognition of the target dC or nucleotide flipping, or both.

DNA sequences containing [2" – ^{2}H] deuterated derivatives of thymidine and deoxyuridine (Figure 2.1B) were synthesized and incorporated into 12 bp oligonucleotides (Figure 2.1A). Each solid DNA sample was hydrated to 10 ± 1 waters per nucleotide. Under these conditions, local hydration is substantially complete so local dynamics are essentially solution-like; introduction of additional waters only adds to bulk hydration, facilitating global tumbling around the helical axis and negating the advantages of the SSNMR technique.

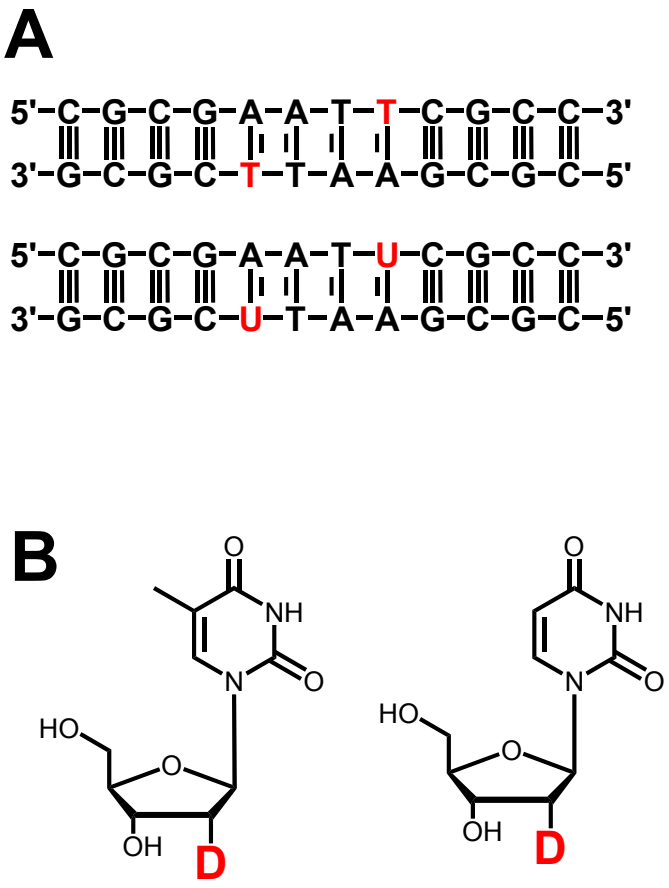

Figure 2.1. (A) Palindromic DNA sequences used in this study. The red indicates the residues containing the [2" – ^{2}H] label. (B) Location of the deuterium label on the thymidine (left) and deoxyuridine (right) residues indicated in red in part A.

This work addresses the question of whether dynamic properties of the lesion-containing DNA furanose ring are the sole or primary property that facilitates recognition of uracil by UNG in U:A base pair contexts. We use deuterium SSNMR as it represents a powerful technique for directly observing dynamic properties in macromolecules. Quadrupole echo lineshape, $<T_{1Z}>$, and $<T_{2e}>$ relaxation data were collected for both sequences in Figure 2.1A, and computer modeling was performed. The results indicate that the relaxation times are identical within the experimental error, the solid lineshapes are essentially indistinguishable above the noise level, and our lineshapes are best fit with a model that does not have significant local motions. Therefore, U:A base pair furanose rings appear to have essentially identical dynamic properties as a normal T:A base pair, and the local dynamics of the furanose ring may not play an important role for uracil recognition in U:A base pairs.

2.2 Theory

The theory of general internal molecular dynamics as probed by deuterium SSNMR is well-documented.[39-40] The fundamental theoretical relationships used to interpret the spectroscopic data describing the internal motions of DNA subunits, in particular the furanose ring, have been described.[23, 34] A dynamically modulated deuterium NMR lineshape is obtained via a quadrupolar echo experiment [$(\pi/2)_x - \tau - (\pi/2)_y - \tau$ – acquire)]. The response of the deuterium spin system starting at the peak of the echo is governed by the equation of motion:

$$\frac{dm(t)_{\pm}}{dt} = A_{\pm} m(t)_{\pm} \tag{2.1}$$

where m(t) is the time domain response and

$$A_{\pm} = i\omega_{\pm} + \pi \tag{2.2}$$

where ± labels either the transition from $m = -1$ to $m = 0$ or from $m = 0$ to $m = +1$, π is a matrix composed of site exchange rates, and $\omega_{\pm}$ is a diagonal matrix with non-zero elements that are the orientation dependent frequencies

$$\omega_{\pm}^{i} = \pm\frac{3}{4}\frac{e^2qQ}{\mathrm{h}}\sum_{a=-2}^{+2} D_{0,a}^{(2)}\left(\Omega_i^{PC}\right) D_{a,0}^{(2)}\left(\Omega^{CL}\right) \tag{2.3}$$

where the super/subscript i denotes the i^{th} of N structural sites. The crystal-fixed frame of the i^{th} site is related to the principal axis system of the EFG tensor by the solid angle $\Omega_i^{PC} = \left(0, \theta_i, \phi_i\right)$. As the sample is assumed to be polycrystalline an additional solid angle $\Omega^{CL} = \left(\phi_{CL}, \theta_{CL}, 0\right)$ relates the crystal-fixed frame to the laboratory-fixed frame.

The effect of the second 90° pulse is to reverse quadrupolar precession. Therefore if detection of the free induction decay is initiated at the top of the echo, the solution of equation 2.1 is:

$$m_{\pm}(t) = \vec{1} \cdot e^{A_{\pm}(t+\tau_2)} \cdot e^{A^*_{\pm}\tau_1} \cdot \vec{m}_0 \tag{2.4}$$

where $A^*_{\pm} = -i\omega_{\pm} + \pi$. If the matrices T and T* diagonalize A and A* respectively, i.e. $T^{-1}AT = \lambda$ and $(T^*)^{-1} A^* T^* = \lambda^*$, the time domain signal may be expanded as

$$\begin{aligned} m(t) &= \vec{1} \cdot T e^{\lambda(t+\tau_2)} T^{-1} \cdot T^* e^{\lambda^* \tau_1} (T^*) \cdot \vec{m}_0 \\ &= \sum_{j,k,l,m,n} T_{jk} e^{\lambda(t+\tau_2)} T^{-1}_{kl} T^*_{lm} e^{\lambda^*_m \tau_1} (T^*)^{-1}_{mn} m_{0,n} \\ &= \sum_k b_k e^{\lambda_k t} \end{aligned} \tag{2.5}$$

The orientation-dependent lineshape is just the Fourier transform of equation 2.5

$$I(\omega, \tau_1, \tau_2, \Omega_{CL}) = \mathrm{Re} \int_{-\infty}^{+\infty} m(t) e^{-i\omega t} dt = \mathrm{Re} \sum_k \frac{b_k}{\lambda_k - i\omega} \tag{2.6}$$

In a polycrystalline sample the lineshape is averaged over all orientations:

$$I(\omega, \tau_1, \tau_2) = \int_0^{2\pi} d\phi_{CL} \int_0^{\pi} d\theta_{CL} \sin\theta_{CL} I(\omega, \tau_1, \tau_2, \phi_{CL}, \theta_{CL}) \tag{2.7}$$

Equation 2.7 is the theoretical form for the motionally modulated deuterium lineshapes presented in this paper. For kinetic models involving jumps between discrete sites equation 2.7 is evaluated numerically using the program MXET1,[40] which also allows calculation of lineshapes modulated by motions around multiple axes. In helical DNA at hydration levels of $W = 10$ and higher, simulations include a uniform rotation

around the long helical axis[41] in addition to strictly local motions of the furanose rings, or other subunits such as the backbone C5'-methylene. A second motional axis would require an additional transformation in equation 2.3 representing the motion around the helical axis.

T_{2e} relaxation is determined by plotting the decay of the signal intensity as a function of quadrupole echo delay. This value is most sensitive to slow motion, on the order of microseconds to milliseconds. For short oligonucletoides, this analysis probes the longitudinal motion about the helical axis and possible slower motions of the furanose ring, which might include pucker inversions. Powder averaged spin-lattice relaxation is sensitive to motions on the nanosecond to microsecond timescales. For short oligonucleotides, this type of timescale will be most sensitive to differences in local motions of the appropriate subunits, which in this case is the furanose ring. Spin lattice relaxation times reported in this paper are obtained by simple saturation recovery. The partly recovered orientation dependent lineshape is, as a function of the recovery time t_r:

$$I\left(\omega,\tau_1,\tau_2,t_r,\Omega_{CL}\right)=(1-e^{-t_r/T_1})I\left(\omega,\tau_1,\tau_2,\infty,\Omega_{CL}\right) \tag{2.8}$$

In equation (8) the spin lattice relaxation time for deuterium nuclei is

$$\frac{1}{T_1}=\frac{3}{16}\left(\frac{e^2qQ}{\mathrm{h}}\right)^2\left(J_1\left(\omega,\theta_{CL},\varphi_{CL}\right)+4J_2\left(2\omega,\theta_{CL},\varphi_{CL}\right)\right) \tag{2.9}$$

where

$$J_q\left(q\omega\right)=\mathrm{Re}\int_{-\infty}^{+\infty}C_q\left(t\right)e^{-i\omega t}dt \tag{2.10}$$

In equation 2.10 $C_q(t)$ is the autocorrelation function defined as:

$$C_q\left(t\right)=\sum_{p,p'}D_{pq}^{(2)*}\left(\Omega_{CL}\right)D_{p'q}^{(2)}\left(\Omega_{CL}\right)\left\langle D_{0p}^{(2)*}\left(\Omega_{PC}\left(0\right)\right)D_{0p'}^{(2)}\left(\Omega_{PC}\left(t\right)\right)\right\rangle \tag{2.11}$$

where

$$\left\langle D_{0p}^{(2)*}\left(\Omega_{PC}(0)\right) D_{0p'}^{(2)}\left(\Omega_{PC}(t)\right)\right\rangle = \int d\Omega_{PC}(0) \int d\Omega_{PC}(t) D_{0p}^{(2)*}\left(\Omega_{PC}(0)\right) D_{0p'}^{(2)}\left(\Omega_{PC}(t)\right) P\left(\Omega_{PC}(t), t \middle| \Omega_{PC}(0)\right) W\left(\Omega_{PC}(0)\right) \tag{2.12}$$

$P\left(\Omega_{PC}(t), t \middle| \Omega_{PC}(0)\right)$ is the conditional probability of the C-D vector reorienting to $\Omega_{PC}(t)$ from $\Omega_{PC}(0)$ at a time *t* earlier, and $W\left(\Omega_{PC}(0)\right)$ is the a priori probability of the bond being oriented at $\Omega_{PC}(0)$.

2.2.1 Models of DNA Helix Motion. In addition to simulating the motion of the particular sugar rings in question, it was necessary to resolve other types of motion that were present in the sample. For hydration levels as low as $W = 10$, collective bending and torsional motions can be neglected for short DNAs. Also, uniform end-to-end tumbling can be neglected, as this type of motion is restricted in the solid-state, even with a sample of intermediate hydration ($10 < W < 20$). Previous work has shown that uniform rotation of the DNA around the helix axis occurs at $W = 10$ and above, and is effectively simulated by a six-site jump,[41] with a half angle of $\theta = 20°$ (orientation of the local dynamic axis of the C2'-^{2}H2" bond with respect to the longitudinal helix axis), values of $\phi = 0°, 60°, 120°, 180°, 240°$, and $300°$ for the six sites, and a rate constant of $k \approx 1.0 \times 10^4$ Hz. Use of these parameters for the overall helix motion has produced good agreement in previous work for several different DNA samples with different types of local motions occurring. These parameters are generally considered well-determined and remain constant for the simulations of the local motions. The resulting simulated spectra are superpositions of the uniform helical rotation and the local motion of the furanose ring.

2.2.2 Models of Furanose Ring Motion. To simulate the deuterium lineshapes for 2″ deuterons, a modeling framework previously successful in simulating furanose lineshapes has been utilized. The cyclic nature of the furanose ring reduces the number of independent geometrical parameters required to describe the displacement of the C-^{2}H bond. The basic structural features arise from the Herzyk and Rabczenko model[42] where the Cartesian coordinates of the j^{th} heavy atom in the furanose ring are:

$$x_j = r_j\left[\sin\alpha_j \times \sin 2\left(\phi + \frac{4\pi j}{5}\right) - \cos\alpha_j \times \cos 2\left(\phi + \frac{4\pi j}{5}\right) + R_j \cos\alpha_j\right]$$
$$y_j = -r_j\left[\cos\alpha_j \times \sin 2\left(\varphi + \frac{4\pi j}{5}\right) - \sin\alpha_j \times \cos 2\left(\phi + \frac{4\pi j}{5}\right) + R_j \sin\alpha_j\right] \qquad (2.13)$$
$$z_j = \left(\frac{2}{5}\right)^{1/2} q\cos\left(\varphi + \frac{4\pi j}{5}\right)$$

where, α_j is the polar angle locating the j^{th} bond, r_j is the radius of the projection of the atomic pseudo-rotation trajectory onto the plane of the undistorted ring, R_j is the distance from the geometric center of the planar five-membered ring to the center of the projection of the j^{th} trajectory onto the plane of the undistorted furanose ring, q is the puckering amplitude (in Å), and ϕ is the pseudo-rotation phase. Given the structural constraints of a furanose ring, the only adjustable parameter is the pucker amplitude q.

Equation 2.13 describes the coordinates of the heavy atom framework of the furanose ring as a function of ϕ. The model can be generalized to include the reorientation of the C-^{2}H bond.[23-24, 34] For each value of pseudo-rotation phase angle ϕ there corresponds a set of angles (θ_{PC}, ϕ_{PC}) specifying the orientation of a C2'-^{2}H2″ bond relative to a coordinate system fixed to the framework of the planar furanose ring. The set

of angles (θ_{PC}, ϕ_{PC}) are then used to calculate the site frequencies $\omega_{\pm}^{i}$ according to equation 2.3.

The trajectory of the ^{2}H2″ deuteron corresponding to this set of angles is a slightly curved ellipse,[23-24, 34] and the angular dispersion increases with puckering amplitude q. If the value of q is 0.4 Å, the total excursion along the direction of the major axis of the ellipse of 52º and the total excursion along the direction of the minor axis is about 20º degrees, whereas for q = 0.2 Å the corresponding excursions along the major and minor semi-axes are about 28º and 10º, respectively.

To calculate the deuterium lineshape with equations 2.4 – 2.7, information is needed regarding the atomic trajectory as well as a form for the operator π in equation 2.2. Jumps between discrete sites is a good approximation when barriers separating the sites approach 10 $k_{B}T$, corresponding to a barrier of about 25 kJ/mol at T = 300K. Instead of treating the motion of the C-^{2}H bond as a jump between discrete sites, the C2'-^{2}H2″ bond can be envisioned as diffusing over these low energy barriers. π then has the form of a steady state Fokker-Planck operator:

$$\pi = D\left[\frac{\partial^2}{\partial\phi^2} + \frac{1}{k_B T}U'(\phi)\frac{\partial}{\partial\phi} + \frac{1}{k_B T}U''(\phi)\right] \tag{2.14}$$

where $U(\phi)$ is the external potential in which the C-^{2}H bond diffuses and D is the orientation-independent diffusion coefficient associated with the motion of the C-^{2}H bond.

Solving equation 2.1 using the form for π given in equation 2.14 is difficult even for relatively simple forms for $U(\phi)$. Nadler and Schulten[43] have introduced a finite

difference approximation for equation 2.14, where π is represented by a tri-diagonal matrix with elements defined by:

$$\begin{aligned} \pi_{ij} &= \frac{1}{\tau}\left(\frac{W_i}{W_{i\pm1}}\right)^{1/2};\ j = i \pm 1 \\ \pi_{ij} &= -\left(\pi_{i,j-1} + \pi_{i,j+1}\right);\ j = i \\ \pi_j &= 0, \text{otherwise} \end{aligned} \tag{2.15}$$

where in equation 2.15 $W_i = \frac{e^{-U(\phi_i)/k_BT}}{Q}$ and Q is the partition function. The diffusion coefficient, D, can be expressed in terms of the correlation time (τ_c) as well as the kinetic rate constant (k) and the unit angular step (δ) where $\frac{1}{2\tau_c} = D = \frac{k\delta^2}{2}$. This formulation has been used to describe the motion of heme groups in proteins,[44] amino acid side chains in proteins[45] and polymethylene chains in lipids.[46] The continuous trajectory is therefore discretized into a small (*ca*. 10) number of individual steps, which are then incorporated into the MXET1 rate matrix. The rate constant k and the associated diffusion coefficient D represent transitions between individual sites along this discretized trajectory. Each site then has associated angular components owing to the 3D orientations of a C-^{2}H bonds going through pseudorotation, and site populations according to equation 2.15.

To simulate the motion of the furanose ring, a form of the potential $U(\phi)$ must be chosen. Initial studies modeling highly dynamic furanose rings from ^{2}H SSNMR used a two-site jump model,[30, 32] where the implicit assumption is that the two sites represent the bottom of potential energy wells corresponding to two primary pseudorotation conformers, and that there is a large ($\geq 10\ k_BT$) barrier between the wells. Other models[34] built upon this work to use a two well potential energy surface[47-

[49] with barriers on the order of 5 $k_B T$. Further development of different potentials includes unequal double-well, weighted double-well, single-well, and harmonic energy surfaces.[23-24] All of these potentials can be approximated with simple analytical equations, making them good candidates for use in this analysis. Their equations and graphical representations are given in Table 2.1 and Figure 2.2.

The lowest energy region for each type of surface, represented by the bottom of each well, was assumed to be a value of ϕ = 180°, which corresponds very closely to the C2'-*endo* pseudorotation conformation. This was considered to be our lowest energy conformation, as it is the primary furanose ring conformation observed in high resolution structures of DNA.

To determine best fit equilibrium lineshapes, a library of hundreds of simulated spectra are generated varying the puckering amplitude q (equation 2.13), the rate constant k between sites along the pseudorotation trajectory (equation 2.14), the form of the potential $U(\phi)$ and the 'barrier height' U_0 (both from Table 2.1), offering a small number of independent variables in the analysis.

Type of Surface	Analytical Expression
Equal double well	$U(\phi)=\frac{U_0}{2}(1-cos\,2\phi)$
Unequal double well	$U(\phi)=\frac{U_{0_1}}{2}(1-cos\,2\phi), 0<\phi<\pi$; $U(\phi)=\frac{U_{0_2}}{2}(1-cos\,2\phi), \pi<\phi<2\pi$
Weighted double well	$U(\phi)=\frac{U_{0_1}}{2}(1-cos\,2\phi), \frac{\pi}{2}<\phi<\frac{3\pi}{2}$; $U(\phi)=\frac{U_{0_1}}{2}(W(1-cos\,2\phi)), \frac{3\pi}{2}<\phi<\frac{\pi}{2}$
Single well	$U(\phi)=\frac{U_0}{2}(1-cos\,\phi)$
Harmonic	$U(\phi)=\frac{\kappa}{2}(\phi-\phi_0)^2$

Table 2.1. Analytical expressions of the *U(ϕ)* angular dependent potential energies used for simulation library.

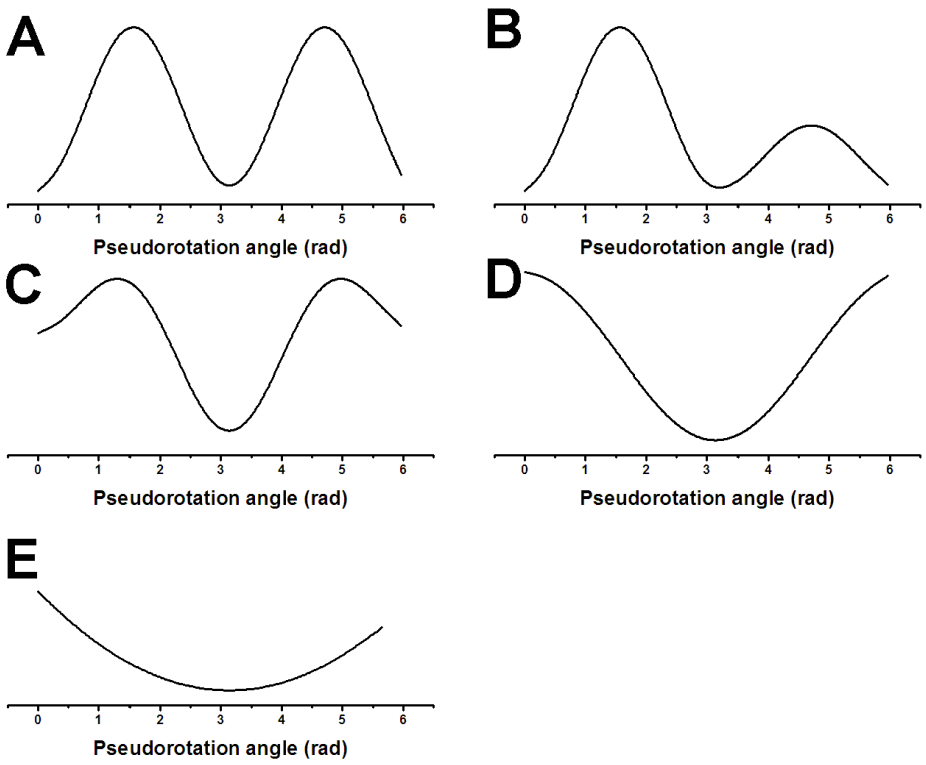

Figure 2.2. Angular dependent potential energies used in simulations described in text. (A) Equal double well. (B) Unequal double well. (C) Weighted double well. (D) Single well. (E) Harmonic.

2.3 Methods

2.3.1 Synthesis of Selectively Deuterated DNAs. [2"- ^{2}H]-2'-deoxyuridine and [2"- ^{2}H]-thymidine nucleosides for deuterium dynamics experiments were prepared using well-established methods.[27] Deuteration was confirmed by proton solution NMR of the labeled nucleoside monomers. Loss of integrated signal intensity of the appropriate peaks and changes in splitting patterns of coupled protons are recorded for all labeled monomers. Conversion to the 5'-O-(dimethoxytrityl)- 3′-O-(2-cyanoethyl-N,N diisopropylphosphoramidite derivatives was also performed using well-established methods.[28]

Automated DNA synthesis was outsourced to SynGen Inc. (Hayward, CA). 2 × 10 μmol syntheses were performed with the labeled amidites discussed above, using standard automated synthesis procedures. The following DNA 12-mer constructs were used, where the underlined residue contains the deuterium label: 5'-CGCGAATTCGCG-3' (T:A sample) and 5'-CGCGAATUCGCG-3' (U:A sample). The control sequence is the well-known Dickerson dodecamer,[29] as furanose ring motion within this sequence has been extensively studied via solid-state deuterium NMR.[30-34] These prior investigations made this sequence a particularly attractive control, and any changes to the appropriate properties due to the replacement of a thymine with a single uracil replacement of a thymine could be monitored and compared to previous work.

