Fall 2016

Modeling The 3-Dimensional Structure Of D(Cgcgaattcgcg) And Its 8-Oxo-Da5 Adduct With 1H Nmr Noesy Refinements

Christopher Miles Reynolds

MODELING THE 3-DIMENSIONAL STRUCTURE OF d(CGCGAATTCGCG) AND ITS 8-OXO-dA5 ADDUCT WITH ^{1}H NMR NOESY REFINEMENTS

A Masters Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Chemistry

By

Christopher Miles Reynolds

December 2016

MODELING THE 3-DIMENSIONAL STRUCTURE OF d(CGCGAATTCGCG)

AND ITS 8-OXO-dA5 ADDUCT USING ^{1}H NMR NOESY REFINEMENTS

Chemistry

Missouri State University, December 2016

Master of Science

Christopher Miles Reynolds

ABSTRACT

Since the characterization of the oligomer **d(CGCGAATTCGCG)** has been published by Dickerson et al., computational studies have been carried out to produce an accurate 3D model. These models are important for visualizing how certain DNA repair enzymes, such as the glycosylases, recognize sites of damage by signatures of local 3D distortion. Using ^{1}H NOESY-generated internuclear distances to replicate the model of this oligomer and a derivative with an 8-oxo-dA5 lesion, we propose characteristics of helical distortion that DNA glycosylases might use for identifying this form of damage. In addition, this method of comparison can be used to study the repair signatures of other known DNA lesions.

KEYWORDS: oxo-A, NMR, NOESY, DNA, base-excision repair, Gaussian, density functional theory, 3D structure, Dickerson dodecamer.

This abstract is approved as to form and content

_____

Dr. Gary Meints
Chairperson, Advisory Committee
Missouri State University

MODELING THE 3-DIMENSIONAL STRUCTURE OF d(CGCGAATTCGCG)

AND ITS 8-OXO-dA5 ADDUCT USING ^{1}H NMR NOESY REFINEMENTS

By

Christopher Miles Reynolds

A Masters Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Masters of Science, Chemistry

December 2016

Approved:

_____
Dr. Gary Meints

_____
Dr. Matthew R. Siebert

_____
Dr. Dean Cuebas

_____
Dr. Kyoungtae Kim

_____
Dr. Julie Masterson: Dean, Graduate College

ACKNOWLEDGEMENTS

I owe my greatest thanks to Dr. Gary Meints. When he took me into his group, my science progression was slipping through the cracks. At the time, my former research advisor dismantled his projects and moved out of the country. I had to start over and Dr. Meints accepted me even though I had no expertise in NMR. His kind, thorough, and patient coaching helped me develop confidence and independence with this instrument. Thanks to his honesty, I also feel I've become a more agreeable and less biased person.

In addition, I would like to thank Dr. Alan Schick for all the mentoring and support he has given me these past few years. I really enjoyed and feel very grateful to him for introducing me to BASF and what it's like to do research and attend meetings with industrial chemists. Dr. Schick also shaped me into a better human and I will always look up to him for his wit and clarity.

Although I did not pursue an organometallic route in research, Dr. Nikolay Gerasimchuk has been a tremendous influence in my life for kindling an autodidactic spirit in me. By introducing many wonders of the chemical world, he has made me relentlessly excited to learn more.

Drs. Shujun Su, Matthew Siebert, and Dean Cuebas taught me how to navigate the theory and software behind the computational chemistry used in this thesis. Without them, I would be lost. Dr. Adam Wanekaya helped me realize how much I value metrology and it was always a pleasure to discuss matters with him.

I am very grateful to Geoffrey Manani, Michael Nothnagel, Jonathan Hardin, Chad Hagan, Aaron Proctor, Chunling Cao, and Stephen Kramer for their shared interests and accessibility to questions throughout my graduate studies. Also, my family cannot be thanked enough for keeping me awake and on my toes this whole time. Lastly, I would like to thank Lisa Kirchner for being my candlelight. Ever since we met, she has always been there for me and I will always be there for her.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. BIOLOGICAL BACKGROUND

1.1 Introduction

Deoxyribonucleic acid (DNA) is a modular and repetitive molecule with multiple levels of structure. It was identified by Friedrich Miescher in 1868 and demonstrated as a genetic carrier in 1944 by Avery, MacLeod, and McCarty.[1] In the famous experiment by James Watson and Francis Crick, DNA was posited as a double helix, held together by hydrogen bonds.[2] As the genetic blueprint and legacy of all life, DNA is necessarily protected and repaired when modified. While organisms capitalize on the diversity offered by mutation, such as in the adaptive immune system and by evolution, many regions are necessarily conserved and result in harm when altered. Accordingly, repair processes are vital.