22 mg (T:A) and 42.8 mg (U:A) of DNA were weighed, salted to 10% w/w with NaCl, heated to 75ºC for 10-15 minutes, cooled to room temperature for 1 hour for annealing, then frozen using liquid nitrogen and lyophilized. To remove residual HDO, samples were re-dissolved in deuterium-depleted water and then lyophilized twice. The

dry sample was packed into a 4 mm sample chamber (standard NMR tube, cut to fit static solid probe), and placed in sealed hydration chambers (79% relative humidity) over saturated salt solutions containing deuterium-depleted water.[35] Water addition (W = number of waters per nucleotide) to the samples was monitored gravimetrically to a hydration level of 10 ± 1. Samples were then sealed and allowed to equilibrate an additional week before use in NMR experiments. Water introduced to salted nucleic acids progressively populates the grooves, bases, and backbone in a well-characterized fashion, beginning with the phosphates and eventually reaching limiting local hydration such that subsequent addition of water contributes primarily to the surrounding water shells.[36] At W=10, local hydration of the nucleic acid backbone and bases is substantially complete.[36-38] Global tumbling, however, is still partially suppressed at this hydration level. The measured dynamics once hydration has reached W=10 should thus be reflective of the local motions present in solution, while global motions are suppressed.

2.3.2 Solid-State NMR Spectroscopy. All experiments were performed on a 14 T Avance NMR spectrometer (Bruker Biospin, Billerica, MA), using a home-built singly tuned solid-state NMR probe. The sample coil was 4 mm in internal diameter, giving a deuterium $\pi/2$ pulse of 1.8 μs. Unless otherwise specified, the pulse delay between the two pulses in the quadrupolar echo sequence was 40 μs.

Following shifting to the echo, deuterium free-induction decays were processed with 2 kHz exponential apodization prior to Fourier transformation. $\langle T_{2e} \rangle$ relaxation times reported in this paper are obtained by varying the echo delay; echo amplitudes were extracted as a function of the delay between the quadrupole echo $\pi/2$ pulse, and fit to a single-exponential function. The time constant of that function is simply the $\langle T_{2e} \rangle$. Spin

lattice relaxation times reported in this paper are obtained by saturation recovery; echo amplitudes were extracted as a function of the delay between acquisitions, and fit to a single-exponential function. The time constant of that function, making the reasonable assumption that a hard pair of $\pi/2$ pulses followed by a free induction decay leaves the system saturated, is simply the $<T_1>$.

2.4 Results

Unsymmetrized quadrupole echo lineshapes for DNAs containing T:A and U:A base pairs are given in Figure 2.3. Visual inspection of the lineshapes indicates they are essentially identical within the limits of the noise level. The T:A lineshape has some modest differences when compared to initial studies on an identical DNA sequence,[31] but these could be attributed to T_{2e} effects due to the difference in field strength, or perhaps small differences in hydration or sample conditions.

Figure 2.4 shows the raw data and fits for the $<T_{2e}>$ relaxation experiments obtained for both the T:A and U:A samples. $<T_{2e}>$ relaxation times are listed in Table 2.2. The values for these relaxation times are equal within the experimental error, and it is concluded that the slow motions within the two samples are effectively indistinguishable.

Figure 2.5 shows the raw data and fits for the saturation recovery experiments to obtain $<T_1>$ for both the T:A and U:A samples. The powder averaged spin-lattice relaxation times are listed in Table 2.2. The values for these relaxation times suggest that the faster motions within the two samples are also indistinguishable. The relaxation results coupled with the essentially identical quad echo lineshapes suggest that the two samples have indistinguishable dynamic properties.

The simulated spectrum that provides the best fit to the symmetrized experimental results is shown in Figure 2.6. It was generated by superpositions of the uniform helical rotation and the local motion of the furanose ring. The local motion was simulated as described above. The best fit simulation arises by using a

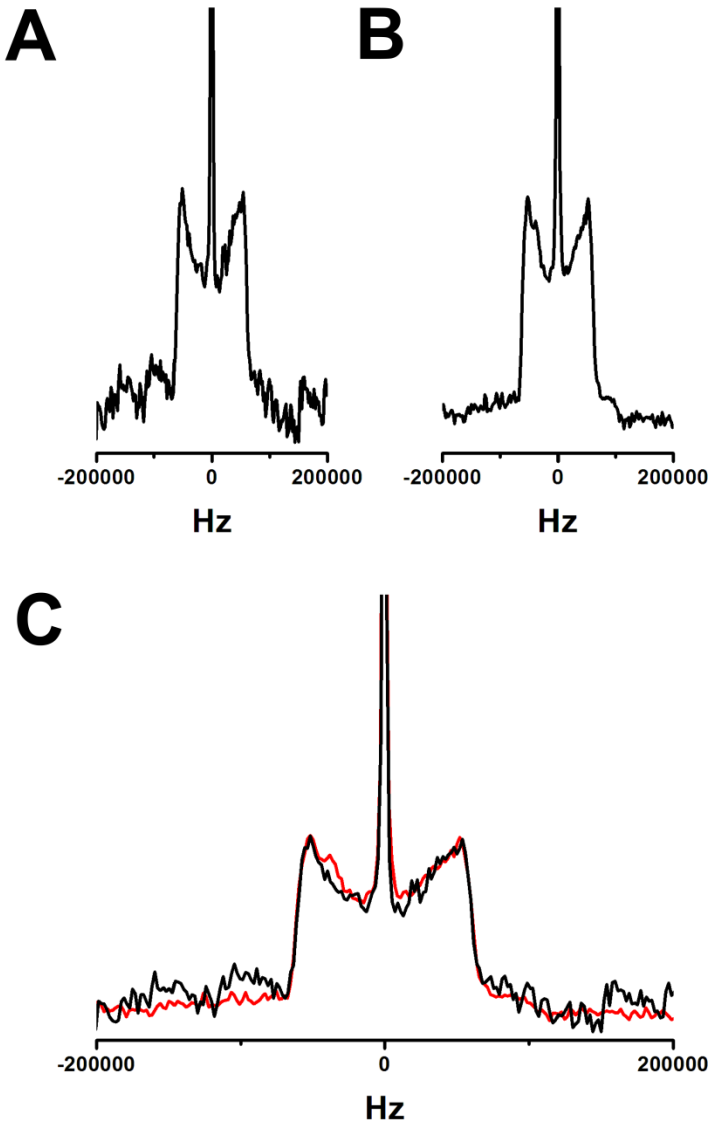

Figure 2.3. Comparison of quadrupole echo lineshapes for (A) T:A and (B) U:A samples and overlayed (C) where the T:A sample is in black and the U:A sample in red. Note that central isotropic peaks are residual HDO.

Relaxation time	T:A sample	U:A sample
$\langle T_{2e} \rangle$	67.3 (± 1.0) μs	60.2 (± 2.2) μs
$\langle T_1 \rangle$	52.6 (± 12.8) ms	45.3 (± 15.4) ms

Table 2.2. Comparison of relaxation times between T:A and U:A samples.

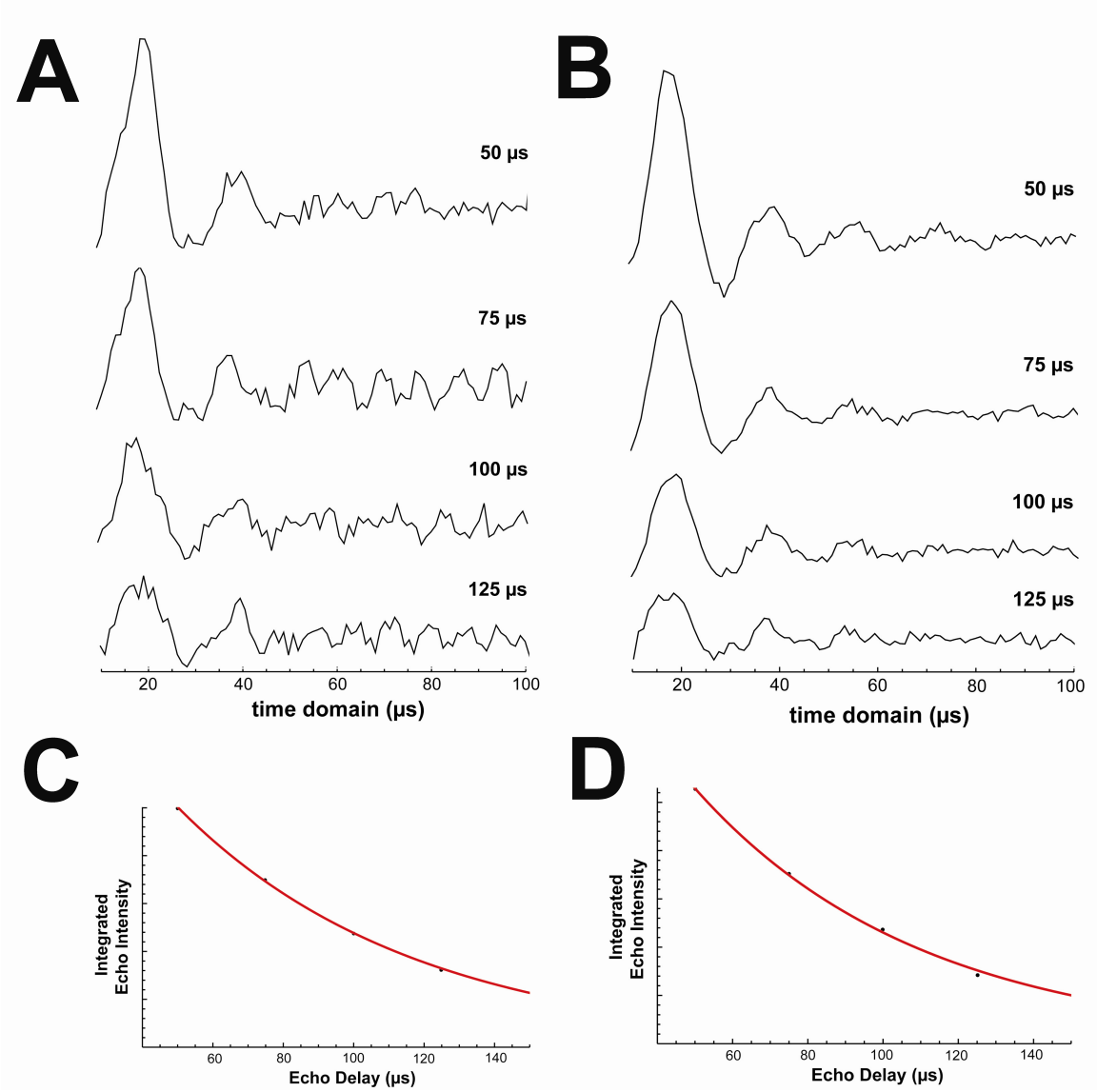

Figure 2.4. $<T_{2e}>$ data and fits. (A) FIDs for T:A. (B) FIDs for U:A. (C) Exponential fit to data from part A for the T:A sample. (D) Exponential fit to data from part B for the U:A sample.

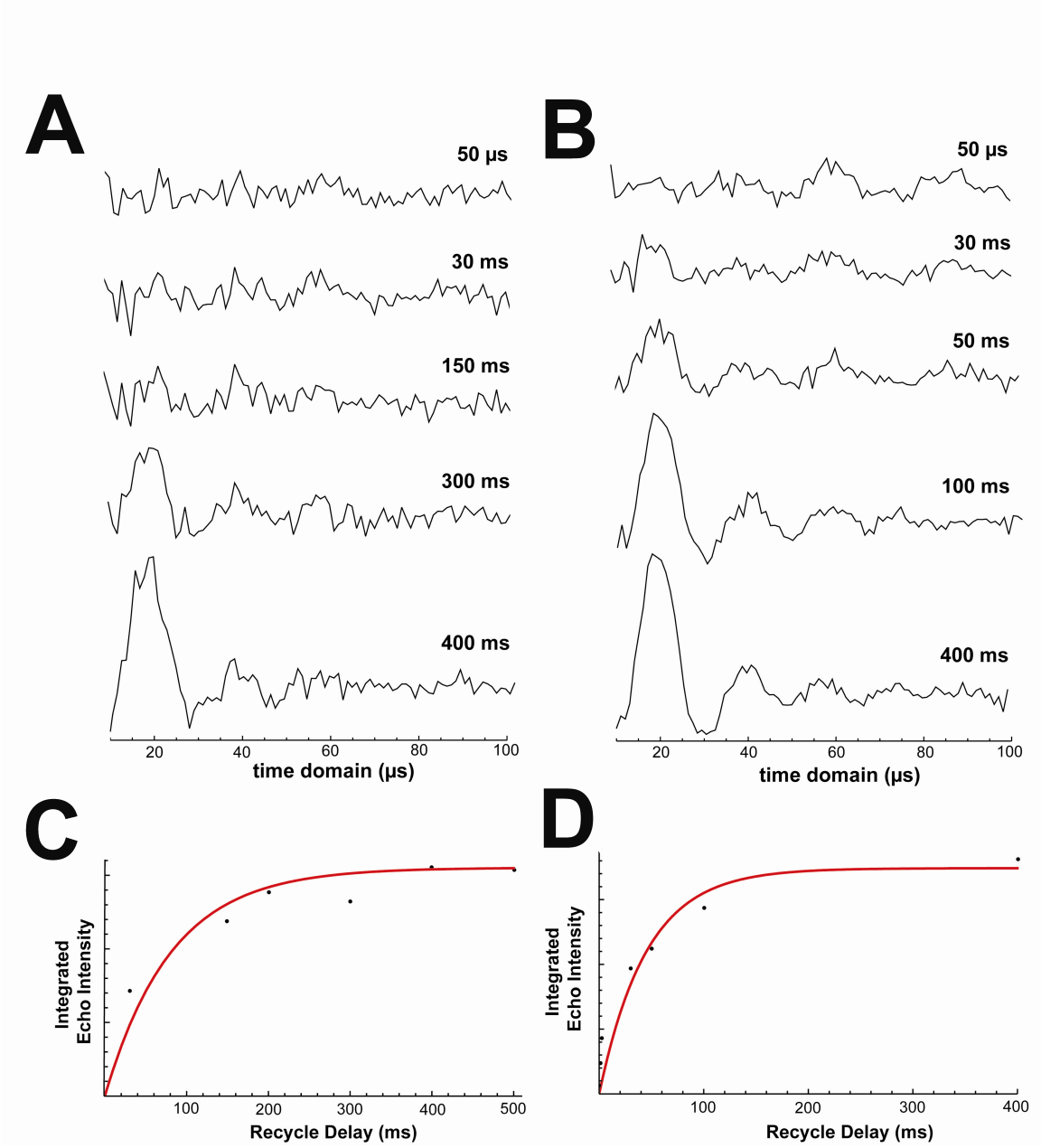

Figure 2.5. <T_1> data and fits. (A) FIDs for T:A. (B) FIDs for U:A. (C) Exponential fit to data from part A for the T:A sample. (D) Exponential fit to data from part B for the U:A sample.

Hookean/harmonic potential for the potential energy surface (Figure 2.2e), with a value of U_0 of 5.5 k_BT. The rate constant between sites along the pseudorotation trajectory is 1.0×10^7 Hz, which leads to a value of 2.0×10^6 rad^2/sec for the diffusion constant D. The puckering amplitude is $q = 0.5$ Å. Additionally, there is a structural difference between this simulation and those performed previously. The angle of the C-^{2}H bond relative to the longitudinal helix axis was determined to be $\theta = 25°$.

It can be shown that the local motions present contribute significantly to the overall lineshape. Figure 2.7 builds up the motions one at a time. First, in Figure 2.7a is a static lineshape. Figure 2.7b contains only the helical motions present due to the filled first hydration shell. Figure 2.7c shows a small amplitude (10 degree) local libration of the C-^{2}H bond, and Figure 2.7d repeats the best fit simulation confirming that there is contribution from non-activated motion.

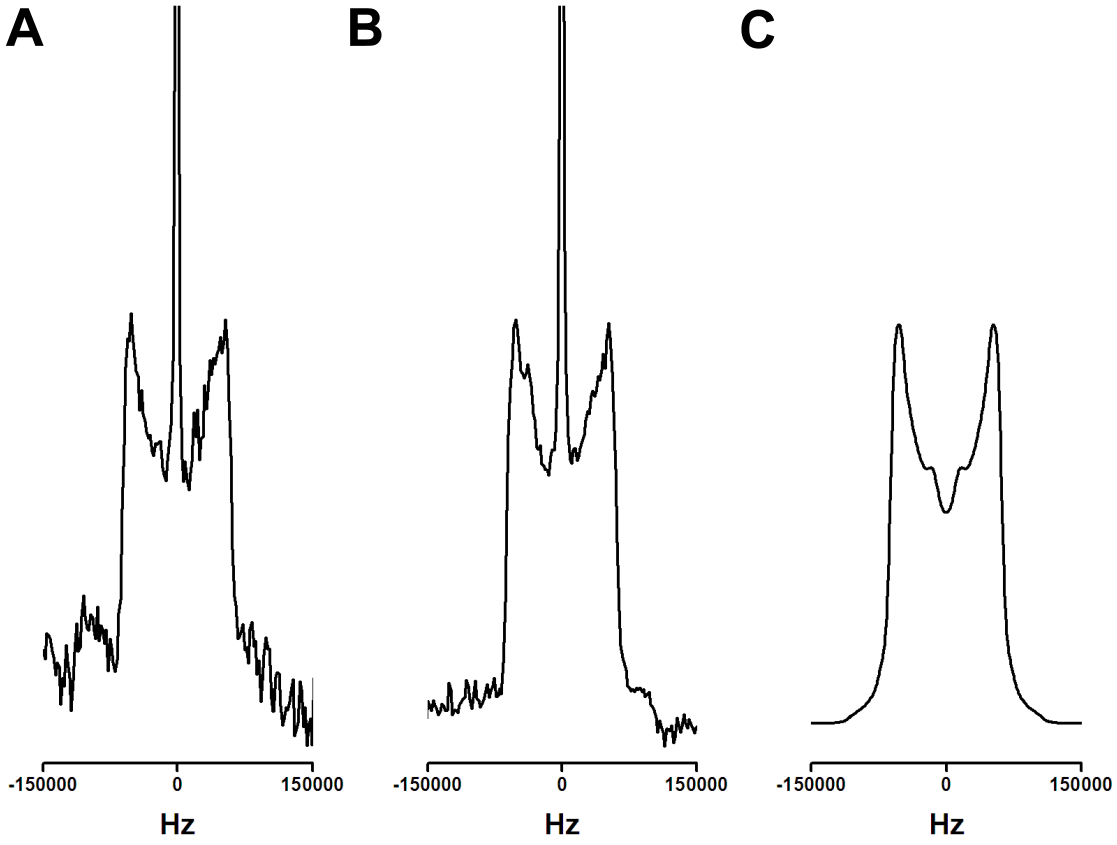

Figure 2.6. Comparison of the T:A (A) and U:A (B) experimental lineshapes to the best fit simulation (C). The specific simulation parameters are discussed in the text. Note that central isotropic peaks are residual HDO.

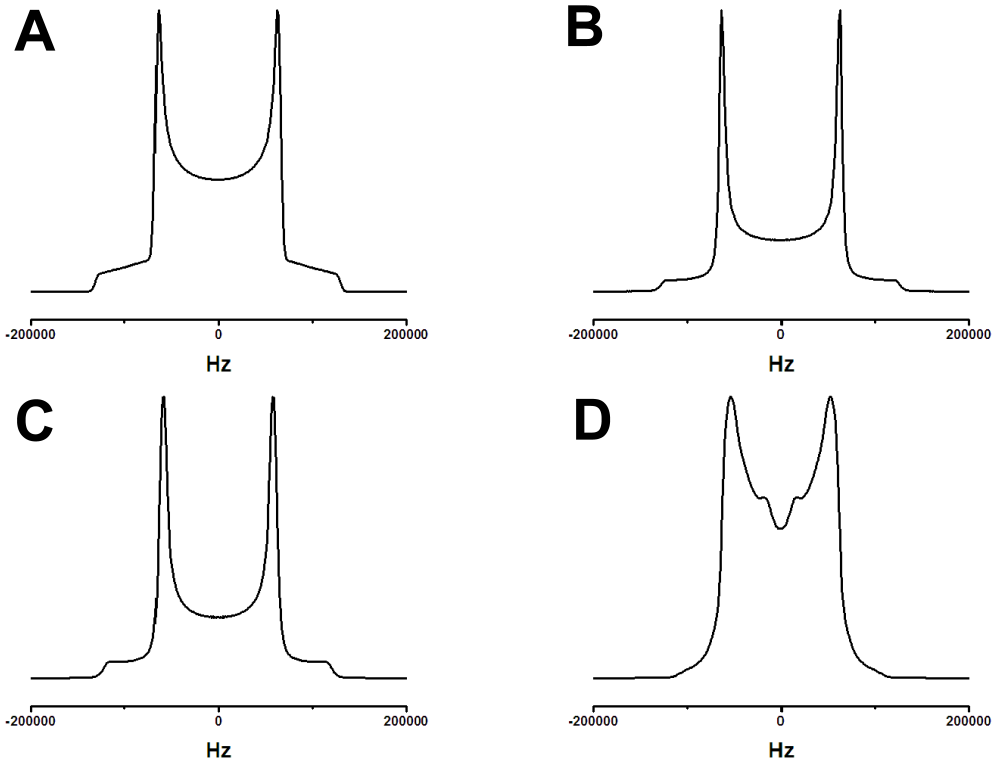

Figure 2.7. Simulation build-up for the best fit to the experimental data. (A) Static deuterium lineshape. (B) Lineshape containing only 6-site slow helical rotation. (C) Lineshape with small angle (10°) libration for local motion of the sugar ring superimposed on the slow helical motion. (D) Best fit simulation with specific details discussed in the main text.

2.5 Discussion and Conclusion

The results from this work lead to the conclusion that there is essentially no difference in the dynamics of the furanose rings for DNA containing uracil compared to a thymine in the same sequence context. This conclusion was drawn from the lack of discernable differences in quadrupole echo lineshapes, $\langle T_1 \rangle$ and $\langle T_{2e} \rangle$ values. These results are analogous, but less dramatic to those seen in the M. *HhaI* binding sequence.[23] The *HhaI* methyltransferase uses a nucleotide flipping mechanism during its conversion of cytosine to 5-methy-cytosine[26] similar to UNG. Deuterium SSNMR indicates a significant difference in motions between the target deoxycytidine residue and the surrounding sequence.[23] However, the differences in the equilibrium lineshapes and simulated lineshape parameters between the deoxycytidine substrate sequence and the methylated analogue are more subtle. It was concluded that the unmethylated site has a modestly larger puckering amplitude (0.40 Å for C versus 0.35 Å for 5-methyl-C).[23] This indicates that while there is a significant sequence dependence on local dynamic properties, there may be a smaller dependence upon modifications of the target DNA base. The results herein for furanose ring lineshapes and relaxation times are analogous in that they compare samples that contain uracil versus thymine which, like cytosine and 5-methyl-cytosine, only differ by the presence of a single methyl group. Given the identical base pairing and the small steric difference between the two bases, it is perhaps not surprising that demethylation of the DNA leads to no detectable dynamic perturbation.

It must be stated that the lack of distinct spectral features suggests a difficulty in providing non-degenerate simulations that fit the experimental lineshapes. The results

from the simulations presented do offer some insight into the local motions as well. Although the harmonic potential is without a significant energy 'barrier' to conformational rearrangement, it can be seen in Figure 2.2 that the angular dependence to the barrier is significant, and the pseudorotational conformations outside the bottom of the well will not be highly populated. Indeed, this is indicated in Figure 2.8, which displays populations of sites within the trajectory used in the best fit simulation and compares them to the site populations of an unequal two well potential (Figure 2.2b) with barrier heights 5 k_BT and 1 k_BT. These data suggest that these sugar rings, undergoing motion within this harmonic well, are only sampling a small number of conformations. This result is consistent with previous data for the T:A indicating small angle motion,[31] that even though the overall angular excursion of the sugar ring is large (0.5 Å), only a small number of these sites are highly populated.