The regulation and mechanism of DNA repair has been studied intensively, but the detailed process in which enzymes recognize damage is uncertain. Traditional models of enzyme-substrate recognition by molecular groups cannot always apply for DNA, since it is a redundant molecule and docking seems to be more dependent on local helical distortions and dynamics. In this thesis, I hope to clarify part of the repair signature for DNA damage by focusing on the 8-oxo-dA lesion, which is where the hydrogen atom at the 8^{th} carbon in adenosine is replaced by oxygen. The lesion, inserted at position A5 (underlined) into CGCG<u>A</u>ATTCGCG, known as the Dickerson dodecamer,[3] will be studied by way of ^{1}H NOESY to glean insights about its 3D structure. This project represents the initial phase in modeling the 8-oxo-dA damage in DNA by setting into motion techniques by which internuclear distances are discovered. In addition, methods for incorporating these distances into computer-simulated models will be discussed.

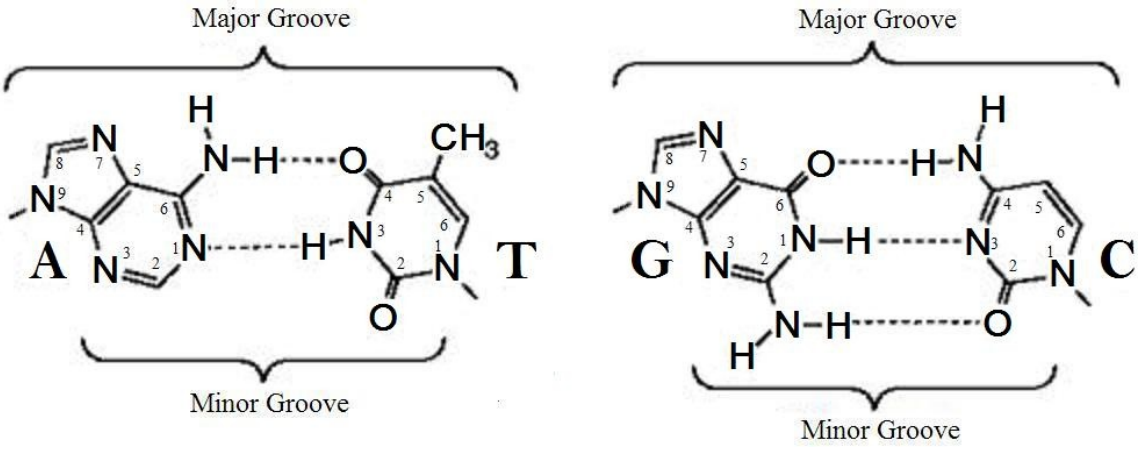


Figure 1.2.1. General structure of the 4 nitrogenous bases, showing the numbering scheme, base-pairing, and the origin of the two grooves in DNA.Dashed lines indicate Watson-Crick base pairs. The nitrogens N1 of the pyrimidines and N9 of the purines connect to the deoxyribose moiety of DNA (not shown). From left to right: A = Adenine, T = Thymine, G = Guanine, C = Cytosine.

1.2 DNA Structure

The primary structure of DNA is its sequence of units, the nucleoside monophosphates, also known as nucleotides. These are composed of a 5-membered sugar (deoxyribose) linked to a phosphate and one of 4 nitrogenous bases: Adenine (A), Guanine (G), Cytosine (C), or Thymine (T). Nucleosides are connected by phosphates covalently in a polyester chain, called the phosphodiester backbone at the 5′ and 3′ ends of the sugar, so-named by the number given to the carbons at those positions. At the 1′ (anomeric) position, the base is connected by a β-glycosidic linkage, which indicates that the stereochemistry (R or S) of the two stereocenters is not the same. The nitrogenous bases are further classified as pyrimidines or purines, which respectively have one or two heterocyclic rings. In healthy DNA, the purines (A and G) form hydrogen bonds with the pyrimidines (T and C) on the opposing strand specifically as A=T and G≡C, although

there are exceptions, such as where G-G hairpins and quadruplexes occur naturally.[4] These forces collectively create the biopolymer DNA with the paired strands running antiparallel to each other.

The secondary structure of DNA reveals how and in what arrangement bases pair, illustrated in tandem with tertiary structure in F1.2.2. In secondary structure, names are ascribed to the few configurations in which hydrogen bonds connect bases, modeled after Watson and Crick,[2] Hoogsteen, [5] and others.[6] In addition, bases can pair with non-neighbors on the opposing strand or members on the same strand, although this is generally reserved for RNA and viral DNA. Such forms can produce bulges, hairpin loops, cloverleaves, and other local features which, together with proteins or other molecules, can fold into structures that have catalytic functionality.[7,8,9] Lastly, the nearly planar arrangement of the bases in double-stranded DNA (dsDNA) is packed close enough that electrostatic attractions occur, known as π-stacking. In DNA, stacking is generally a result of π-orbital overlap and substituent effects, but the actual contribution of each is an area of disagreement and active research.[10,11]

Single-stranded DNA (ssDNA) automatically pairs with a neighboring strand under physiological conditions and disruption of this is termed denaturation or melting.[12,13] The rejoining of strands after denaturation, or annealing, is a useful technique for ensuring that all of a DNA sample is double stranded for analysis.[14] This process involves heating the DNA to its melting point, followed by cooling to room temperature.

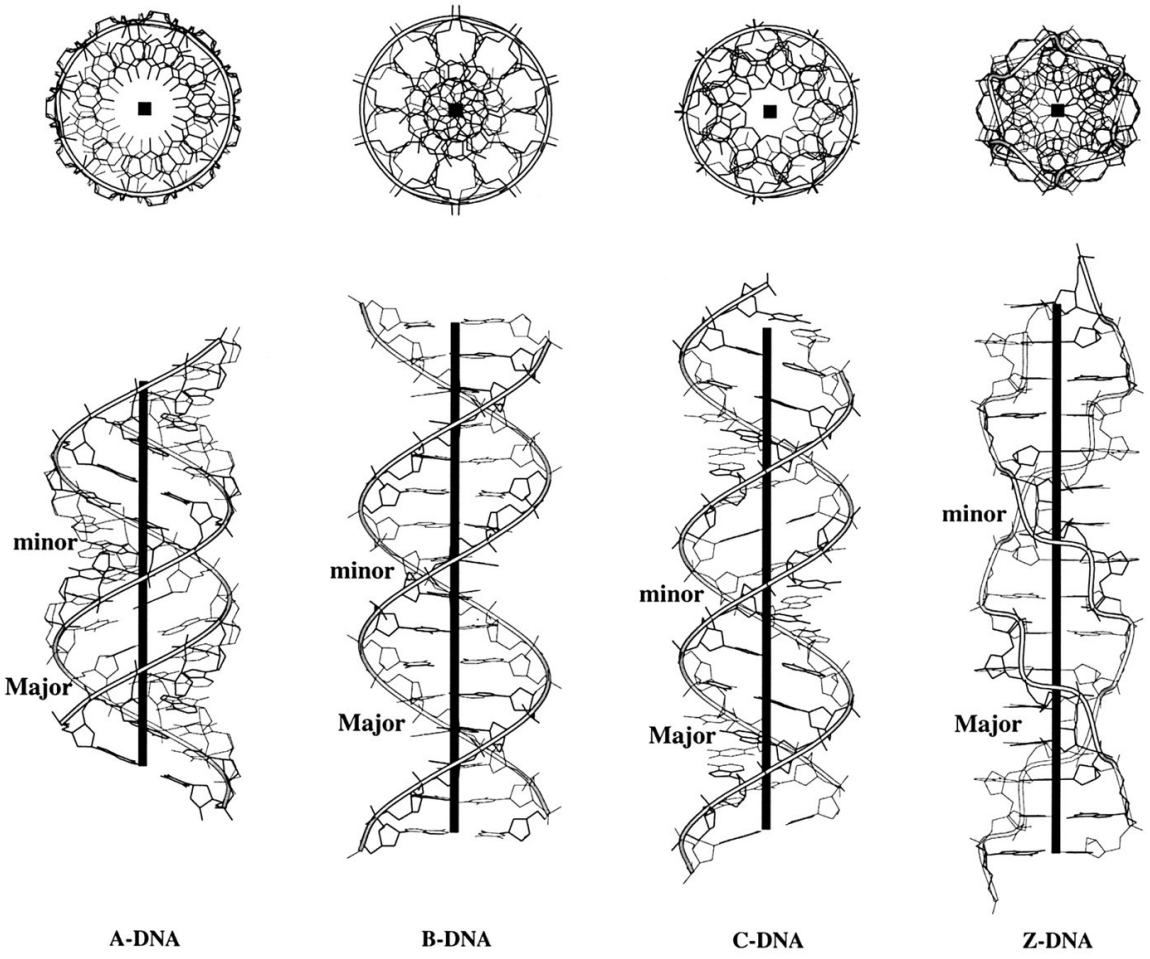


Figure 1.2.2. Differences in helicity seen by four of the many possible tertiary structures of DNA. B-DNA is the most common form in organisms. In addition, secondary structure is visualized by focusing on the differences in base pairing. Replicated with permission from Xiang-Jun Lu & Wilma K. Olson.[51]