What is the importance, if any, of these similar motions and what conclusions can be drawn from this work? Taken initially, these results may seem to contradict the hypothesis of a dynamic component to lesion recognition, and this is certainly a possible interpretation. In particular, the replacement of thymine with uracil while base pairing with adenine represents one of the least significant structural alterations and this type of structural disruption may provide relatively small effects on the overall DNA properties. Sequences containing other lesions are currently being examined.

The differences between the U:A and T:A samples are indistinguishable, indicating that purely dynamic recognition process via local motions of the furanose

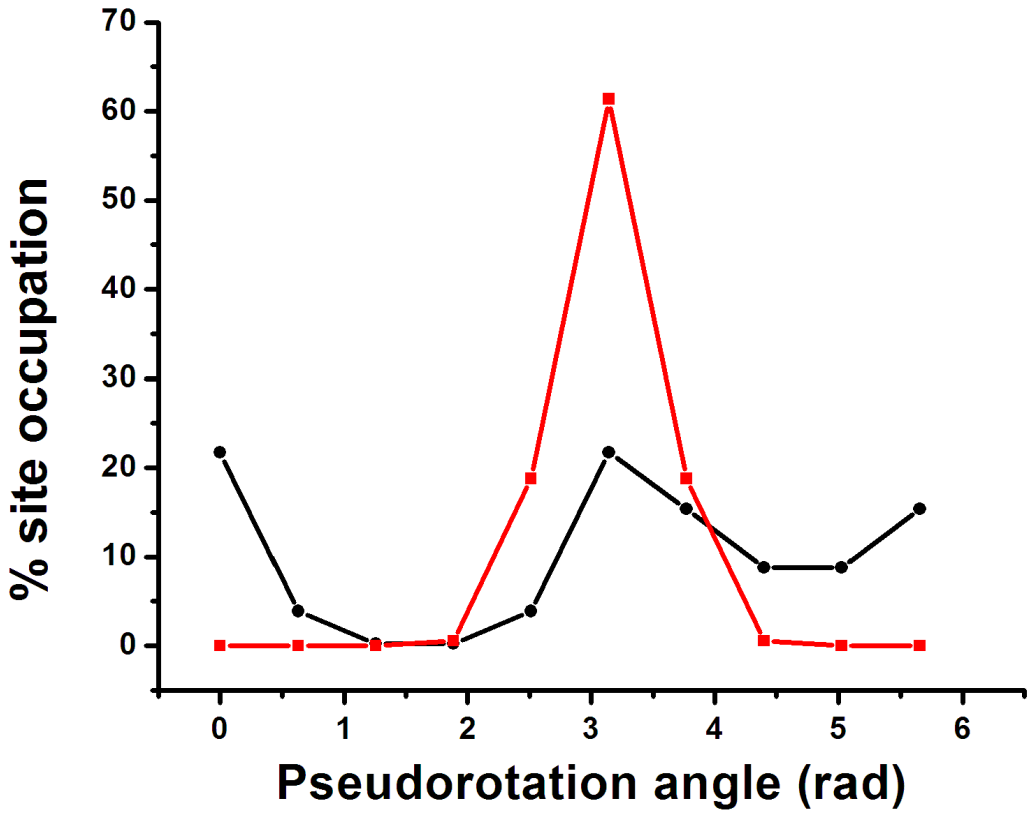

Figure 2.8. Overlay of the site populations for the trajectories of the potential energy used in the best fit simulations (red) for U:A and T:A compared to an arbitrary unequal double (Figure 2.2(B)) well of barrier heights 5 k_BT and 1 k_BT (black). The symbols represent the ten sites along the discretized trajectory. specific details discussed in the main text.

ring is unlikely to be the sole arbiter of specificity. Indeed, other aspects including steric hindrance and electronic properties of the 5-substituent and thermal stability of base pairing partners have been proposed as other means of discrimination.[50] The current work agrees well with established data on several fronts. First, uracil base-paired with adenine provides lower activity in uracil DNA glycosylases from various organisms[50-53] as well as other DNA glycosylases with uracil as a substrate,[15] relative to U:G base pairs. Additionally, U:A base pairs have significantly different melting thermodynamic properties relative to U:G base pairs (as well as other base pairing contexts).[15] If there is a dynamic component to recognition, they may be exhibited more strongly in a U:G base pair than in a U:A base pair.

Sequence context may also play an important factor. It has been stated that for uracil recognition, that sequence plays a greater role than base pairing in determining repair enzyme activity.[51-53] When a large number of sequences were tested for activity, the flanking sequence for the samples in this work (i.e. 5'-AATUCG-3') offered a relatively low activity compared to the optimum flanking sequence (5'-GCAUAA-3').[51-52] Flexibility in the nearest neighbors of a uracil has also been proposed as a aspect of recognition from fluorescence studies.[54] Sequence dependence may ultimately result from differences in base stacking energies between the uracil and its neighbors. Differences in base stacking have been suggested as playing a role in recognition of thymine in T:G base pairs.[15, 54]

If there is a dynamic component of discrimination and/or recognition, there may be other reasons that differences do not appear in these two samples. First, the important subunits responsible may be other than the furanose ring. In fact, the backbone

methylene group in C:G base pairs have more significantly different local motions than in 5-Me-C:G base pairs.[24] The backbone moiety contains dihedral angles that must be altered significantly during the nucleotide flipping process, and it may be these that have significantly lowered energy barriers in lesion-containing DNAs. Second, the timescale of motions may be beyond the region accessible by the experiments performed, and it has been proposed that U:G removal happens faster than U:A removal.[52]

This work offers insight into the local dynamic properties of DNA containing the single base lesion uracil base paired to adenine. To our knowledge, this work represents the first site-specific ^{2}H SSNMR study of damaged DNA dynamics. Given the importance of understanding the functional roles of dynamics in biomolecular recognition, analysis and quantification of these motions becomes important in determining their role. Our results suggest the presence of a uracil base pairing with adenine does not alter the dynamics of the furanose ring from lineshape and relaxation analysis. Therefore, a dynamic component of uracil recognition via the furanose ring is unlikely to be the sole arbiter of specificity for uracil DNA glycosylase.

References

1. Schärer, O. D. *Angew. Chem. Int. Ed.* **2003,** *42,* 2946.
2. Sousa, M. M. L.; Krokan, H. E.; Slupphaug, G., *Mol. Aspects. Med.* **2007**, *28*, 276.
3. Krokan, H. E.; Standal, R.; Slupphaug, G. *Biochem. J.* **1997**, *325*, 1.
4. McCullough, A. K.; Dodson, M. L.; Lloyd, R. S. *Annu. Rev. Biochem*. **1999**, *68*, 255.
5. Stivers, J. T.; Jiang, Y. L. *Chem. Rev.* **2003**, *103*, 2729.
6. Fromme, J. C.; Verdine, G. L. *Adv. Protien Chem*. **2004**, *69*, 1.
7. Huffman, J. L.; Sundheim, O.; Tainer, J. A. *Mut. Res.* **2005**, *577*, 55.
8. Zharkov, D. O.; Grollman, A. P.; *Mut, Res.* **2005**, *577*, 24.
9. Roberts, R. J.; Cheng, X., *Annu. Rev. Biochem.* **1998**, *67*, 181.
10. Slupphaug, G.; Mol, C. D.; Kavli, B.; Arvai, A. S.; Krokan, H. E.; Tainer, J. A. *ature* **1996**, *384*, 87.
11. Lukin, M.; de los Santos, C., *Chem. Rev.* **2006**, *106*, 607.
12. Delort, A.-M.; Neumann, J. M.; Molko, D.; Hervé; Téoule, R.; Tran Dinh, S. *Nucleic Acids Res.* **1985**, *13*, 3343.
13. Ghosh, M.; Kumar, N. V.; Varshney, U.; Chary, K. V. R. *Nucleic Acids Res.* **2000**, *28*, 1906.
14. Dong, J.; Drohat, A. C.; Stivers, J. T.; Pankiewicz, K. W.; Carey, P. R. *Biochemistry* **2000**, *39*, 13241.
15. Liu, P.; Theruvathu, J. A.; Darwanto, A.; Lao, V. V.; Pascal, T.; Goddard III, W.; Sowers, L. C. *J. Biol. Chem*. **2008**, *283*, 8829.
16. Stivers, J. T. *Prog. Nucleic Acid Res. Mol. Biol*. **2004**, *77*, 37.
17. Yang, W. *Cell Res.* **2008**, *18*, 184.
18. Cao, C.; Jiang, Y. L.; Stivers, J. T.; Song, F. *Nature Struct. Mol. Biol.* **2004**, *11*, 1230.
19. Parker, J. B.; Bianchet, M. A.; Krosky, D. J.; Friedman, J. I.; Amzel, L. M.; Stivers, J. T., *Nature* **2007**, 449, 433.
20. Sundina, A. E.; Volkov, E. M.; Kubareva, E. A. *Biocatalysis* **2000**, *41*, 121.
21. Meints, G.; Drobny, G. *Biochemistry* **2001**, *40*, 12436.
22. Miller, P. A.; Shajani, Z.; Meints, G. A.; Caplow, D.; Goobes, G.; Varani, G.; Drobny, G. P. *J. Am. Chem. Soc*, **2006**, *128*, 15970.

23. Meints, G. A.; Miller, P. A.; Pederson, K.; Shajani, Z.; Drobny, G. P. *J. Am. Chem. Soc.* **2008**, *130*, 7305.

24. Pederson, K.; Meints, G. A.; Shajani, Z.; Miller, P. A.; Drobny, G. P. *J. Am. Chem. Soc*, **2008**, *130*, 9072.

25. Echodu, D.; Goobes, G.; Shajani, Z.; Pederson, K.; Meints, G. A.; Varani, G.; Drobny, G. P. *J. Phys. Chem. B,* **2008**, *112*, 13934.

26. Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. *Cell* **1994**, *76*, 357.

27. Robins, M.J.; Wilson, J. S.; Hansske, F. *J. Am. Chem. Soc.* **1983**, *105*, 4059.

28. Gait, M.J., ed. *Oligonucleotide Synthesis: A Practical Approach*. Practical Approach Series, IRL Press: Oxford, **1984**.

29. Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, *149*, 761.

30. Huang, W.-C.; Orban, J.; Kintanar, A.; Reid, B. R.; Drobny, G. P. *J. Am. Chem. Soc.* **1990**, *112*, 9059.

31. Hatcher, M. E. *A Solid-State Deuterium NMR Investigation of the Local Dynamics of Nucleotides in the EcoRI Restriction Endonuclease Binding Site*; University of Washington: Seattle, WA, **1996**.

32. Hatcher, M. E.; Mattielo, D. L.; Meints, G. A.; Orban, J.; Drobny, G. P. *J. Am. Chem. Soc.***1998**, *120*, 9850.

33. Geahigan, K. B.; Meints, G. A.; Hatcher, M. E.; Orban, J.; Drobny, G. P. *Biochemistry* **2000**, *39*, 4939.

34. Meints, G. A.; Karlsson, T.; Drobny, G. P. *J. Am. Chem. Soc.* **2001**, *123*, 10030.

35. Lide, D. R. *Handbook of Chemistry and Physics*, CRC Press: Boca Raton, Fla., USA, *84*, 15-26, **2003-2004**.

36. Falk, M.; Hartman, K. A.; Lord, R. C. *J. Am. Chem. Soc.* **1962**, 84, 3843.

37. Wang, A. C.; Kennedy, M. A.; Reid, B. R.; Drobny, G. P. *J. Magn. Res.* **1994**, 105, 1.

38. Schurr, J. M.; Fujimoto, B. S.; Diaz, R.; Robinson, B. H. *J. Magn. Res.* **1999**, 140, 404.

39. Torchia, S.; Szabo, A., *J. Magn. Reson.* **1982**, *49*, 107.

40. Vold R. R.; Vold, R. L. In *Advances in Magnetic and Optical Resonance*; Warren, W., Ed.; Academic Press: San Diego, Vol.16, *85–171*, **1991**.

41. Alam, T.; Drobny, G.P. *Chem. Rev.* **1991**, *91*, 1545.

42. Herzyk, P.; Rabczenko, A. *J. Chem. Soc. Perkin Trans. II* **1985**, 1925.

43. Nadler, W.; Schulten, K., *J. Chem. Phys.* **1986**, *84*, 4015.

44. Nadler, W.; Schulten, L. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 5719.

45. Wittebort, R.; Szabo, A. *J. Chem. Phys.* **1978**, *69*, 1722.

46. Wittebort, R. J.; Olejniczak, E. T.; Griffin, R. G. *J. Chem. Phys.* **1987**, *86*, 5411.

47. Levitt, M.; Warshel, A., *J. Am. Chem. Soc.* **1978**, *100*, 2607.

48. Olson, W. K., *J. Am. Chem. Soc.* **1982**, *104*, 278.

49. Saenger, W. *Principles of Nucleic Acid Structure*; Springer Verlag: New York, 1984.

50. Liu, P.; Burdzy, A.; Sowers, L. C. *Chem. Res. Toxicol*. **2002**, *15*, 1001.

51. Nilsen, H.; Yazdankhah, S. P.; Eftedeal, I.; Krokan, H. E. *FEBS Lett.* **1995**, *362*, 205.

52. Slupphaug, G.; Eftedal, I.; Kavli, B.; Bharati, S.; Helle, N. M.; Haug, T.; Levine, D. W.; Krokan, H. E. *Biochemistry*, **1995**, *34*, 128.

53. Bellamy, S. R. W.; Baldwin, G. S. *Nuc. Acids Res*. **2001**, *29*, 3857.

54. Allawi, H. T.; SantaLucia Jr., J. *Biochemistry*, **1997**, *36*, 10581.

Chapter 3

Theoretical and Experimental Solid-State NMR Study of a Melamine-Cyanuric Acid Complex

Summary

The highly insoluble adduct formed between melamine and cyanuric acid is widely used as a cosmetic and polymer additive, and has been implicated in animal and human kidney failure resulting from the ingestion of melamine-tainted food-products. While the canonical structure of the complex is a layered graphitic two-dimensional hexagonal latttice, a low-quality X-ray structure reported a monoclinic distortion.

We have used ^{1}H, ^{2}H, and ^{13}C solid-state NMR to investigate this material. Our results reveal a highly permeable porous structure, which allows access of small-molecules to the interior of the crystallites and rapid exchange of hydrogens between crystals and ambient water vapor. Quadrupolar coupling constants (C_Q) obtained by sideband analysis of the chemical-shift-resolved ^{2}H MAS SSNMR spectra indicate the presence of only two chemically distinct deuterons (rather than the five predicted from the crystal structure) and indicate a strong N-H...N hydrogen bond between the cyanuric acid donor and melamine acceptor. The longer NH…O hydrogen bond (N…O ~ 2.96 Å) has a significantly larger C_Q value of 133 kHz compared to the C_Q of the shorter NH…N hydrogen bond (N…N ~ 2.86 Å) at 90 kHz. In contrast to previous work[2] ^{1}H and ^{13}C MAS spectra similarly show two chemically distinct hydrogens and carbons respectively, consistent with the hexagonal structure; we obtain ^{2}H linewidths substantially less than

the previously reported splittings due to the monoclinic distortion. Careful inspection of previously published ^{2}H and ^{13}C MAS spectra suggests that the apparently increased multiplicity is not a result of lattice inequivalence, but rather, is a consequence of mis-setting of the magic angle, a species of artifact which has in the past led even experienced spectroscopists astray. Finally, powder x-ray diffraction shows no evidence of microcrystalline structure in the complex, and is consistent with two-dimensional sheets stacked in a disordered fashion.

3.1 Introduction

The recent controversies regarding the industrial melamine contamination of human and animal foodstuffs have brought intense international interest in this compound.[1] With a high mass percentage of nitrogen, melamine is frequently added to animal feed to give it false "high protein" content. Melamine is used along with cyanuric acid as a non-halogen flame retardant, a basis for supramolecular self-assembly[3], a potential solid lubricant, and in many plastic products.[2] The nature of the cyanuric acid-melamine complex (CAM) has been studied now for decades, with Ranganathan publishing the crystal structure after the complex was painstakingly grown under hydrothermal conditions.[4] The challenge to obtaining quality crystallographic data prior to Ranganathan's work is that the CAM complex is insoluble in organic solvents and rapidly precipitates from aqueous solution.

Structurally, the non-covalent nature of the CAM complex is of particular interest. While hydrogen bonds have been readily probed by many scientific means, little is currently known about their contribution to the robust structure of CAM and its formation via molecular self-assembly.[5] IR spectroscopy revealed that CAM is held together by hydrogen bonding.[6] It has been shown both experimentally and theoretically that NMR chemical shift values are highly dependent on the strength of the hydrogen bond.[7-9] Using solid-state NMR, Berglund and Vaughan compiled deuterium quadrupole coupling constants for a variety of O-H…O hydrogen bonds. They showed that the coupling constants decrease from 275 kHz for the infinite $r_{o...o}$ (corresponding to the isolated O-H group on calcium hydroxide) to about 50 kHz for the shortest $r_{o...o}$ distance.[7] Therefore, a

systematic solid-state NMR study of the CAM complex can provide structural insight pertaining to the relative strengths of these hydrogen bonds.

Structural insight is most easily achieved through studying localized interactions. For example, in an NMR system the two most localized interactions are the proton or deuteron-specific chemical shift or the quadrupole coupling constant (C_Q). Previously, we have shown that short O-H. . .O hydrogen bonds show a positive dependence of the deuterium quadrupole coupling constant with temperature by determining the value of $d(C_Q)/dT$ (kHz/K) from deuterium SSNMR spectra.[9] Little has been done to determine if these same trends are witnessed in longer, nitrogen-based hydrogen bonds. The presence of two distinct nitrogen-containing hydrogen bonds in the CAM complex makes it an ideal system of study.

Previously published x-ray diffraction[4] and SSNMR data[15] seemed to be inconsistent with the conclusion that the CAM complex forms highly symmetric, graphitic-like sheets. For these previously-mentioned studies, only crystals grown under hydrothermal conditions were suitable for single-crystal x-ray diffraction. These studies reported an overall monoclinic distortion of the complex that was inconsistent with our preliminary quantum calculations. Follow up work by the Damodaran group[15] used ^{13}C, ^{2}H and ^{15}N SSNMR methods that seemed to confirm the monoclinic crystal structure proposed by Ranganathan's original work. These seemingly inconsistent results with our initial calculations led to our study of the crystal structure with SSNMR.

3.2 Theory

3.2.1 The Pake powder pattern and magic angle spinning. Solid-state NMR is mostly focused around the analysis of single crystal and powder samples. In solution NMR, rapid molecular tumbling in solution averages orientation-dependent, or anisotropic, effects. While these effects are not completely averaged out in solution state, they are sufficiently averaged over the time range of a typical NMR experiment, therefore resulting in the lack of spectral frequencies associated with these anisotropic interactions. Anisotropic effects in solid samples arise from inequivalent nuclear sites within the magnetic field, leading to the super position of many spectral frequencies associated with each nuclear orientation. This is especially problematic in powder samples and the resulting spectral shape, called the "powder pattern," can obscure any valuable molecular properties one might wish to obtain from an NMR experiment.

Magic angle spinning (MAS) is a technique developed in the 1960's[19] that is widely used in solid-state NMR to remove or lessen anisotropic effects of chemical shielding, heteronuclear and homonuclear dipolar coupling, and quadrupolar coupling. Rapid rotation of a sample at the magic angle can be used to reduce a broadened powder pattern to one well-resolved line representing the isotropic NMR frequency. However, to completely average out the anisotropic interactions, the frequency of sample rotation must be significantly greater than that of the anisotropic interaction. Modern MAS probes can effectively spin at > 50 kHz while our facility is limited to a 15 kHz rotational frequency. This often results in anisotropic interactions that are not completely averaged, resulting in additional spectral lines known as 'spinning sidebands'. These spinning sidebands radiate out in either direction from the central isotropic frequency at an interval

equal to the spinning speed. At lower spinning speeds, it can often be difficult to differentiate the central isotropic frequency from the spinning sidebands. However, the central frequency will not change position as a function of the spinning speed whilst the spinning sidebands will.

Magic angle spinning is an effective way of obtaining high resolution spectra because all second-rank tensor interactions share a common $(3\cos^2\theta - 1)$ dependence, where θ is the angle between the most distinct principal axis of the tensor and the magnetic field. By spinning rapidly at an angle θ_R: $(3\cos^2\theta_R - 1) = 0$, or in other words $\theta_R \sim 54.74°$, where θ_R is the angle between the spinning axis and the applied magnetic field (Figure 3.1). Optimization of this magic angle is imperative to the quality of spectra obtained by MAS SSNMR. Significant line broadening and the lack of well-resolved spinning sidebands can occur if the magic angle is even slightly "off angle." It is common practice to use the ^{79}Br nucleus in KBr to optimize the angle, as no decoupling is required to get narrow, well-resolved lines and sidebands. KBr is ideal in that the quadrupolar nucleus, ^{79}Br, is abundant and rapidly relaxes so that angle optimization can be done with single scans using either the transformed data or the free induction decay (FID). Optimization of the magic angle with the FID is particularly convenient as the presence of rotational echoes in the FID means there is no need to perform a Fourier transform after each scan. Figure 3.2 compares the FID and spectrum of a pseudo-optimized magic angle for KBr (figure 3.2 A and C) versus an angle that has not been optimized (figure 3.2 B and D).

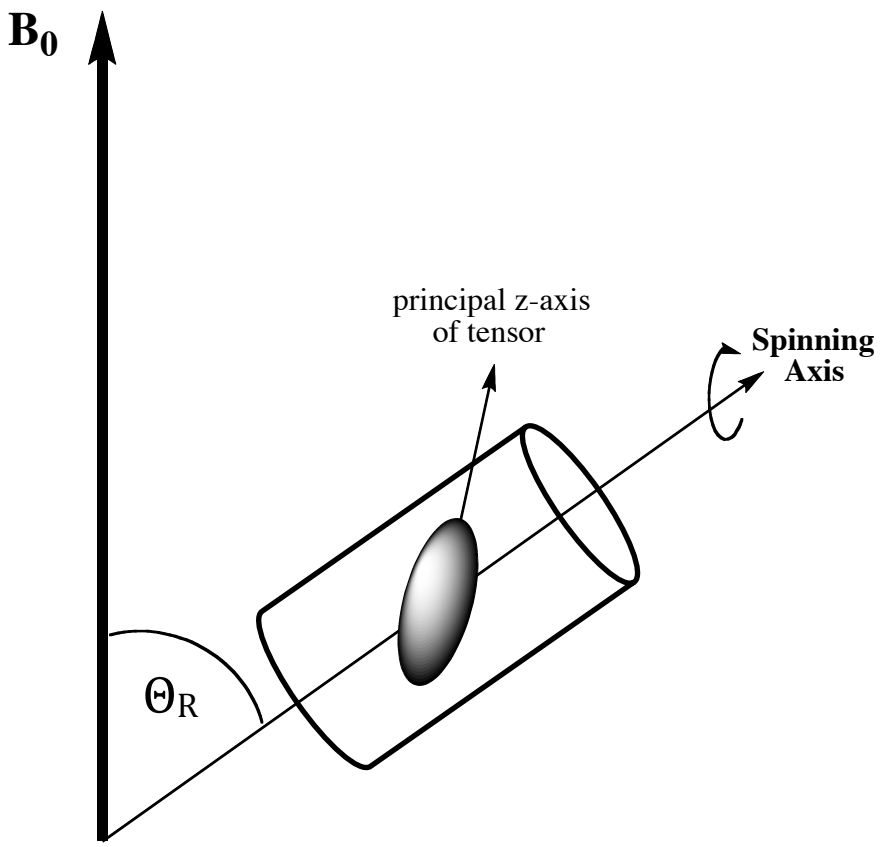

Figure 3.1. Magic angle spinning places a sample in a cylindrical rotor at an angle Θ_R (54.47°) with respect to the applied magnetic field (B_0). The ellipsoid in the picture represents the tensor of the anisotropic interaction, and labeled is the distinct principal axis of the tensor in the principal axis frame.

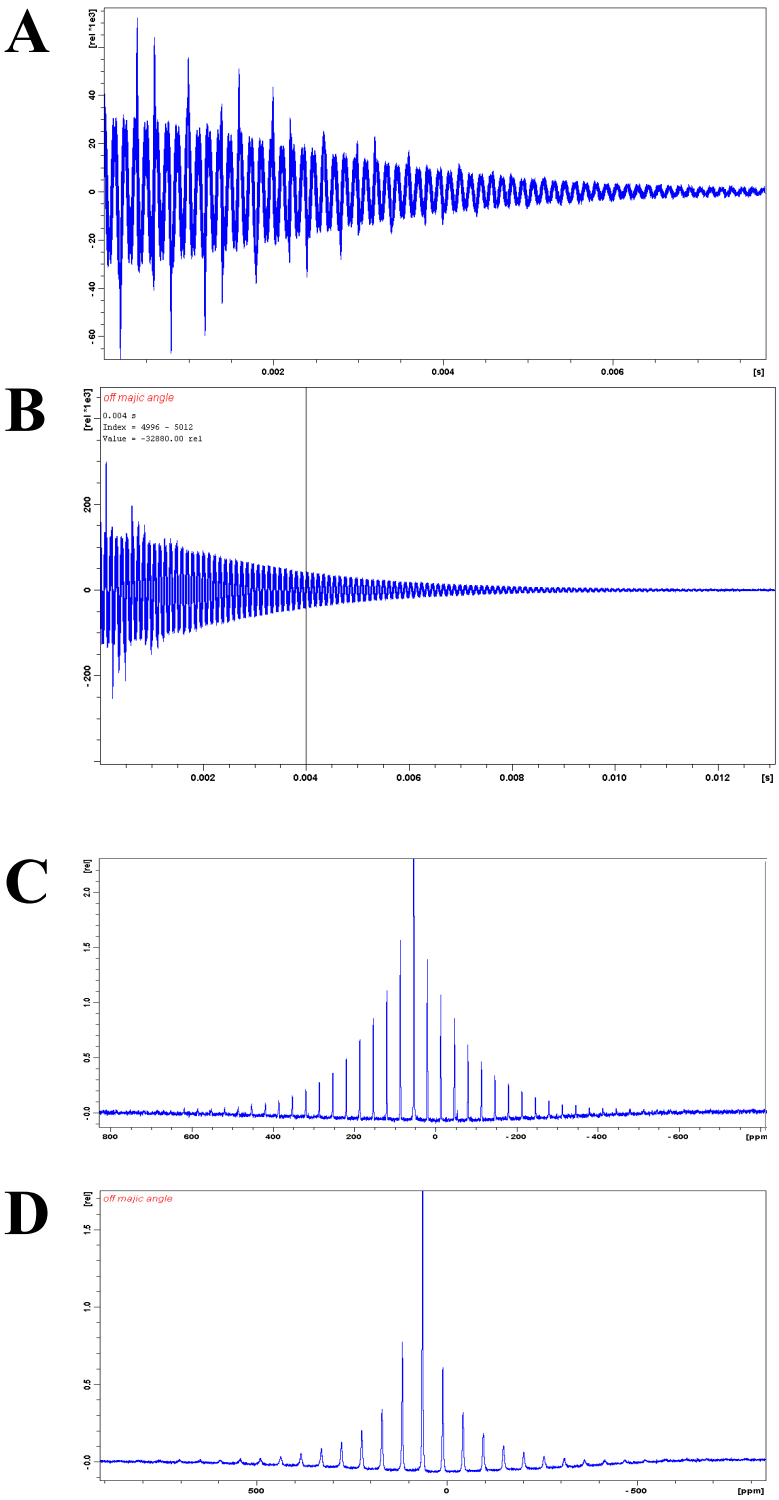

Figure 3.2. ^{79}Br FIDs and spectra of KBr demonstrating the importance of optimizing the magic angle. (A)The FID when the magic angle is well-set. Notice the proliferation of rotational echoes compared to (B) the resulting FID from a poorly set angle. (C) The spectrum of KBr with a well-set angle compared to (D) the spectrum from a poorly set angle.

Notice in the non-optimized FID the complete lack of any rotational echoes and the presence of only a few pairs of spinning sidebands.

Upon spinning a sample at the magic angle, the Larmor frequency evolves into

$$\omega_S = -\omega_0\left[\overline{\sigma} + A_1\cos(\omega_R t+\gamma) + B_1\sin(\omega_R t+\gamma) + A_2\cos(2\omega_R t+2\gamma) + B_2\sin(2\omega_R t+2\gamma)\right] \quad (3.1)$$

where

$$A_1 = \frac{2}{3}\sqrt{2}\sin\beta\cos\beta\left[\cos^2\alpha\left(\sigma_{xx}^{PAF}-\sigma_{zz}^{PAF}\right)+\sin^2\alpha\left(\sigma_{yy}^{PAF}-\sigma_{zz}^{PAF}\right)\right]$$

$$B_1 = \frac{2}{3}\sqrt{2}\sin\beta\sin\alpha\cos\alpha\left(\sigma_{xx}^{PAF}-\sigma_{yy}^{PAF}\right)$$

$$A_2 = \frac{1}{3}\left[\left(\cos^2\alpha\cos^2\beta-\sin^2\alpha\right)\left(\sigma_{xx}^{PAF}-\sigma_{zz}^{PAF}\right)+\left(\sin^2\alpha\cos^2\beta-\cos^2\alpha\right)\left(\sigma_{yy}^{PAF}-\sigma_{zz}^{PAF}\right)\right]$$

$$B_2 = \frac{2}{3}\cos\beta\sin\alpha\cos\alpha\left(\sigma_{xx}^{PAF}-\sigma_{yy}^{PAF}\right) \quad (3.2)$$

In the previous equations, ω_R is the rotational frequency about the magic angle and α, β and γ are the Euler angles that rotate the principal axis frame to the rotor axis frame. Further analysis of equation 3.1 shows the time-independent isotropic chemical shielding (σ) as well as two terms that oscillate about the central frequency by ω_R and $2\omega_R$. If ω_R is significantly larger than the chemical shielding anisotropy, the magnitude of the oscillating frequency is negligible and no sidebands are present.[12, 17-18]

3.2.2 Cross polarization magic angle spinning (CPMAS). The observation of dilute spin systems offers another challenge in solid-state NMR. In the absence of strong homonuclear dipolar interactions, relaxation times for some sites can be prohibitively long. Additionally, the signal-to-noise is almost invariably worsened due to the dilute nature of the spin. Excessively long relaxation times can significantly increase

experimental times. Additionally, dilute spin systems require thousands of scans for well-resolved spectra, which also leads to long experiment times. Posing an additional problem in the study of dilute spin systems is the broadening effect of heteronuclear dipolar coupling when the dilute spin is adjacent to a relatively abundant spin. This is especially problematic for the SSNMR of organic and biological solids as the dilute $^{13}C/^{15}N$ spins are often near abundant ^{1}H spins.

A combination of high-power decoupling coupled with the use of a specific pulse program can alleviate these issues. In this particular pulse program, called cross-polarization (CP), the dilute spin derives its magnetization from an adjacent and abundant spin-1/2 nucleus (often ^{1}H in organic compounds). An initial 90° excitation pulse on the abundant spin creates transverse magnetization in the rotating frame. An on-resonance contact pulse is then applied to the abundant spin system and the spins of both nuclei are "spin-locked" along a like axis of the rotating frame. In the rotating frame, we label the field acting on the nuclei B_1 since the contact pulse is on-resonance. The precession of each nucleus in the rotating frame will depend on the magnitude of the rf applied as well as the nuclei's respective gyromagnetic ratio. A condition called the Hartmann-Hahn condition (named in honor of the individuals who first discovered cross-polarization) must be met where the precession of both spins are equal, i.e. $\gamma_I B_{1,I} = \gamma_S B_{1,S}$. This is achieved by setting the amplitude of the rf pulses for each spin in a way that the 90° pulse lengths are equivalent for each spin. The Hartmann-Hahn condition ensures that the energy barrier between the two spin states for each nucleus is equivalent, so that transfer of polarization from one spin (the abundant spin) to the other spin (the dilute spin) results

in no net change in energy and no net change in polarization. The general pulse program for the cross-polarization experiment can be found in Figure 3.3.

3.2.3 Side band spinning analysis: the Herzfeld Berger method. Magic angle spinning in SSNMR can provide both the isotropic and anisotropic parameters of a solid sample. Maricq and Waugh[20] originally showed that careful measurement of all sideband intensities can be used to determine the chemical shift anisotropy, but this method is not always feasible because of sideband overlap. Analysis of a smaller subset of spinning sidebands by methods set forth by Herzfeld and Berger[12] can provide the chemical shift anisotropy and the asymmetry parameter. As indicated in equations 3.1 and 3.2, the intensities of the sidebands are intrinsically related to the tensor elements of the chemical shielding.

As mentioned, sidebands flank the central isotropic resonance by integer multiples $N\omega_R$ of the rotational frequency. The relative intensity of the Nth sideband is given by:

$$I_N = \frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi} |F|^2 \, d\alpha \sin\beta \, d\beta \tag{3.3}$$

where

$$F = \left(\frac{1}{2\pi}\right)\int_0^{2\pi} \exp\left[i\left(-N\theta + \Delta_-\tau_-(\alpha,\beta,\theta) + \Delta_+\tau_+(\beta,\theta)\right)\right] d\theta \tag{3.4}$$

and the values τ_- and τ_+ are dependent on the angles and, therefore, the intensities of the sidebands are a complex function of only Δ_- and Δ_+ variables. Estimates of the Δ_- and Δ_+ variables can be made and the integration completed in order to minimize and stabilize the magnitude of I_N. Experimental sideband intensities must be normalized in order to directly compare experimental parameters to the theoretical parameters.

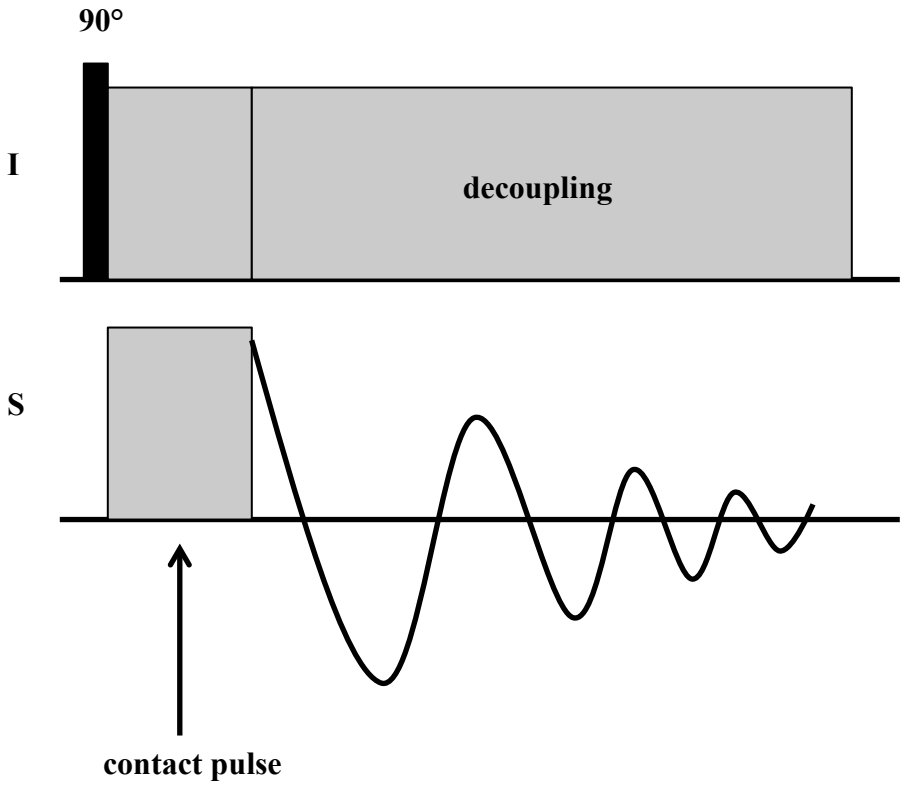

Figure 3.3. The cross-polarization pulse sequence. The abundant spin is given as I and the dilute spin is given as S.

This requires that all experimental sidebands be accurately measured which, as previously noted in the work of Maricq and Waugh, is very difficult. However, this normalization concern can be alleviated by working with the ratios of sideband intensities with respect to the intensity of the central isotropic chemical shift, I_N/I_M where M=0 and N=±1, ±2, ±3 and so on. The need to effectively cover the full range of the tensor elements of the chemical shift and asymmetry parameters remain. Using the typical convention of the tensor elements, $\sigma_{33} > \sigma_{22} > \sigma_{11}$, Herzfeld and Berger define two new variables

$$\rho = \frac{(\sigma_{11} + \sigma_{33} - 2\sigma_{22})}{(\sigma_{33} - \sigma_{11})} \tag{3.5}$$

$$\mu = \frac{(\gamma H_0)(\sigma_{33} - \sigma_{11})}{\omega_R} \tag{3.6}$$

It is apparent that μ will always be positive for positive values of rotational frequency, and that ρ has a minimum value of -1 and a maximum at +1. In the original work, specific orientations were chosen so that $\Delta_- = \mu$ and $\Delta_+ = -\mu\rho$ and then these values were substituted into Equation 3.4 to obtain computed sideband intensities. Values of μ ranging from 0 to 14 were plotted against ρ to obtain contour plots of the intensity ratios, so that these plots could be used to obtain ρ and μ from the ratio of sideband intensities taken from a spinning sideband spectrum. Modern computing technology significantly simplifies this process; the sideband pattern is integrated, the integrated intensities inputted into a script based on the Nth sideband labeling scheme, values for the chemical shielding and asymmetry parameters are estimated and the script performs a least-squared fit analysis. This process is repeated until the calculated error is minimized and stabilized.

3.2.4 Nuclear magnetic parameters and hydrogen bond environment. The magnitude of nuclear magnetic parameters depends on the chemical environment of that nucleus. In their study of the strongly hydrogen bonded system of potassium hydrogen malonate ($KHCH_2(COO)_2$), Berglund and Vaughan[20] found that protons in a strong hydrogen bond generally have larger chemical shielding anisotropies. Additionally, both the anisotropy and the isotropic chemical shift correlate with hydrogen bond strength in that the isotopic component of the chemical shielding tends to decrease (relative to the standard) with increasing strength of the hydrogen bond. X-ray and neutron diffraction data[22] of the potassium hydrogen malonate crystal showed a $R_{O...O}$ of 2.468 corresponding to an anisotropy ($\Delta\sigma$) of 27.6 ppm and a chemical shift of 20.5 ppm. By comparison, potassium hydrogen maleate has a slightly stronger hydrogen bond with a $R_{O...O}$ of 2.437.[21] As expected, the anisotropic component of the chemical shielding is increased to 30.3 ppm and the chemical shift is increased to 20.9 ppm.

Later experiments by Berglund and Vaughan[23] would confirm that these trends with hydrogen bond strengths could be extended to other nuclear magnetic parameters, mainly, the quadrupole coupling constant of deuterium in hydrogen bonds. Taking the published NMR and crystal structure data for a variety of strongly hydrogen bonded solids, they found that the trend in deuterium quadrupole coupling constants was opposite to the anisotropic component of the chemical shielding, meaning, the quadrupole coupling constant decreased in magnitude with increasing hydrogen bond strength. Conglomerate plots of these trends for the Berglund Vaughan study can be found in figure 3.4.

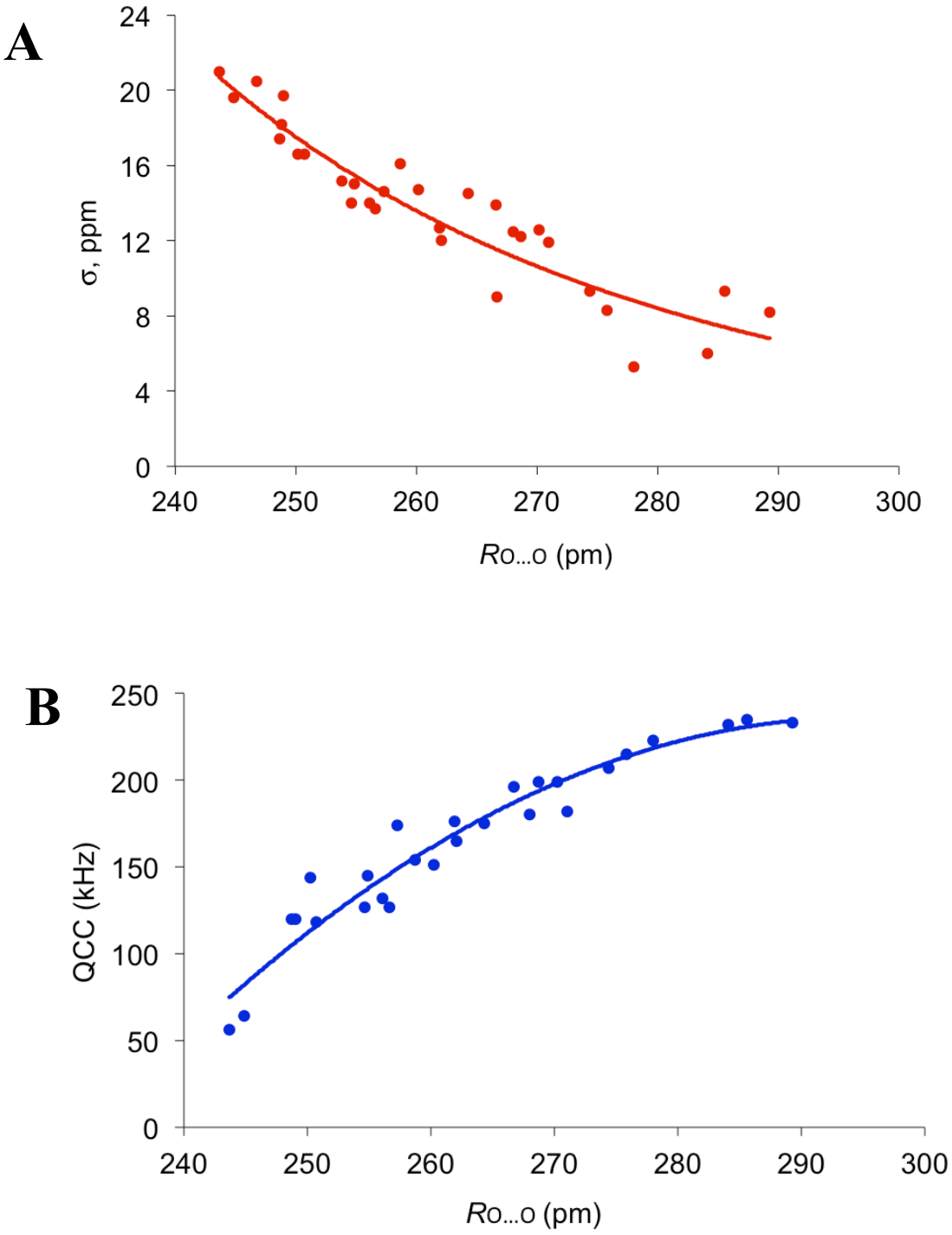

Figure 3.4. Plots showing the trend of (A) chemical shielding anisotropy and (B) deuterium quadrupole coupling constant in a variety of strong hydrogen bonded solids.[23]

3.3 Methods

3.3.1 Computational chemistry studies. Electronic structures were optimized using the GAMESS[16] software package at the B3LYP/6-311++G(2d,p) level of theory. To reduce computation time, the tri-fold rosette structure of the CAM complex was optimized first using a simplified trimer model in which cyanuric acid was the central molecule surrounded by three melamine molecules (Figure 3.5). All of the melamine atoms' molecular coordinates were frozen with exception of the nitrogen atom involved with the N-H. . .N hydrogen bond connecting the cyanuric acid to the melamine. Once the cyanuric acid molecular coordinates and the N-H. . . N hydrogen bonds were optimized, the remaining melamine atoms were unfrozen and the entire trimer was optimized. The procedure was repeated for the melamine-centered trimer (Figure 3.6) and once optimized, the coordinates were combined and the entire 174 atom tri-fold rosette structure was optimized.

3.3.2 Sample Preparation. Melamine and cyanuric acid were purchased from Sigma-Aldrich. The protonated complex (Sample A, Figure 3.10) was prepared by combining with rapid stirring 0.01 M aqueous solutions of cyanuric acid and melamine, filtering, and drying with a vacuum pump overnight. The perdeuterated complex sample (Sample B, Figure 3.10) was prepared similarly; however, deuterium oxide purchased from Sigma-Aldrich was used to prepare the solutions. The species to be deuterated was prepared in a deuterium oxide solution while the non-deuterated species prepared in water and then mixed thoroughly, filtered, and dried via vacuum. Once thoroughly dried, the samples were sealed and frozen to prevent any isotopic exchange.

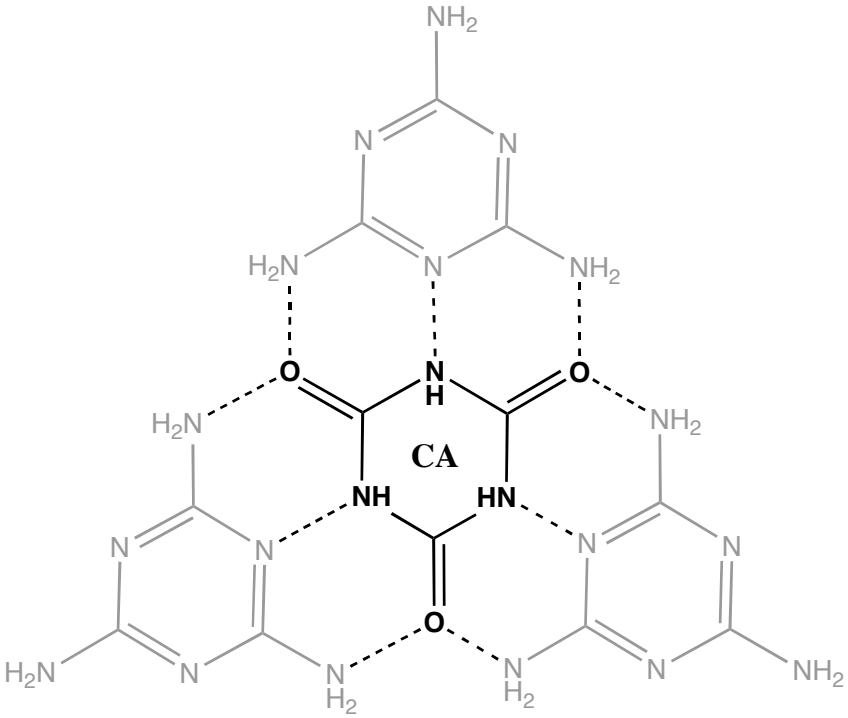

Figure 3.5. Simplified rosette structure used to optimize the central cyanuric acid molecule (bold) in the CAM complex. The molecular coordinates of the surrounding melamine molecules were initially frozen to decrease computational time (grey).