The twisting of dsDNA is its tertiary structure, and usually takes the form of a right-handed double helix, but also has left handed and triple helical forms. Those commonly encountered in the biological setting are named by letter and can be seen in F1.2.1. There are two gaps between rounds of the helix named the major and minor grooves. The differences in phosphate torsion angles and sugar puckering between these forms of DNA are the primary factors that account for the distances in these grooves,

which in turn affect how much base stacking is involved. F1.2.2 provides another way to differentiate the major groove and minor grooves.

The compaction of the helix is mathematically treated with a linking number, Lk_o, which is a combination of twist (tw) and writhe (wr). Twist is the number of times one strand of DNA is wrapped around the other in the right-handed (clockwise) direction, while writhe is the number of times the axis of a closed loop of DNA is coiled. For reference, B-DNA has 10.4-10.5 base pairs per turn.[15] The longer it is, the greater the number of turns and thus the higher the linking number. Since B-DNA is a right-handed helix, its linking number is positive and any deviation from this is considered supercoiling.

Positive supercoiling condenses DNA, while negative supercoiling unwinds it. Since local unwinding is necessary for transcription and translation; natural dsDNA exhibits negative supercoiling in those regions. The stress of supercoiling creates emergent structures, such as solenoids and figure-eights which primarily depend on the number of coils, pH, and structural proteins to support the topology. Supercoiling is initiated in cells by topoisomerases, which cleave the phosphodiester backbone of one strand of dsDNA (Type I topoisomerase) or both (Type II). Type I topoisomerases change the linking number of dsDNA by causing one strand to rotate about the other, then repair the nick to preserve the changes. Type II topoisomerases cleave both strands of DNA and insert a double-stranded fragment.[16] In this way, the degree of helical constriction is modulated to grant site-specific regulation. Owing to its relationship with damage recognition by repair enzymes, the helicity of the DNA studied in this thesis will be semi-empirically quantified and explored in the Discussion section.

Nuclear DNA is wrapped around histones and other structural proteins. This comprises a quaternary structure, the chromosomes, which pack, protect, and govern accessibility further. DNA is rarely naked outside the loci of transcription and replication. Although an important feature in the landscape, quaternary structure will not be discussed in detail because it is largely a result of DNA-protein interactions.[17] Even with these levels of protection, damage does occur.

This collection of research will refer to numbering scheme in F1.2.3 extensively. Each number on the bases represents a carbon in the cycle. If a hydrogen atom is attached to it, it is named after that number.

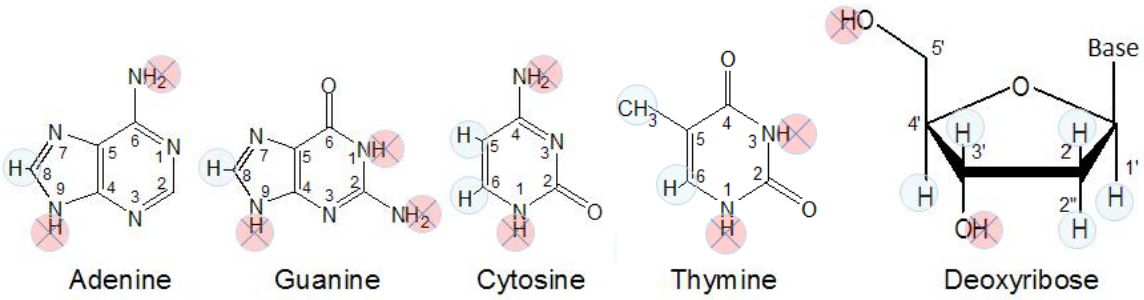


Figure 1.2.3. Numbering scheme of the four major DNA bases and sugar. Blue circles indicate non-labile ^{1}H nuclei, while red circles indicate those nuclei which are eitherabsent or exchangeable in the solvent. Only Hydrogens encircled in blue provide NMR signals in the experimental setup.

To distinguish between hydrogens with non-degenerate NMR signals pointing in different directions on the same carbon, the prime (′) and double prime (″) are also used. For the experiments outlined in this thesis, the 1′, 2′, 2″, and 3′ hydrogen nuclei on the sugar will be measured. For pyrimidines, the C5H and C6H will be recorded, with the exception of thymines, which have a methyl group instead of C5H. On the purines, only the C8H will be measured. By tradition and because of a similar chemical shift, the C8H on purines will be referred to as the Base proton, as will the C6H on the pyrimidines. In

addition, each nucleotide will be numbered by its position in the oligomer, from 1 to 12, as C1, G2, C3, etc.