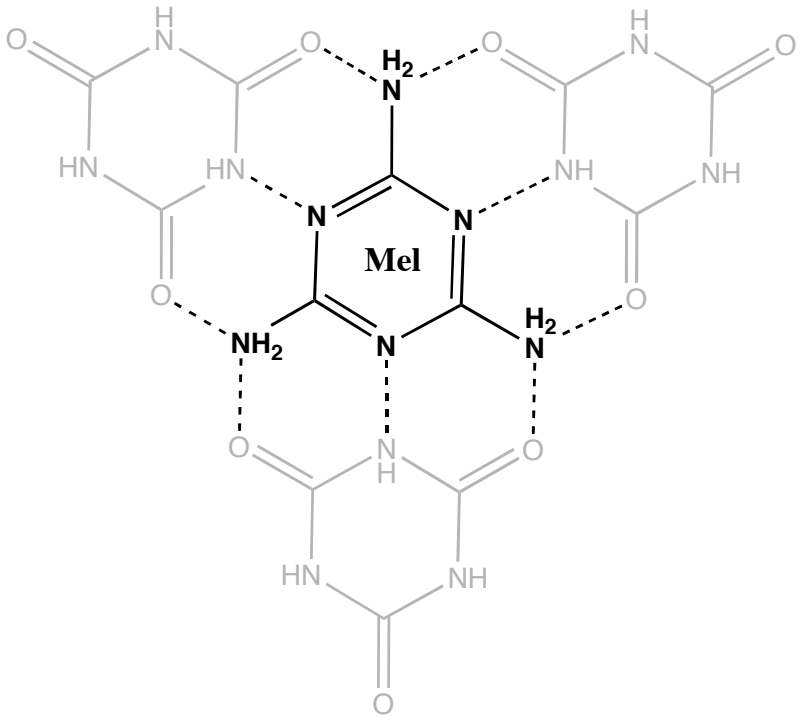

Figure 3.6. Simplified rosette structure used to optimize the central melamine molecule (bold) in the CAM complex. The molecular coordinates of the surrounding cyanuric acid molecules were initially frozen to decrease computational time (grey).

The samples for the isotopic exchange study were prepared as indicated above, though the samples were left in a deuterium vapor equilibrated desiccator to promote isotopic exchange. The chamber was equipped with the dried protonated complex and a receptacle of deuterium oxide along with a fan to ensure saturation of the deuterium isotope. Additionally, perdeuterated CAM was placed in the desiccator with protonated water to promote the back-exchange of deuterons off the CAM complex.

3.3.3 NMR Experiments. Spectra were acquired with a Bruker Avance spectrometer operating at 14.1 T (92.102 MHz for deuterium, 600.0 MHz for proton, and 150.864 MHz for carbon-13) using a simple one pulse sequence with a $\pi/2$ pulse of 2.0 μs for the proton spectra. Deuterium MAS SSNMR was performed using a Bloch decay and high power proton decoupling with a 4.0 μs $\pi/2$ pulse. The CPMAS protocol for the ^{13}C work included a 30 s second delay and a 2 ms mixing time. The proton $\pi/2$ pulse was 3.70 μs at 8 dB power. MAS rotational frequency was 15 kHz for both deuterium and proton spectra collected at ambient temperatures, while various spinning speeds for the CPMAS were completed to determine chemical shielding parameters of the carbon atoms. All proton and deuterium spectra were referenced to isotopic frequencies of residual protons and deuterium in deuterium oxide and the CPMAS data referenced to TMS. Based on the formulas derived by Herzfeld and Berger,[12] MAS and CPMAS sideband intensities were used along with least squared fit analysis to obtain experimental quadrupole coupling constants (C_Q) and asymmetry parameters (η).

3.4 Results and Discussion

The GAMESS electronically optimized rosette structure of the CAM complex is shown in Figure 3.7. Cyanuric acid and melamine are isoelectronic and the overall structure of the complex optimizes to a symmetric D_{3h} symmetry. This result suggests a highly symmetric hydrogen bonded supramolecular structure which is counter to the relatively asymmetric nature of the single-crystal grown under hydrothermal conditions in the previously published x-ray diffraction paper.[4] Previous SSNMR studies of CAM had suggested that while both cyanuric acid and melamine are independently symmetric in their monomeric forms, the inequivalence of the hydrogen bonds in the complex impose an overall asymmetry due to distinct chemical environments. Specifically, those nitrogen, carbon and proton atoms on the two-fold symmetric axis were chemically distinct from those atoms off axis. This suggests a 1:2 ^{2}H multiplicity in the cyanuric acid and a 2:2:2 ^{2}H multiplicity in the melamine as labeled in Figure 3.8. The original authors claimed that while the exocyclic amino protons in the melamine molecule are chemically equivalent, in the supramolecular structure these protons form an inequivalent number of hydrogen bonds: one, two, or none leading to the overall magnetic inequivalence. Furthermore, they noted that this same multiplicity is not observed in the splitting patterns of the carbon and nitrogen atoms on melamine, to which they attribute to chemical shift dispersion. Our preliminary computational work suggests otherwise and, as shown below, seems to be confirmed with powder SSNMR studies.

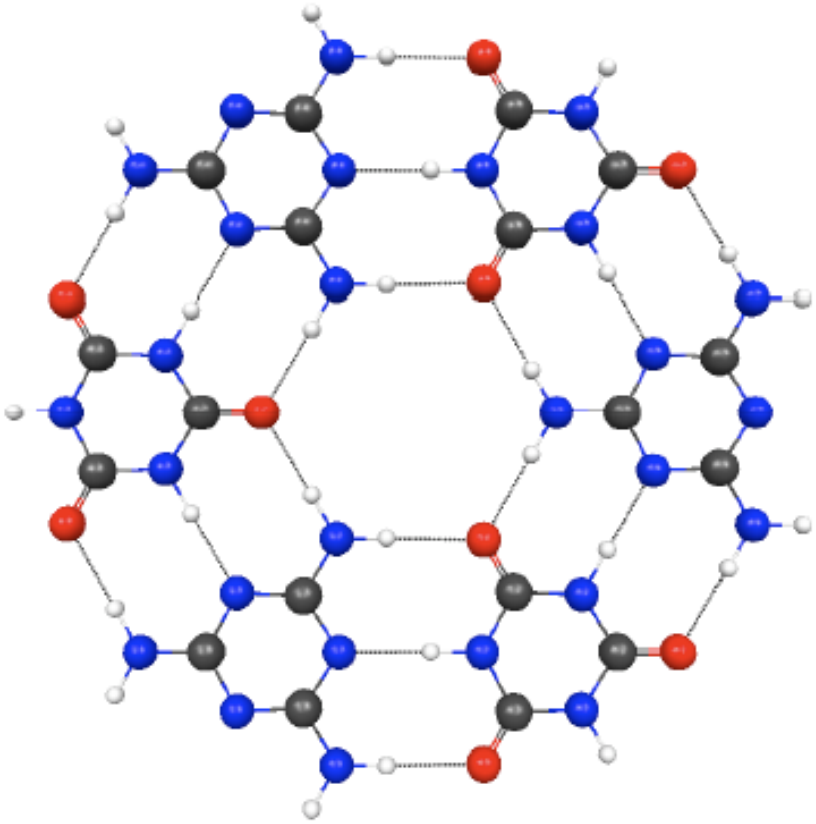

Figure 3.7. B3LYP/6-311++G(2d,p) optimized structure of the hexagonal 2D lattice of the melamine-cyanuric acid complex. The rosette structure optimizes to an overall D_{3h} symmetry.

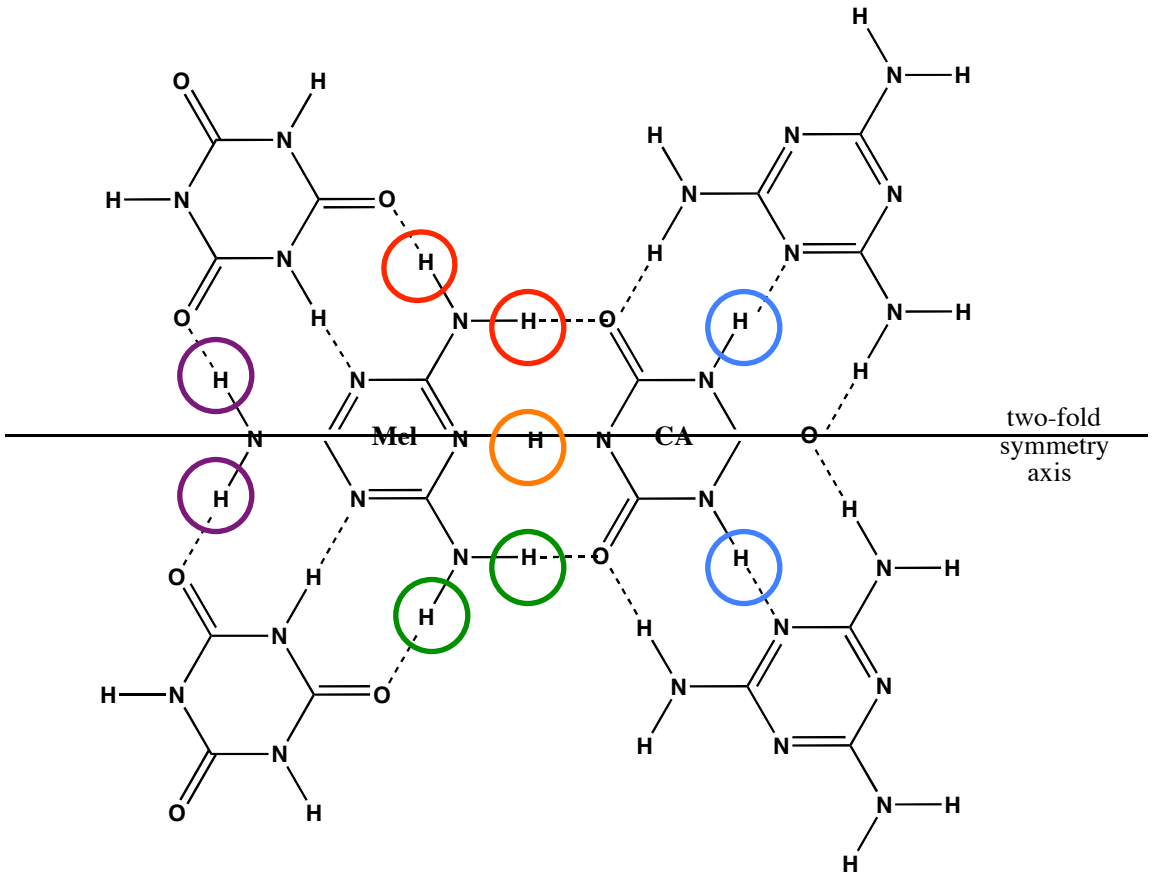

Figure 3.8. Published deuterium SSNMR chemical shift multiplicity for the CAM complex. The deuterons are cyanuric acid is 2:1 (blue:orange) while melamine is 2:2:2 (red:green:purple).

MAS SSNMR spectra of the protonated and deuterated complex at ambient temperatures are shown in Figure 3.9. The relatively narrow proton lineshape downfield corresponding to the cyanuric acid proton/deuteron in Figure 3.3a shows resolved signals in both the proton spectrum and the deuterium spectrum. The presence of two chemically inequivalent protons on the melamine species is witnessed by the two distinct chemical shift values. Additionally, the amino protons are in such proximity that they lead to dipolar broadening of the lineshape as seen in both the proton and the deuterium SSNNMR spectra. Upon analysis of the proton SSNMR spectrum, this dipolar coupling is evident due to the broadening of the center band as well as increased intensities of the sidebands. The proton spectrum of the CAM complex shows two distinct chemical shift values, one narrow line downfield relative to the wider line upfield. The distinct chemical shifts confirm the presence of two separate hydrogen bonds in the complex.

The downfield shift at 14.5 ppm corresponds to the shorter amine hydrogen bond from the N-H of the cyanuric acid to the cyclic N of the melamine. The upfield shift at 7.5 ppm corresponds to the longer amide hydrogen bond of the N-H from melamine hydrogen bonded to the carbonyl on the cyanuric acid. These assumptions are verified by the difference in the experimentally determined deuterium quadrupolar coupling constants. Sideband spinning analysis of the ^{2}H SSNMR spectrum showed that the shorter 2.87 Å amine bond of the cyanuric acid deuteron in the supramolecular structure has a decreased C_Q of 135 kHz relative to the longer amide bond in the monomeric form with a deuteron C_Q of 178 kHz.

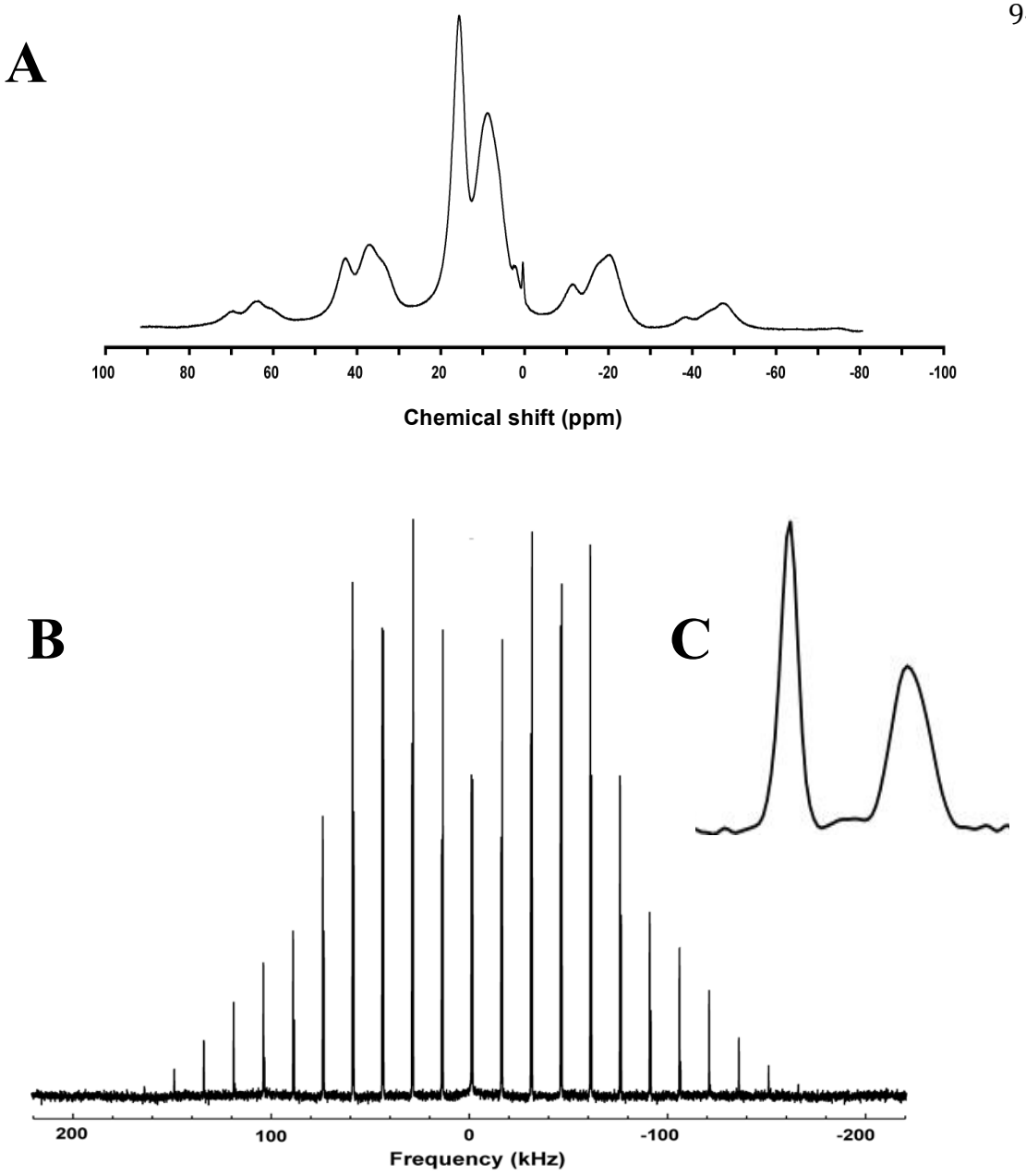

Figure 3.9. (a) ^{1}H SSNMR of the ^{1}H-labeled CAM complex. (b) ^{2}H SSNMR of the perdeuterated CAM complex. (c) Close-up view of the central frequency of the ^{2}H spectrum showing the two chemical environments corresponding to the cyanuric acid deuteron (downfield) and the melamine deuterons (upfield).

Likewise, the longer amide bond of the melamine deuteron in the CAM complex results in a larger C_Q value of 201 kHz compared to the shorter amine bond of melamine, with deuteron C_Q values of 192 and 222 kHz. These experimental results for the CAM complex are tabulated and compared to the monomeric forms of cyanuric acid and melamine in Table 3.1. The variance of the deuteron quadrupole coupling constant in the melamine species can be attributed to the fact that in the crystal structure, while exocyclic amino groups are chemically equivalent relative to reach other, the protons on the exocyclic amine are not chemically equivalent. This is attributed to the fact that one hydrogen bonded to the cyclic nitrogen atom of the adjacent melamine and the other proton is not involved in hydrogen bonding.

Precipitation of the complex from solution occurs at a rate that exceeds the isotopic exchange of deuterium, as noted in the selectively labeled CAM spectra shown in Figure 3.10. Proton powder SSNMR spectra of the selectively labeled (protonated, Sample A; deuterated, Sample B) complex show chemical shift values and lineshapes only relevant to the unlabeled constituent. Conversely, chemical shifts and lineshapes for the deuterium spectra correspond only to the labeled constituent. The proton spectrum of the CAM complex with ^{2}H labeled cyanuric acid shows a dipole-broadened lineshape due to the presence of amine protons on the melamine constituent. Likewise, the proton lineshape corresponding to the ^{2}H labeled melamine in the CAM complex is relatively narrow, since a majority of the dipolar broadening due to the melamine protons has been reduced due to the isotopic exchange.

Species	C_Q (kHz)	C_Q (kHz)	Bond Type	Bond length (Å)
	Theoretical	Experimental		
Cyanuric acid	181	178	N-H…O	2.83
Cyanuric acid (complexed)	140	135	N-H…N	2.87
Melamine	206	192, 222	N-H…N	3.0 – 3.1
Melamine (complexed)	207	201	N-H…O	2.96

Table 3.1. Tabulated data of theoretical quadrupole coupling constants compared to our experimentally determined quadrupole coupling constants of deuterons within the CAM complex. Included are hydrogen bond lengths for the cyanuric acid and melamine species both singularly and within the complex. Single point calculations of the electronic field gradient for the optimized structures of melamine and cyanuric acid provided theoretical C_Q values.

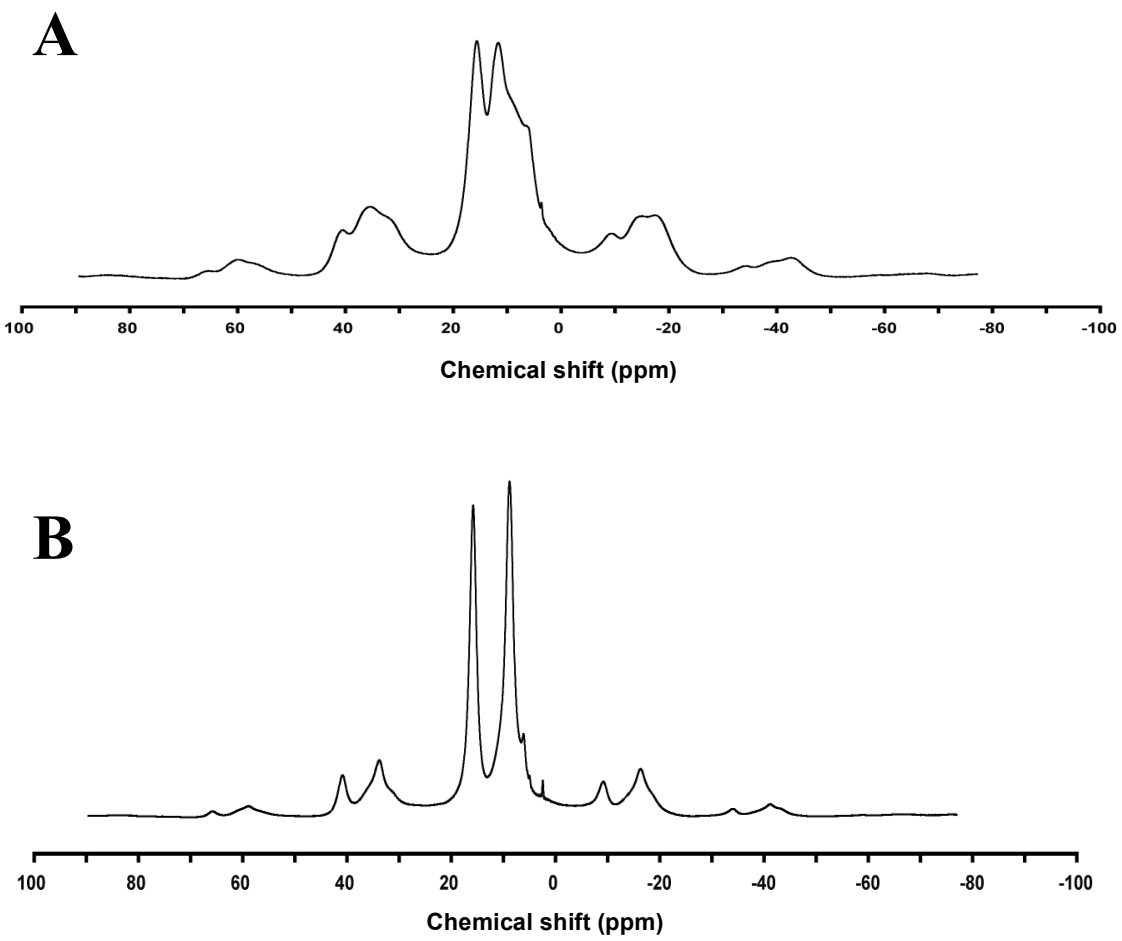

Figure 3.10. (a) Proton SSNMR spectrum of the isotopically labeled ^{2}H-Cyanuric acid, ^{1}H-Melamine complex. (b) Proton SSNMR spectrum of the isotopically labeled ^{1}H-Cyanuric acid, ^{2}H- Melamine complex.

We compared our ^{2}H and ^{13}C CPMAS (Figure 3.11) SSNMR lineshapes to those previously published[15] in an attempt to understand the chemical shift splitting discrepancy. Our experimental CPMAS lineshapes were far simpler compared to the variety of splitting patterns found in the original work. For instance, the results in the original SSNMR work showed a distinction in the chemical shifts of the atoms on the axis of symmetry compared to those atoms off the axis of symmetry. If the supramolecular CAM structure was actually hexagonal, one would expect chemical and magnetic equivalence of these atoms. Inspection of our SSNMR lineshapes showed two chemically distinct hydrogens and carbons, respectively, rather than an extensively split lineshape, consistent with the hexagonal structure we obtained via quantum calculations. While fine pattern chemical shift distinction can easily be concealed in a pattern with significant broadening, we obtained ^{13}C and ^{2}H linewidths substantially less than the previously reported lineshapes. Careful inspection of previously published ^{2}H and ^{13}C MAS spectra suggests that the apparently increased multiplicity due to the monoclinic distortion of the supramolecular structure is not a result of an overall lattice inequivalence, but rather, is most likely attributed to a mis-setting of the magic angle. Direct comparison of our central frequency lineshapes for both ^{2}H and ^{13}C to the previously published lineshapes can be found in Figure 3.12.

In an attempt to confirm our quantum mechanical studies that indicated an overall hexagonal structure of the CAM complex, we performed powder x-ray diffraction on the complex formed at ambient temperatures and pressures (Figure 3.13). The overall simplicity of the powder pattern seems to confirm a highly symmetric structure, rather than the monoclinic distortion previously reported[4] for single crystals grown under

hydrothermal conditions. We can confirm by a combination of NMR and x-ray diffraction studies that the solid obtained under ambient conditions is distinct from the single-crystal obtained via hydrothermal synthesis.

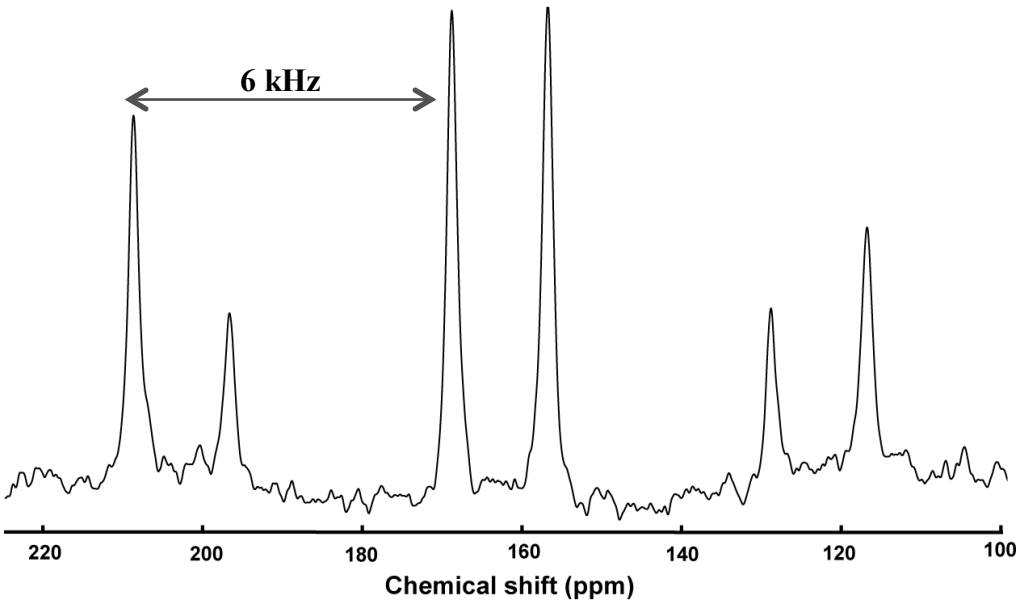

Figure 3.11. ^{13}C cross-polarization SSNMR spectrum of the CAM complex with a 6 kHz MAS rotational frequency.

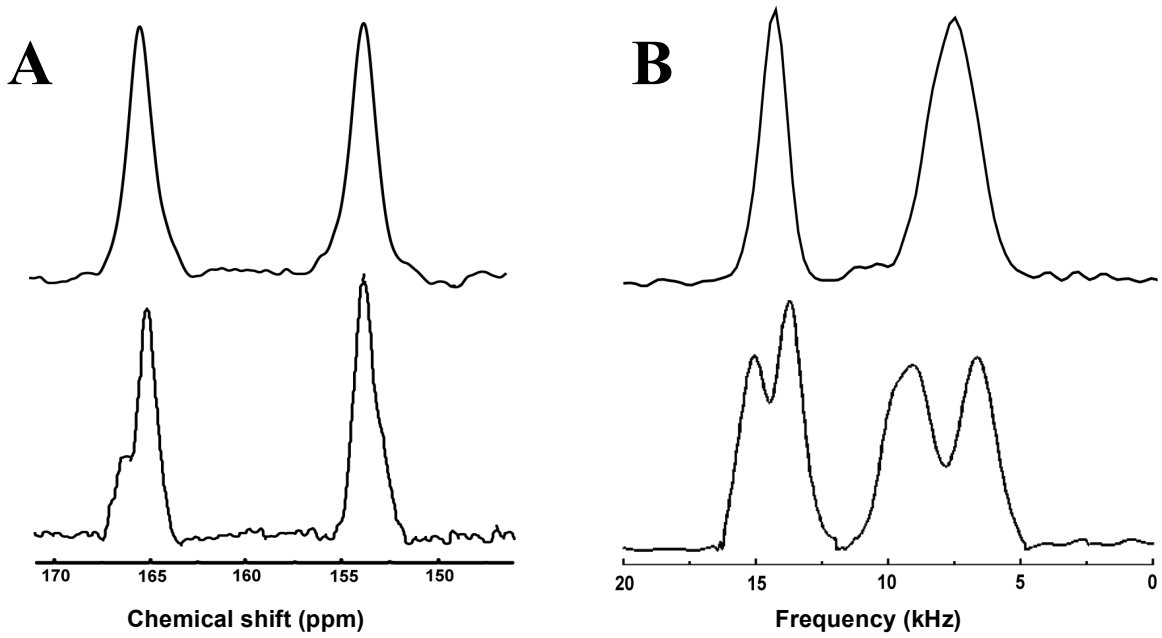

Figure 3.12. Comparison of our (a) ^{13}C and (b) ^{2}H experimental SSNMR central frequencies (top) to those previously published[15] (bottom) for the CAM complex.

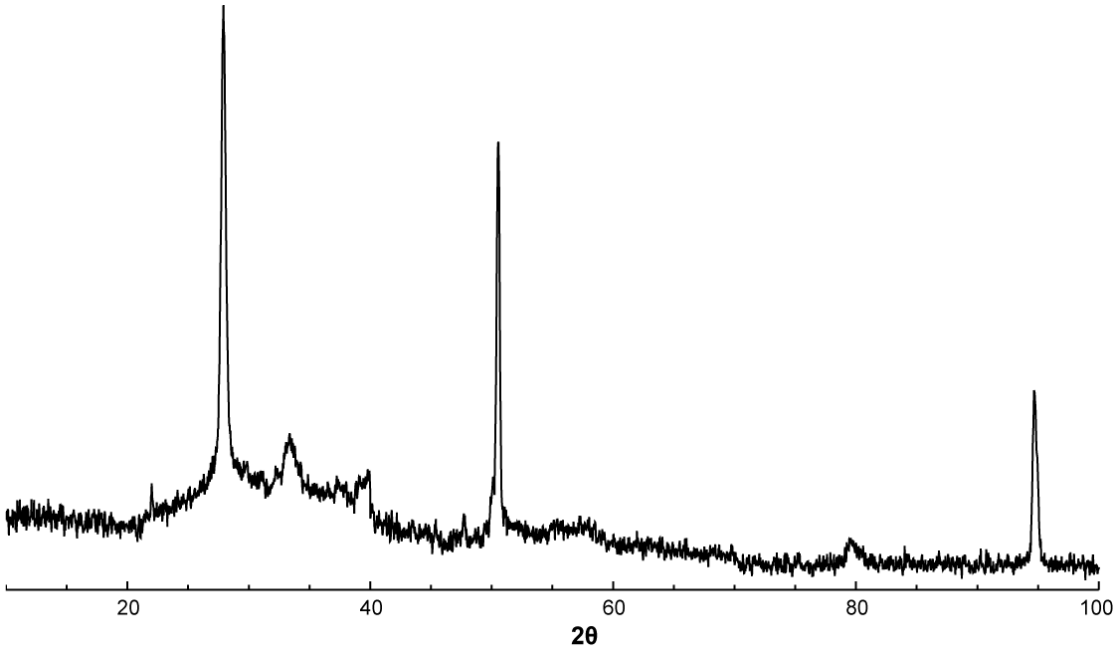

Figure 3.13. Powder x-ray diffraction of the CAM complex

3.5 Conclusion

The room temperature complex formed between melamine and cyanuric acid in neutral conditions appears to be unlike the single crystals formed previously via hydrothermal conditions. The crystal structure of the sample formed under ambient conditions appears to be a highly symmetric two-dimensional layered hexagonal structure, rather than the monoclinic structure previously reported. We have confirmed this with extensive quantum mechanical methods as well as ^{1}H, ^{2}H, and ^{13}C SSNMR. Our solid-state NMR results vary greatly from those previously reported in that our results show two unique proton and carbon environments indicating a highly symmetric adduct rather than a distinction between the proton and carbon atoms that are on-axis versus those that are off-axis. The overall simplicity of our powder X-ray diffraction studies seem to confirm an overall hexagonal structure rather than the monoclinically distorted crystal structure reported previously. Further investigation of the previously published SSNMR data suggests that the magic angle was mis-set resulting in what looked like an increased multiplicity in the lineshape, leading the original investigators to assume that this was sufficient evidence of the monoclinic lattice inequivalence as previously reported by Ranganathan.

Additionally, the relative magnitudes of the deuterium quadrupolar coupling constants directly correlate to the length of the hydrogen bonds as reported by Ranganathan, *et al*. Very strongly shifted deuterium and proton chemical shift data of the cyanuric acid N-H---N bond in the complex agrees with the relatively small experimentally calculated C_Q value and the small bond length as reported.

References:

1. Xin, H.; Stone, R. *Science*. 322(5906), 1310-1311.
2. Wang, Y.; Wei, B.; Wang, Q. *J. Cryst. Spect. Res*. **1990**, *20*(**1**), 79-84.
3. Perdigão, L.M.A.; Champness, N.R.; Beton, P.H. *Chem. Comm*. **2006**, 538-540.
4. Ranganathan, A.; Pedireddi, V.R.; Rao, C.N.R. *J. Am. Chem. Soc*. **1999**, *121*, 1752-1753.
5. Seto, C.T.; Whitesides, G.M. *J. Am. Chem. Soc*. **1993**, 115, 905-916.
6. Finkel'shtein, A.I.; Rukevich, O.S. *Journal of Applied Spectroscopy*. **1983**, 38, 327-30.
7. Berglund, B.; Carson, D.G.; Vaughan, R.W. J. Chem. Phys. **1980**, *72*(**2**), 824-827.
8. Zhao, Z.; Dvorak, M.; Silvernail, C.; Belot, J.; Harbison, G.S. *Solid State Nuclear Magnetic Resonance*. **2002**, *22*, 363-372.
9. Zhao, X.; Rossi, P.; Barsegov, V.; Zhou, J.; Woodford, J.N.; Harbison, G.S. *J. Mol. Str*. **2006**, *790*, 152-159.
10. Wiebenga, E.H. *J. Am. Chem. Soc*. **1952**, *74*, 6156-7.
11. Hughes, E.W. *J. Am. Chem. Soc*. **1941**, *63*, 1737-52.
12. Herzfeld, J.; and Berger, A.E. *J. Chem. Phys*. **1980**, *73*(**12**), 6021-6030.
13. Ditchfield, R. *J. Chem. Phys*. **1976**, *65*(**8**), 3123-33.
14. Prins, L.J.; Neutenboom, E.E.; Paraschiv, V.; Crego-Calama, M.; Timmerman, P.; Reinhoudt, D.N. *J. Org. Chem*. **2002**, *67*(**14**), 4808-4820.
15. Damodaran, K., Sanjayan, G. J., Rajamohanan, P. R., Ganapathy, S., and Ganesh, K.N. *Org. Lett.* **2001**, *3*(**12**).
16. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comp. Chem*. **1993**, 14, 1347.
17. Duer, Melinda J. *Introduction to Solid-state NMR Spectroscopy*. **2004,** Blackwell Publishing, Ltd.
18. Xiongjian, Wu. Thesis for Ph.D.
19. Andrew, E.R. and Eades, R.G. Discuss. *Faraday Soc*. **1962**, 34(38).

20. Berglund, Bo; Carson, Douglas G.; Vaughan, Robert W. *J. Chem. Phys*. **1980**, *72*(**2**), 824-827.
21. Goedkoop, J.A. and MacGillavry, C.H. *Acta Crystallogr*. **1957**, *10*(**125**).
22. Currie, M. and Speakman, J.C. *J. Chem. Soc.* **1970**.
23. Berglund, Bo and Vaughan, Robert W. *J. Chem. Phys*. **1980**, *73*(**5**), 2037-2043.

Chapter 4

Thermodynamic Isotope Effects in the NMR Spectra of Partially Deuterated Amino Acids

Summary

A common assumption in NMR is that isotope labeling is not perturbative. This is often not so for deuterium. Theory and experiment demonstrate that fully deuterated ($–ND_3^+$) and partially deuterated ($–NDH_2^+$) amino groups in three distinct crystalline amino acids have substantially different ^2H spectra, because of partitioning of deuterium out of strong hydrogen bonds, which disrupts averaging by thermally activated hops of the group. Thermochemical quantum calculations allow us to compute deuterium partitioning between the three sites, and reproduce the effect. Such computations predict similar but somewhat smaller effects in partially deuterated methyls, particularly if one site is sterically crowded. Accurate computation of methyl and amino group spectra and spin relaxation requires accounting for these effects.

4.1 Introduction

Deuterium solid-state NMR is powerful technique for investigating molecular dynamics[27-30] (see Chapter 2) and gaining structural insight[30-31] (see Chapter 3). This is mainly due to the nature of deuterium's electronic quadrupole interaction. This orientation-dependent quadrupole interaction arises from the interaction of the nuclear electric quadrupole moment with the electric field gradient. Hence, the quadrupole interaction acts as a very convenient structural probe due to its dependence on electron density. Furthermore, its relatively small magnitude (ranging from 140-220 kHz in organic compounds) is small enough that is does not completely distort the NMR spectrum unlike quadrupole lineshapes of nuclei with significantly larger electric quadrupole moments. However, the magnitude of the quadrupole interaction remains large enough that it can still act as a sensitive probe of the electronic environment. Additionally, the relative simplicity of deuterium's spin dynamics compared to other quadrupole nuclei allow for complete analysis of density matrices in complicated pulse sequences and relaxation processes. Another simplification is that the 2nd-order quadrupolar contribution to the lineshape is often insignificant for deuterium compared to its $I > 1$ counterparts. This adds to the overall ease of theoretical descriptions and lineshape simulations.[32-33]

4.2 Theory

4.2.1. The thermodynamic isotope effect. The structural, thermochemical, and geometric effects upon isotopic substitution have been well described.[1,3-4,22-25] Recently,

many experimental schemes have been developed that use isotopic substitution as a tool in understanding biological reactions and structures. Substituting in a heavier isotope increases the mass and changes the zero point energy, thus perturbing the equilibrium and kinetics.[23] Early and extensive study of the elementary hydrogen molecules (H_2, HD, and D_2) revealed that many thermochemical properties including heats of fusion and vaporization, molar volumes, and others changed upon substitution of the heavier isotope.[2,25] These early studies also reveled that isotopic effects were small except for those involving hydrogen and the effects decrease substantially with increasing atomic weight.

The isotopic exchange process was generalized by Harold Urey[2] by:

$$aA_1 + bB_2 \rightleftharpoons aA_2 + bB_1 \tag{4.1}$$

in which the (2) subscript indicates the heavier isotope. From this, the equilibrium constant could be written as a ratio of the partition functions.

$$K = \frac{\left(Q_{A_2} / Q_{A_1} \right)^a}{\left(Q_{B_2} / Q_{B_1} \right)^b} \tag{4.2}$$

It was noted through experimental and computational work that the ratio of the partition functions decreases with increasing temperature, indicating that the effect would be more pronounced for experimental or calculations being completed at room temperature or lower. Additionally, kinetic amide isotopic effects have been used repeatedly for understanding protein folding by directly assaying transition state structures.[1] This effect can be used to extract thermochemical data from the folding and unfolding of proteins by:

$$\Delta\Delta G^{D-H} = -RT \ln \frac{K_{eq}^{D}}{K_{eq}^{H}} \tag{4.3}$$

The equilibrium constants can be written as a function of the isotopic exchange folding and unfolding rate constants:

$$\Delta\Delta G^{D-H} = -RT \ln \frac{k_f^D / k_u^D}{k_f^H / k_u^H} \tag{4.4}$$

Geometric isotope effects in which substitution of heavier isotopes can perturb equilibrium molecular geometry can also occur.[22-24] Geometric perturbations of hydrogen bonds can affect entire molecular structures and physical properties. Known as Ubbelohde[9,10] effects by the solid-state physics community, these effects have been studied extensively in ferroelectric materials and by NMR, diffraction studies, and microwave spectroscopy in liquids and gases.[9-11] More recently, quantum chemical calculations and molecular dynamic simulations have been used to study geometric effects.[8] In a hydrogen bond, denoted A—X· · ·B where X = ^{1}H or ^{2}H, two observations are made upon substitution with the heavier deuterium isotope:

(1) the intramolecular A—X bond shrinks and

(2) the intermolecular X· · ·B elongates.

The shrinkage of the intramolecular bond can be attributed to the decrease in the zero-point energy, which in turn reduces the anharmonicity of the hydrogen atom potential. The theoretical basis for the elongation of the intermolecular bond has so far remained illusive, though attempts to describe the interaction by decomposing the interaction into its energetic parts by computational methods have been attempted. This entails describing

the interaction as a sum of electrostatic, exchange-repulsion, polarization, and charge-transfer and other energies by computational methods such as LMO energy decomposition analysis (LMOEDA),[5] constrained space orbital variation (CSOV),[6] and natural energy decomposition analysis using natural bond orbitals (NBO).[7,8]

In hydrogen bonded solids, substitution of a deuteron in a hydrogen bond results in an overall lengthening of the hydrogen bond and a substantial increase in the phase transition temperature. Extensive NMR studies in the solution-state have demonstrated perturbations of magnetic properties with H/D exchange. As shown in chapter 1, the chemical shielding tensor has an inherent dependence on the nuclear coordinates. Nuclear coordinates are altered upon isotopic substitution, resulting in a measurable change in the chemical shift. NMR studies of intermolecular hydrogen bonding systems have proven to be difficult due to fast proton and hydrogen bond exchange in the solution-state. Some efforts were made to circumvent these issues by using a liquefied, low-freezing Freon mixture (CDF_2Cl–CDF_3) as a solvent in slow exchange regimes.[11,12] In 2001, Lorente *et al.* used ^{1}H, ^{2}H and ^{15}N SSNMR techniques to measure H/D exchange effects in hydrogen-bonded solids. These studies resulted in measurable alterations of the chemical shifts and changes in the magnitude of ^{15}N–$^{1}H/^{2}H$ dipolar coupling.

4.2.2. Deuterium solid-state NMR. It was previously shown that the total NMR Hamiltonian can be written as a sum of terms including the Hamiltonian describing the internal interactions and the Hamiltonian describing the external interactions. The external Hamiltonian is dominated by the Zeeman interaction. For nuclei with $I \geq 1$, the internal interaction is dominated by the quadrupolar interaction. Therefore, the total NMR Hamiltonian for powder static ^{2}H NMR can be expressed as:

$$\mathcal{H}_{powder}^{^2H} = \mathcal{H}_{Zeeman} + \mathcal{H}_{Quad} \tag{4.5}$$

When the static magnetic field is significantly larger than the magnitude of the quadrupolar interaction, the secular approximation can be used to rewrite the total powder static Hamiltonian.[32] It is of the form:

$$\mathcal{H}_{total}^{secular} = -\omega_I \left(1 - \sigma_{iso}\right) I_z + \frac{1}{3}\omega_Q \left(3I_z^2 - \mathbf{I}\cdot\mathbf{I}\right) \tag{4.6}$$

where σ_{iso} is the isotropic component of the chemical shielding, ω_I is the Larmor frequency, and ω_Q is the quadrupolar frequency of the form:

$$\omega_Q = \frac{3e^2qQ}{4I\left(2I-1\right)\hbar} \tag{4.7}$$

The $3e^2qQ/\hbar$ is the quadrupole coupling constant and eQ the nuclear quadrupole moment. Deuterium is $I = 1$ and is therefore a three state system with two degenerate spin transitions of $+1 \leftrightarrow 0$ and $0 \leftrightarrow -1$. In the presence of the quadrupolar interaction, this degeneracy is lifted and the result is two transitions at $\pm\,\omega_Q$ about the Larmor frequency as shown in figure 4.1.[14]

However, the assumption that only one orientation of the EFG tensor with respect to the static magnetic field exists is incorrect for a polycrystalline powder sample. In these types of samples, the orientation of the EFG tensor is random with respect to the magnetic field. The quadrupolar frequency can therefore be expressed in terms of the components of the Euler solid angle in the Wigner rotation matrix. The Wigner rotation

matrix relates laboratory frame tensor components to the principal axis system of the EFG tensor.[14,32]

$$\omega_Q = \frac{3e^2qQ}{8I(2I-1)\hbar}\left[\left(3\cos^2\beta - 1\right) + \eta\sin^2\beta\cos 2\gamma\right] \tag{4.8}$$

If the principal *z*-axis of the deuteron's EFG tensor is parallel to the static field, then β = 0 and the resulting spectrum resembles the generalized spectrum as shown in figure 4.1b. However, a powder sample has a random distribution of EFG tensors relative to the field, resulting in random values for β and γ. What results is a lineshape that is the superposition of all these orientations. This lineshape is known as the powder pattern or Pake doublet (figure 4.2). As evident in figure 4.2, there exists three singularities in the spectrum that correspond to specific orientations of the EFG tensor:

$$\begin{aligned} &\pm\omega_Q && \text{when} \quad \beta = 0^\circ \\ &\mp\frac{\omega_Q}{2}(1+\eta) && \text{when} \quad \beta = 90^\circ,\ \gamma = 0^\circ \\ &\mp\frac{\omega_Q}{2}(1-\eta) && \text{when} \quad \beta = 90^\circ,\ \gamma = 90^\circ \end{aligned} \tag{4.9}$$

The doublet nature of the lineshape is a direct result of the two possible spin transitions. Furthermore, the splitting of the maxima is equal to ¾ the quadrupole coupling constant. This illustrates the experimental efficiency of deuterium; the relatively small C_Q value for deuterium makes the width of the resulting Pake doublet experimentally much easier to deal with. Most other quadrupolar nuclei have considerably larger quadrupole coupling constants. Pake patterns of nuclei with significantly larger quadrupole coupling constants

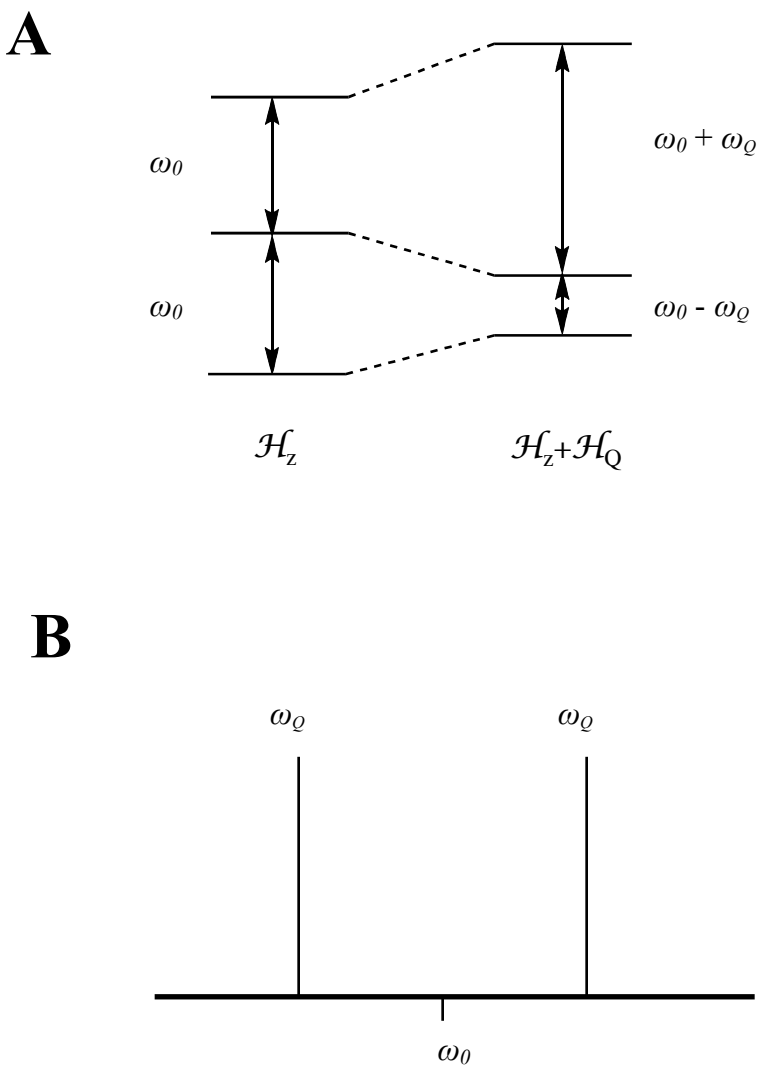

Figure 4.1. (A) First-order quadrupolar effects on the energy levels of a spin $I = 1$ nucleus. The spin transitions are degenerate under the Zeeman interaction but the degeneracy is lifted under the addition of the quadrupolar interaction. (B) The resulting spectrum will be two frequencies at $\omega_0 \pm \omega_Q$.[13]

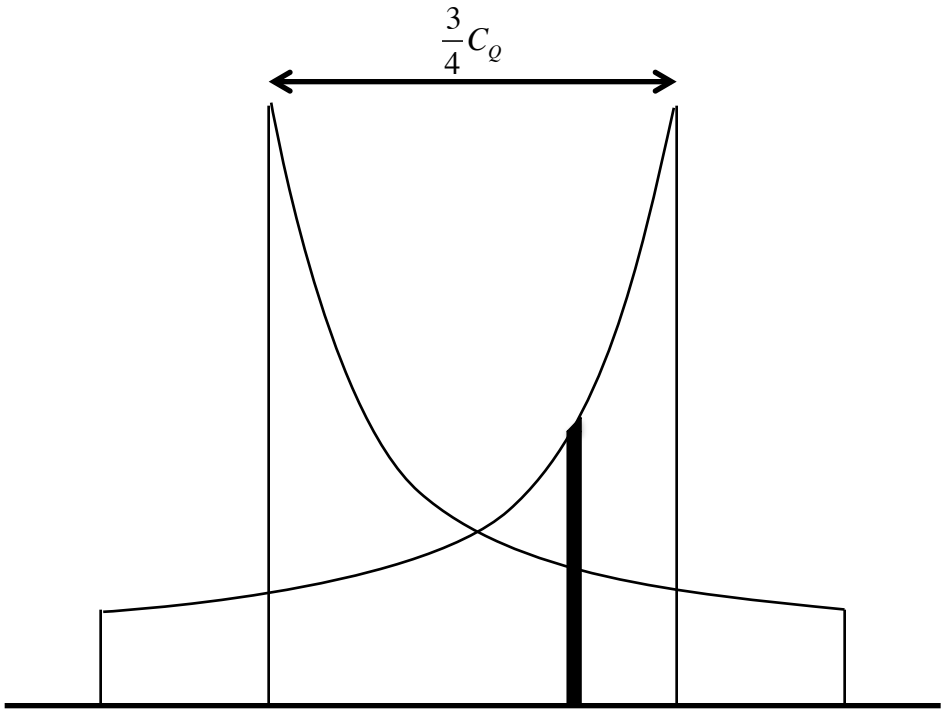

Figure 4.2. Generalized form of the ^{2}H static powder or Pake pattern. The mirrored superposition of the two Pake patterns is due to the two allowed spin transitions. The splitting of the maxima is equal to $\frac{3}{4}C_Q$. The shaded region represents one orientation of the EFG with the field (i.e. a specific set of [β, γ]) and the overall lineshape is the superposition of all these orientations.

result in a significantly wider lineshape that is often difficult or impossible to fit in one spectral window.

While the relatively small quadrupole constant for deuterium results in a narrow spectral lineshape (~200 kHz) compared to other quadrupole nuclei, the lineshape is significantly broadened compared to non-quadrupole nuclei. This can make obtaining a non-distorted experimental lineshape difficult. For this reason, NMR experiments of quadrupole nuclei — especially deuterium — typically utilize a specialized pulse sequence called the quadrupole echo (or solid echo) pulse sequence (figure 4.3). It is of the form $[(\pi/2)_x - \tau_1 - (\pi/2)_y - \tau_2 - \text{acquire})]$ where τ_1 and τ_2 are approximately equal. In practice, however, τ_2 is usually adjusted slightly so that data acquisition begins directly at the echo maximum.

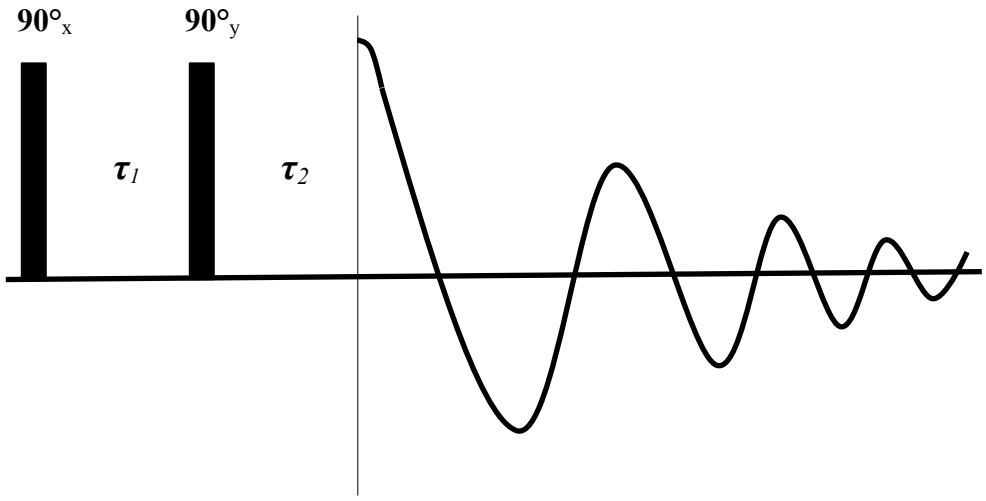

Figure 4.3. The quadrupole echo pulse sequence. τ_1 and τ_2 are approximately equal, but τ_2 is set so that data acquisition begins at the maximum of the FID.

4.2.3. Ab initio calculations of the electric field gradient. For *ab initio* calculations, each component of the electric field gradient tensor is treated as a derivative of the energy:[14]

$$q_{\alpha\beta\gamma}(N) = \frac{\partial E(Q)}{\partial Q_{\alpha\beta\gamma}}\Big|_{Q=0} \tag{4.10}$$

In this expression, Q represents the quadrupole tensor with individual components $Q_{\alpha\beta\gamma}$. These tensors can then be diagonalized to provide the principal components. These principal components are then used to calculate asymmetry parameters (η) and quadrupole coupling constants (C_Q) as explained in previous sections. The quadrupole coupling constant is most often calculated theoretically by evaluating the electric field gradient at each nucleus using *ab initio* methods. The electric field for any point is given as

$$\mathbf{V} = \frac{1}{4\pi\varepsilon_0}\sum \frac{q}{\left|\mathbf{R}-\mathbf{r}_i\right|^3}\left(\mathbf{R}-\mathbf{r}_i\right) \tag{4.11}$$

Ab initio methods usually compute field gradient iteratively as a function of the nuclear coordinates as $V_{ii} = \partial E_i/\partial i$ where E is the electric field and i is the chosen coordinate.[21]

Some careful attention must be paid to the units. *Ab initio* programs calculate and report EFG tensor elements in atomic units: Hartree/Bohr^2. Experimentally obtained coupling constants as obtained though lineshape fitting are usually reported in frequency units (Hz or kHz). Therefore, we must take into account the appropriate conversion factor. As explained in Chapter 1, we pick the largest absolute value of the principal quadrupole

tensor component to be q_{zz}. To convert this to a coupling constant, we must employ the following equation:

$$C_Q = Cq_{zz}Q \tag{4.12}$$

where Q is the deuterium nuclear quadrupole moment (0.286 fm^2) and C is the unit conversion factor previously mentioned. C has been previously calculated[15] and is expressed as:

$$\begin{aligned} C &= \frac{\left(4.35974381\times10^{-18}\,\text{J/Hartree}\right)\left(10^{15}\,\text{m/fm}\right)^2}{\left(5.2917720839\times10^{-11}\,\text{m/Bohr}\right)^2\left(6.62606876\times10^{-34}\,\text{J}\cdot\text{sec}\right)} \\ &= 2.34964781\frac{\text{Mhz}\cdot\text{Bohr}^2}{\text{fm}^2\cdot\text{Hartree}} \end{aligned} \tag{4.13}$$

which provides the theoretically-determined quadrupole coupling constant in frequency units.

4.3 Methods

4.3.1 Sample preparation. Glycine, L-alanine, and deuterium oxide were purchased from Sigma Aldrich. L-Histidine monohydrochloride monohydrate was purchased from Avocado Research Chemicals, Ltd. Glycine hydrochloride was produced by dissolving a saturating amount of glycine in warm 12 M hydrochloric acid. This solution was stored at 3°C to promote crystallization. The crystals were collected via filtration, ground to a fine consistency, and then dried overnight in a warm oven. A homemade Schlenk line was assembled and equipped with a round bottom Schlenk flask. A pre-determined quantity

of solid amino acid was added to the flasks and then purged with nitrogen gas. Approximate 10%, 30%, 60%, and 100% deuterated amino acids were obtained by rapidly dissolving the amino acid in the appropriate amount of $^{1}H_2O$ and $^{2}H_2O$, followed by rapid vacuum evaporation. To ensure ~99% deuteration, the samples representing the perdeuterated samples were dissolved in $^{2}H_2O$ followed by vacuum evaporation. This process was repeated at least three times. The flasks remained sealed to prevent isotopic exchange. The sealed flasks were taken directly for analysis via solid-state NMR, again minimizing contact with the atmosphere to prevent any isotopic exchange.

4.3.2 Solid-state NMR experiments. The deuterated samples of the amino acids were placed directly in sample rotors for analysis via SSNMR. The static ^{2}H NMR spectra were obtained at 14 T on a Bruker Avance spectrometer with a deuterium frequency of 92.102 MHz. The $\pi/2$ pulse was determined to be 4.1 μs for all three samples. A quadrupole echo sequence $[(\pi/2)_x - \tau_1 - (\pi/2)_y - \tau_2 - \text{acquire})]$ was employed with high power proton decoupling on a double-tuned probe. The delays used were 50 μs (alanine) and 30 μs (glycine HCl and histidine·HCl·H_2O) between the two pulses of the quadrupole echo and a 30 μs and 20 μs delay between the last quadrupole pulse and the acquisition period for alanine and glycine HCl/ histidine·HCl·H_2O, respectively. The recycle delay was kept deliberately short at 0.5 s to ensure other exchangeable deuterons with longer relaxation times were omitted from the spectrum (in histidine hydrochloride monohydrate). This recycle delay combined with a dwell time of 4 μs resulted in a spectral window of ±125 kHz.

4.3.3 Computational studies. Neutron and x-ray diffraction data were used to generate multiple unit cells of L-alanine,[16,17] histidine monohydrochloride monohydrate,[18] and

glycine hydrochloride.[19] These unit cells were then truncated down to a minimum number of adjacent molecules representing a single amino acid molecule caged by adjacent molecules of the amino acid, the minimum requirement being that all hydrogen bond partners of the central (caged) amino acid's $-NH_3$ are accounted for. These truncated unit cells were optimized at the B3LYP/6-311++G(2d,p) level of theory. In these optimizations, all molecular coordinates not belonging to the caged amino acid were frozen.

Hessian analysis was performed on the optimized coordinates and zero-point harmonic energies obtained for the eight isotopomers of each amino acid. Electric field gradients were calculated for the three deuterons on the amino acid. Boltzmann distribution analysis was performed based on the calculated differences in zero-point energies to obtain population factors for each isotopomer. The field gradients were averaged over the $3–NHD_2$, $3–NH_2D$, and the $1–ND_3$ isotopomers and weighted by the Boltzmann probabilities. The resulting weighted field gradients were then diagonalized to provide the principal tensor components averaged over the population of the isotopomers. These tensor components were then used to calculate η and C_Q theoretical values for the $–ND_3^+$ species, the $–NDH_2^+$ species, and the $–ND_2H^+$ species.

4.4 Results and Discussion

Optimized amino acid structures used for zero-point energy and field gradient calculations are provided in figures 4.4 through 4.6. Wilson[16] *et al*. first used neutron diffraction data to determine the structure of L-alanine. They determined the three amine

hydrogen bond were relatively symmetric with $r_{N...O}$ 2.836(4), 2.808(3), and 2.861(3) Å, respectively. Experimentally determined hydrogen bond angles ∠N-H…O were determined to be 162.2(6)°, 168.6(5)°, and 162.0(5)°. This is in good agreement with the results of our geometry optimization: $r_{N...O}$, 2.853, 2.815, 2.833 Å and ∠N-H…O, 160.92°, 168.14°, 163.72°.

Al-Karaghouli *et al*. used neutron diffraction to determine the glycine hydrochloride crystal structure. The chloride ion is used to link three glycine atoms into parallel layers, forming strong hydrogen bonds with two of the hydrogen atoms on the amine group of the glycine molecule with $r_{N...Cl}$ = 3.189(1) and 3.140(1) Å and ∠N-H…Cl, 170.8(1)° and 166.4(2)°. The third hydrogen atom on the amine forms two weak bifurcated hydrogen bond interactions: one with the chloride ion of the same molecule and the other with the carbonyl oxygen of a neighboring glycine molecule. In the first interaction, $r_{N...Cl}$, 3.300(1) Å and ∠N-H…Cl, 126.7°. The second interaction has $r_{N...O}$, 2.992(1) Å and ∠N-H…O, 131.6(2)°. Again, our computational results are in decent agreement with these previously reported experimental results. The optimized strong hydrogen bond lengths $r_{N...Cl}$ are 3.112 and 3.149 Å with ∠N-H…Cl bond angles of 176.90° and 168.42°. The bifurcated hydrogen bonds have interatomic distances of 3.290 Å for the $r_{N...Cl}$ interaction and 3.060 Å for the $r_{N...O}$ interaction. These correspond to angles of 120.15° and 137.86° for the ∠N-H…Cl and ∠N-H…O angles, respectively.

Fuess *et al*. reported the crystal structure of histidine hydrochloride monohydrate using neutron diffraction. They found that the histidine molecules are linked together by hydrogen bonds of the amine group with the oxygen atom of an adjacent water molecule and two different chlorine atoms. Their experimental bond lengths were determined to be

3.193 and 3.198 Å for the $r_{N\cdots Cl}$ interactions and 2.790 Å for the $r_{N\cdots O}$ interaction. These correspond to bond angles of 169.0° and 149.4° for the ∠N-H…Cl angles and 168.3° for the ∠N-H…O angle. In relatively close agreement, our computed bond distances are 3.192 and 3.198 Å for $r_{N\cdots Cl}$ and 2.774 Å for $r_{N\cdots O}$, corresponding to bond angles of 169.0 and 149.45° for ∠N-H…Cl and 168.0° for ∠N-H…O.

The spectral comparison of the 10% and 100% deuterated amino acids can be found in figures 4.4 through 4.6. Boltzmann population-averaged field gradients from the computational study as well as harmonic zero-point energies from the Hessian analysis can be found in Table 4.1. Direct comparison of experimental and theoretical magnetic and electronic molecular parameters is tabulated in Table 4.2. In all three amino acids, experimentally determined field gradient tensors depend on the level of deuteration in that the asymmetry parameter decreases with increasing deuteration. There was no observable effect on the quadrupole coupling constant. Observation of the theoretically determined asymmetry parameter and quadrupole coupling constant show the same trend, but do not reproduce the experimental parameters exactly.

As mentioned previously in this chapter, the thermodynamic isotope effect and the relative smaller size of a deuteron in comparison to a proton dictates that any deuteron should preferentially partition itself into weaker hydrogen bonds. For this reason, a single deuteron in a $–NDH_2$ group will preferentially partition into the weakest hydrogen bond more than $1/3^{rd}$ of its time; likewise, it will partition itself into the strongest hydrogen bond less than $1/3^{rd}$ of its time. This phenomenon is represented by the theoretically calculated Boltzmann populations as reported in Table 4.1 and arises from the difference in zero-point energies of the isotopomers. Glycine·HCl has two strong

hydrogen bond interactions and a weak bifurcated hydrogen bond interaction. Additionally, the strong hydrogen bonds differ in length by 0.036 Å. As expected, the calculated Boltzmann populations reflect the differences in hydrogen bond lengths, with a single deuteron population probability of 0.2702 for the strongest hydrogen bond (3.112 Å), 0.3365 for the second strongest hydrogen bond (3.149 Å), and 0.3933 for the weakest hydrogen bond (3.290 Å). This same effect is present in the other amino acids of study as well as the other isotopomers but is lessened due to the similarities in hydrogen bond strengths.

4.5 Conclusion

Isotopic labeling has been used extensively in the NMR studies of various materials. While the assumption that isotopically labeling is non-perturbative in nature is usually a good assumption, this study and previously reported studies shows that this assumption is not valid for the isotopic substitution of hydrogen. Additionally, this work also shows that the non-perturbative nature of isotopically labeling can be an especially bad assumption when determining experimental NMR parameters. This study shows that this observation can most likely be attributed to the thermodynamic isotope effect. We have obtained similar trends upon selective deuteration in NMR parameters for experimental results as well as those parameters obtained theoretically, though careful consideration of averaging the deuteron along vibrational modes should be considered to achieve exact fits in magnitude for the NMR parameters.

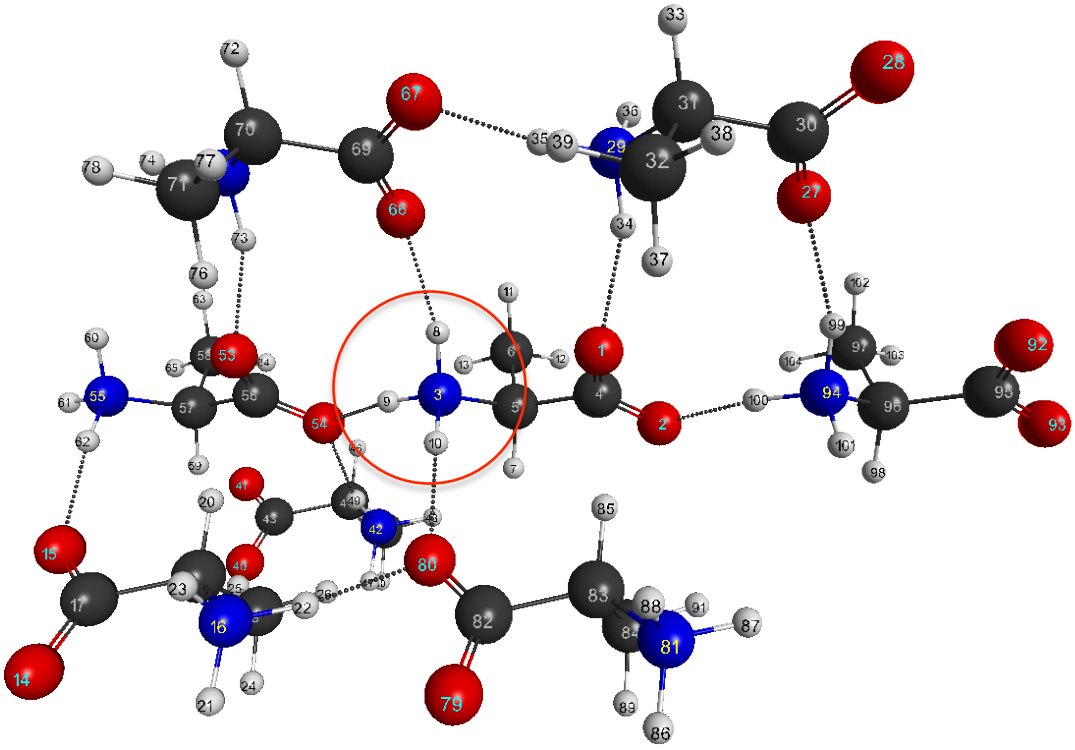

Figure 4.4. Truncated structure of multiple alanine unit cells used for the computational studies. The caged alanine molecule with the amine hydrogen bonds of interest is indicated by atom numbers 1–12.

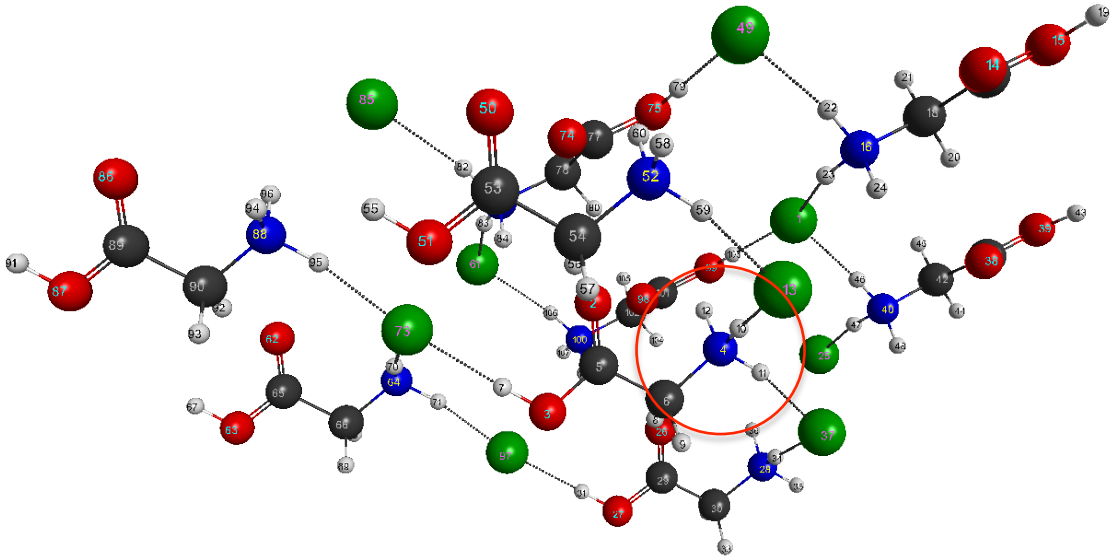

Figure 4.5. Truncated structure of multiple glycine·HCl unit cells used for the computational studies. The caged glycine molecule with the amine hydrogen bonds of interest is indicated by atom numbers 1–12. Notice the bifurcated hydrogen bond of the 12H a chlorine atom of the same molecule and the carbonyl oxygen of an adjacent molecule.

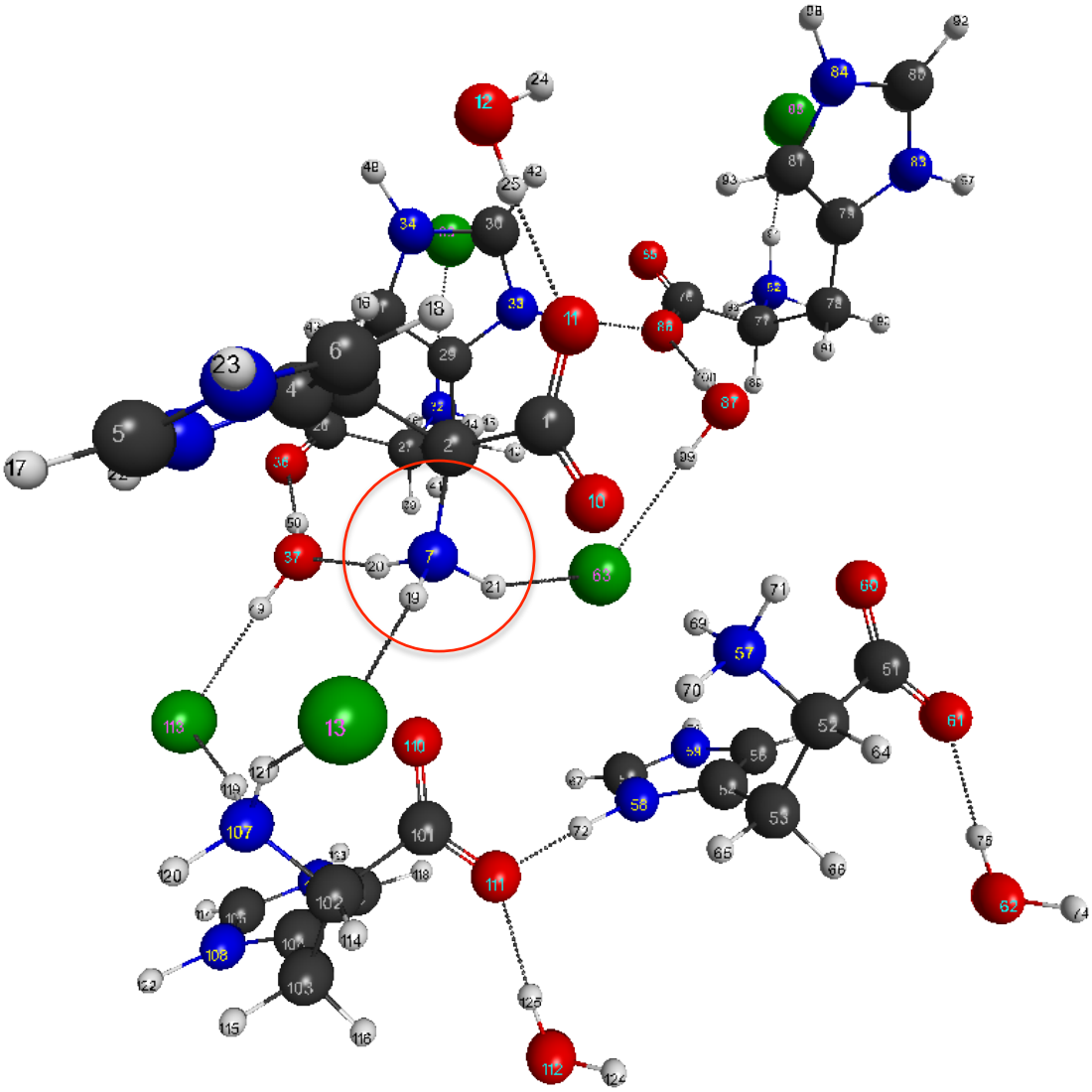

Figure 4.6. Truncated structure of multiple histidine·HCl·H_2O unit cells used for the computational studies. The caged histidine molecule with the amine hydrogen bonds of interest is indicated by atom numbers 1–23.

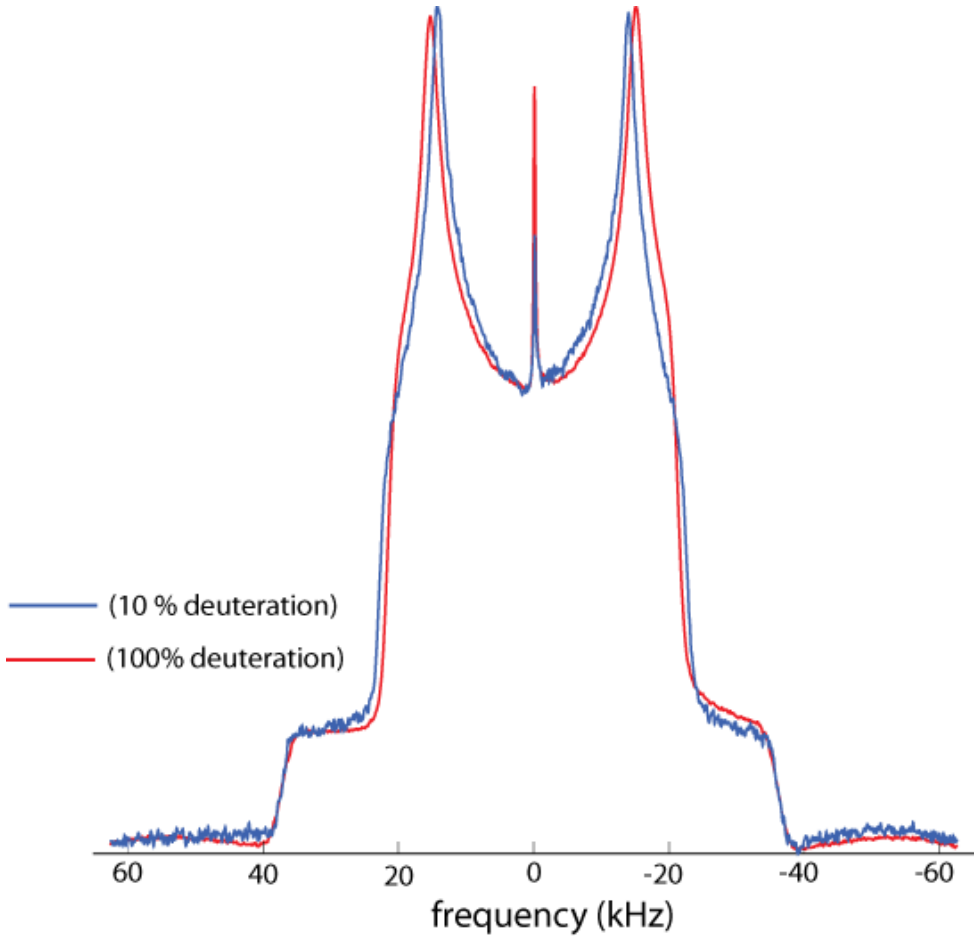

Figure 4.7. Comparison of static ^{2}H NMR spectra of 10% (blue) and 100% (red) deuterated L-alanine.

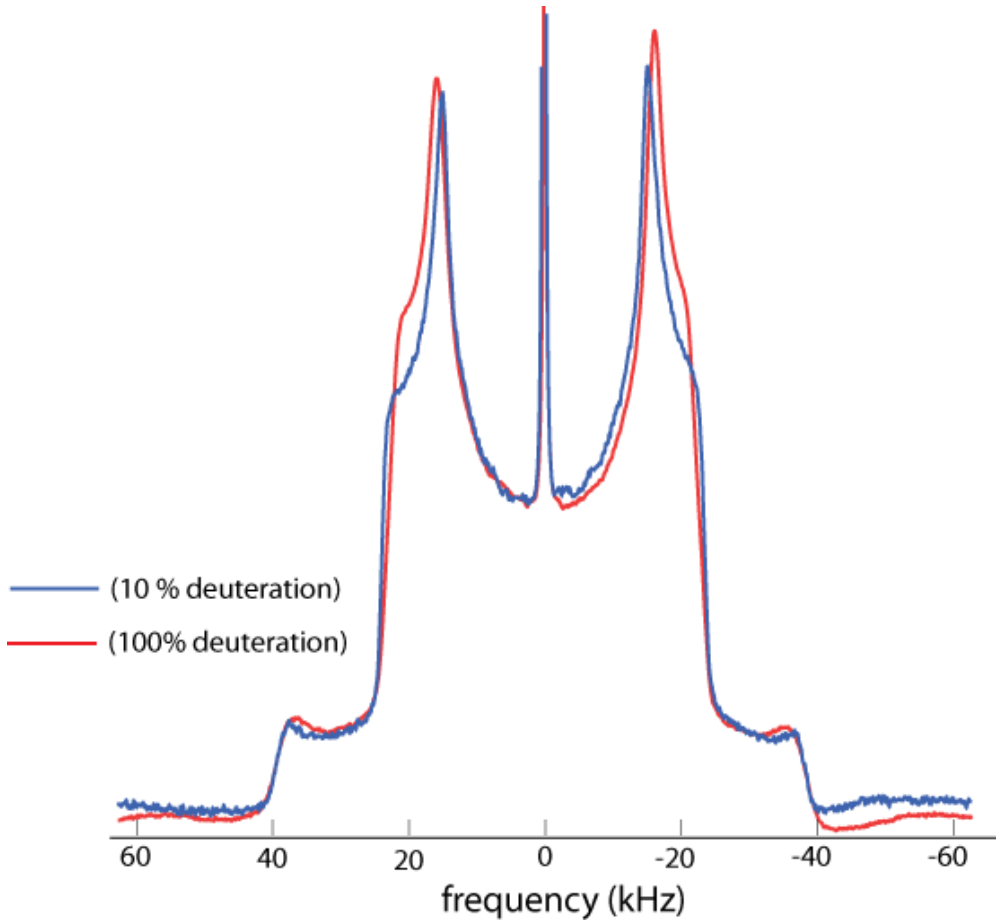

Figure 4.8. Comparison of static ^{2}H NMR spectra of 10% (blue) and 100% (red) deuterated glycine hydrochloride.

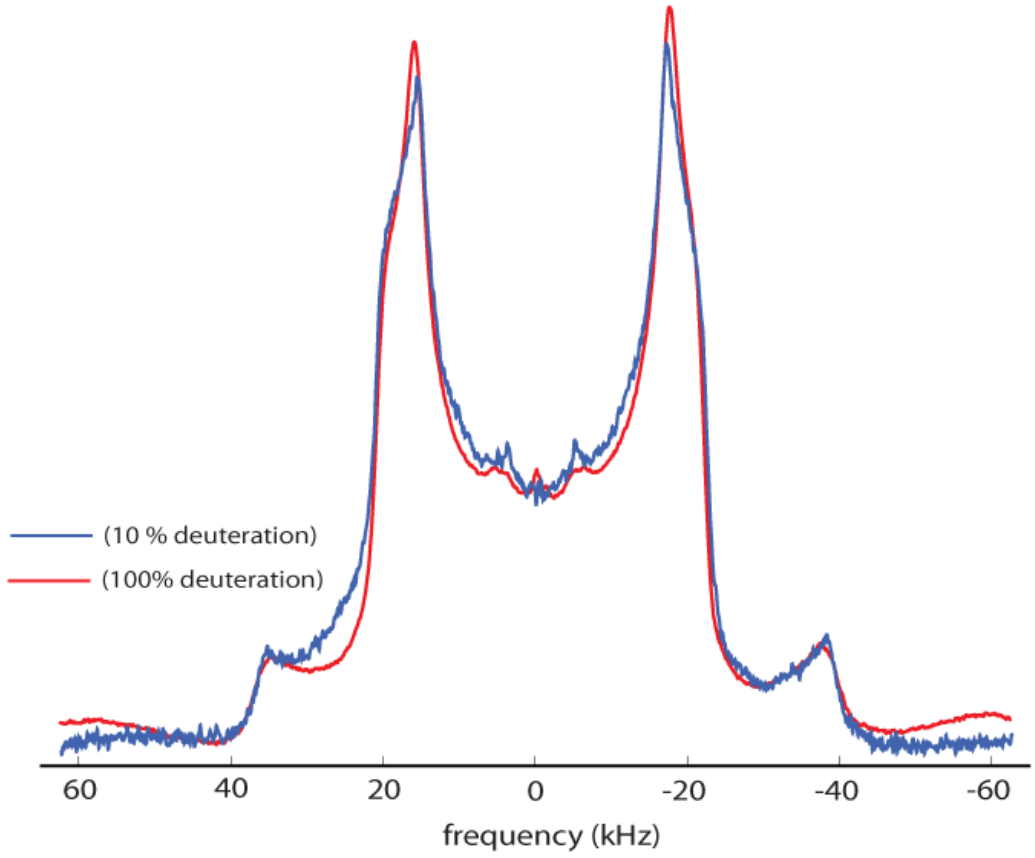

Figure 4.9. Comparison of static ^{2}H NMR spectra of 10% (blue) and 100% (red) deuterated histidine monohydrochloride monohydrate.

ISOTOPOMER	^{2}H number	ZPE (kJ/mol)	Pop. weight	$\|V_{xx}\|$	$\|V_{yy}\|$	$\|V_{zz}\|$
NH_3–Alanine	–	290.907	1	–	–	–
ND_3–Alanine	8, 9, 10	262.576	1	0.08358	0.04376	0.03982
NHD_2–Alanine	8, 9	272.134	0.3247			
NHD_2–Alanine	8, 10	272.078	0.3321	0.08361	0.04484	0.03877
NHD_2–Alanine	9, 10	271.996	0.3431			
NDH_2–Alanine	8	281.591	0.3231			
NDH_2–Alanine	9	281.505	0.3345	0.08367	0.04641	0.03725
NDH_2–Alanine	10	281.446	0.3425			
NH_3–Glycine	–	251.798	1	–	–	–
ND_3–Glycine	10, 11, 12	184.881	1	0.08214	0.08108	0.00106
NHD_2–Glycine	10, 11	194.278	0.2737			
NHD_2–Glycine	10, 12	193.869	0.3229	0.09303	0.08776	0.005274
NHD_2–Glycine	11, 12	193.317	0.4034			
NDH_2–Glycine	10	203.201	0.2702			
NDH_2–Glycine	11	202.657	0.3365	0.10520	0.09321	0.01198
NDH_2–Glycine	12	202.271	0.3933			
NH_3–Histidine	–	527.914	1	–	–	–
ND_3–Histidine	19, 20, 21	499.855	1	0.08794	0.05115	0.03678
NHD_2–Histidine	19, 20	509.068	0.3560			
NHD_2–Histidine	19, 21	509.448	0.3089	0.08806	0.05416	0.03390
NHD_2–Histidine	20, 21	509.275	0.3311			
NDH_2–Histidine	19	518.603	0.3346			
NDH_2–Histidine	20	518.433	0.3583	0.08834	0.05775	0.03059
NDH_2–Histidine	21	518.816	0.3071			

Table 4.1. Computationally determined zero-point energies (ZPE), Boltzmann population weights, and averaged and weighted electric field gradients for each isotopomer of alanine, glycine hydrochloride and histidine monohydrochloride monohydrate.

ISOTOPOMER	THEORETICAL			EXPERIMENTAL	
L-alanine	**NDH_2**	**ND_2H**	**ND_3**	**10%**	**~99%**
C_Q (kHz)	56.22	56.19	56.17	49.0	49.0
η	0.11	0.07	0.05	0.23	0.18
Glycine·HCl	**NDH_2**	**ND_2H**	**ND_3**	**10%**	**~99%**
C_Q (kHz)	57.31	56.43	55.63	53.0	53.0
η	0.62	0.52	0.41	0.22	0.17
Histidine·HCl·H_2O	**NDH_2**	**ND_2H**	**ND_3**	**10%**	**~99%**
C_Q (kHz)	59.37	59.17	59.09	51.0	51.0
η	0.31	0.23	0.16	0.12	0.10

Table 4.2. Comparison of theoretically and experimentally determined NMR parameters for variously deuterated amino acids.

References:

1. Krantz, B.A.; Srivastava, A.K.; Nauli, S.; Baker, D.; Sauer, R.T.; Sosnick, T.R. *Nat. Struc. Bio.* **2002**, *9*, 48-463.
2. Urey, H.C. *Liversidge Lecture at the Chemical Society in the Royal Institution*. **1946**, 562-581.
3. Ikabata, Y.; Imamura, Y.; Nakai, H. *J. Phys. Chem. A*. 2011, *115*, 1433-1439.
4. Northrop, D.B. *Ann. Rev. Biochem.* **2002**, *9*(**6**), 458-463.
5. Su, P.; Li, H. *J. Chem.Phys.* **2009**, *131*, 14102.
6. Morokuma, K. *J. Chem. Phys.* **1971**, *55*, 1236-1244.
7. Bagus, P.S.; Hermann, K.; Bauschlicher, J.C.W. *J. Chem. Phys.* **1984**, *80*, 4378-4386.
8. Glendening, E.D.; Streitwieser, A.J. *J. Chem. Phys*. **1994**, *100*, 2900-2909.
9. Sugimoto, H. *J. Phys. Cond. Matter*. **1998**, *10*, 1237-1246.
10. Shenderovich, I. G.; Limbach, H. H.; Smirnov, S. N.; Tolstoy,P. M.; Denisov, G. S.; Golubev, N. S. *Phys. Chem. Chem. Phys*. **2002**, *4*, 5488–5497.
11. Tolstoy, P. M.; Schah-Mohammedi, P.; Smirnov, S. N.; Golubev, N. S.; Denisov, G. S.; Limbach, H. H. *J. Am. Chem. Soc*. **2004**, *126*, 5621–5634.
12. Pietrzak, M.; Shibl, M. F.; Bröring, M.; Kühn, O.; Limbach, H. H. *J. Am. Chem. Soc.* **2007**, *129*, 296–304.
13. Alam, T.M.; Drobny, G.P. *Chem. Rev.* **1991**, *91*, 1545-1590.
14. Wu, X. *A combined study of experimental measurement and theoretical calculation of NMR property,* Ph.D. thesis, Univeristy of Nebraksa-Lincoln, **2011**.
15. Harbison, G.S.; Kye, Y.; Penner, G.H.; Grandin, M.; Monette, M. *J. Phys. Chem. B.* **2002**, *106*, 10285-10291.
16. Wilson, C.C.; Myles, D.; Ghosh, M.; Johnson, L.N.; and Wang, W. *New J. Chem.* **2005**, *29*, 1318-1322.
17. Lehmann, M.S.; Koetzle, T.F.; Hamilton, W.C. *J. Am. Chem. Soc.* **1972**, *94*(***8***), 2657-2660.
18. Al-Karaghouli, A.R.; Cole, F.E.; Lehmann, M.S.; Miskell, C.F.; Verbist, J.J.; Koetzle, T.F. *J. Chem. Phys.* **1975**, *63*(**4**), 1360-1365.
19. Fuess, H.; Hohlwein, D.; Mason, S.A. *Acta Cryst.* **1977**, *B33*, 654-659.

20. Soda, G. and Chiba, T. *J. Chem. Phys.* **1969**, *50*(**1**), 439-455.

21. Hinchliffe, A. *Computational Quantum Chemistry*. John Wiley & Sons, Ltd., **1988**.

22. Wann, M.; Harbison, G.S. *J. Chem. Phys*. **1994**, *101*, 231.

23. Bulychev, V.P.; Buturlimova, M.V.; Tokhadze, K.G. *Phys. Chem. Chem. Phys*. **2011**, *13*, 14019-14025.

24. Ando, K. *Phys. Rev B.* **2005**, *72*, 172104.

25. Grimison, A. *J. Phys. Chem*. **1963**, *67*(**5**), 692-964.

26. Li, X-Z.; Walker, B.; Michaelides, A. *PNAS*, **2011**, *3.*

27. Eckman, R.R.; Vega, A.J. *J. Phys. Chem*. **1986**, *90*(**19**), 4679-4683.

28. Williams, J.C.; McDermott, A.E. *Biochem*. **1995**, *34*(**26**), 8309-8319.

29. Lin, W-Y.; Blum, F.D. *J. Am. Chem. Soc.* **2001**, *123*(**9**), 2032-2037.

30. Reif, B. *Methods Mol. Biol.* **2012**, *831*, 279-301.

31. Weidner, T.; Breen, N.F.; Li, K.; Drobny, G.P.; Castner, D.G. *PNAS*, *107*(**30**), 13288.

32. Duer, M. J. *Introduction to Solid-State NMR Spectrscopy*. Blackwell Publishing, 2004.

33. Levitt, M.H. *Spin Dynamics: Basics of Nuclear Magnetic Resonance*. Wiley Scientific, 2001.

Chapter 5

Conclusions and Summary

To summarize this work, solid-state NMR can be combined with computational methods and fundamental physical chemistry concepts to study a variety of systems and problems. The nature of the quadrupole moment of deuterium makes it an especially convenient NMR probe for structural and dynamic studies of various compounds. The first project included in this work uses static SSNMR and lineshape simulations to determine if local furanose ring dynamics plays an important role in the BER recognition of thymidine:uracil mismatches in DNA. Careful consideration of water solvation can mimic biological conditions while still employing the benefits of solid-state NMR by restricting gross molecular motions like isotropic tumbling. Relaxation studies of specifically hydrated oligonucleotides were performed to determine if any gross deviation of furanose ring dynamics in the mismatched uracil-containing DNA compared to the thymidine-containing DNA could contribute to enhanced repair protein recognition. While this study showed that local ring puckering dynamics of the 2″ position on the furanose ring do not contribute to the recognition of the thymidine:uracil mismatch, previous studies have indicated that other local dynamics likely contribute to the repair protein recognition for a variety of other DNA lesions.

The second project presented was a study in materials chemistry. Self-assembling organic molecules have long been of interest for their applications in material science. Of

specific interest regarding the cyanuric acid-melamine complex was the illicit doping of human and animals foodstuffs with melamine as a way to increase the measured nitrogen content, the main concern being that once in the presence of cyanuric acid, a highly insoluble adduct forms (CAM) that may block kidney function leading to kidney failure and death. Review of previously published crystal structures and solid-state NMR of the CAM complex prepared via hydrothermal synthesis were significantly different compared to our initial NMR results and computationally optimized structures. Specifically, our results showed that the CAM complex prepared in neutral conditions at room temperature has a highly symmetric hexagonal lattice, compared to the reported monoclinic structure formed via hydrothermal synthesis. The relative strength of the N-H. . . N vs N-H. . . O hydrogen bonds were confirmed by comparing experimental C_Q values and chemical shift values. The trend in the magnitude of these NMR parameters coincides with previously published SSNMR hydrogen bond data. Additionally, we were able to show that the complex precipitates out of aqueous solutions faster than H/D isotopic exchange was able to occur. By selectively deuterating one species in ^2H_2O and mixing it with the other species prepared in ^1H_2O, we were able to obtain ^1H and ^2H MAS SSNMR that showed predominantly one species only.

The final project sought to understand the effects of H/D isotopic labeling on NMR parameters. Isotopic labeling is often considered non-perturbative. However, previously published data suggests that the smaller size of a deuteron in comparison to a proton causes the deuteron to selectively partition itself into weaker hydrogen bonds. Known as the thermodynamic isotope effect, this phenomenon can cause distortions in NMR parameters due to geometric and dynamic deviations upon isotopic substitution.

We successfully demonstrated this effect in three hydrogen-bonded amino acids: L-alanine, glycine hydrochloride and histidine hydrochloride monohydrate. By selectively deuterating each amino acid with 10% and ~99% deuteration, we were able to obtain significantly smaller η values for the perdeuterated species versus the 10% deuterated species via static SSNMR lineshape fitting. We were able to demonstrate this same trend theoretically as well. Boltzmann probability values were determined for each isotopomer from computational zero point energies of optimized structures. These probabilities were then used to average the electric field gradients obtained via DFT methods and then used to determine theoretical C_Q and η values. Since the values obtained via NMR represent the thermal average of the NMR parameters over all accessible states, further computational studies including averaging over the vibrational states would need to be performed to obtain an exact or approximate match in the NMR parameters.

Overall, this project demonstrates the many facets of solid-state NMR. Coupled with lineshape simulations and computational studies, SSNMR can be used to study problems that are more fundamental in nature (thermodynamic isotope effects) to problems that are significantly more complex (dynamics in biomolecules).