1.3 DNA Damage

DNA inherently resists certain forms of modification compared to RNA. As a general rule, DNA incorporates the thymine base, whereas RNA uses uracil. The C5-methyl on thymine has been found to stabilize DNA by contributing to base-stacking interactions.[18] In addition, the presence of a 2′ hydroxyl group on the sugar in RNA leaves it open to hydrolysis by a suitable nucleophile.[18] The 2′ hydroxyl is absent in DNA and is the reason for the "deoxy" affix. In these two aspects of stability, DNA has been considered the more appropriate candidate for representing the genome.

DNA can be damaged in a variety of ways. High energy radiation, such as ultraviolet light can cause dimerization of adjacent pyrimidines and double strand breakage. DNA can be modified by reactive oxygen species (ROS), which are oxygen-containing radicals. Radicals possess one or more unpaired valence electrons, making them highly reactive because the pairing of electrons is very energetically favorable. In this thesis, the form of oxidative DNA damage studied is the 8-oxo-2-deoxyadenosine (8-oxo-dA), seen in F1.3.1. The lesion can form when ultraviolet radiation excites molecular oxygen to its singlet state, which reacts with DNA. Several other pathways to its formation occur in the body as well, including direct attack by hydroxyl radicals. The same type of damage occurs with deoxyguanosine (dG) and the two species have been used as markers of oxidative stress in animal models.[19]

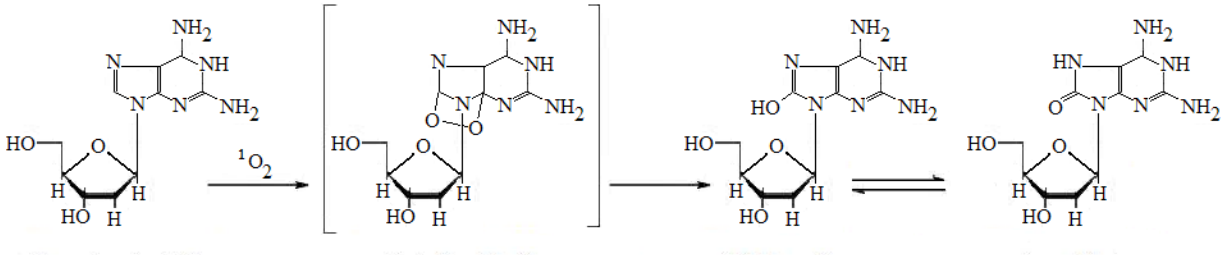


Figure 1.3.1. Reaction of deoxyadenosine with singlet oxygen to form the 8-oxo-2-deoxyadenosine adduct.

Commonly encountered ROS in cells have a lifetime on the order of nanoseconds, such as for OH•, HOO•, and O_2^{-}•, depending on what chemical species they encounter.[20] Singlet oxygen lasts relatively longer, at around 3-4μs in a cell[21] and 48μs[22] in aqueous solution. Nitric oxide NO• is an exception, lasting up to 2s[23] in the blood between erythrocyte transfer to tissues, where it acts as a molecular messenger. The ultimate lifetime of NO• is much longer, as it can participate in many cellular processes.

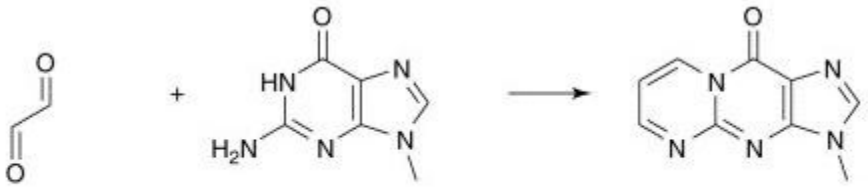


Figure 1.3.2. Reaction of malondialdehyde with guanine to form an alkylated purine derivative.

Evidence exists that ROS can cause damage indirectly by forming reactive byproducts. They can be introduced into the body by cooking and eating, whereby oxygen from the air reacts with heated foods and creates a cocktail of products, some of which are harmful. Malondialdehyde (F1.3.2) is an example that can be encountered by ROS-induced rearrangement of oils during cooking, which can initiate electrophilic reactions and form adducts with proteins and purines.[24]

Reducing sugars can react with amino groups, including those of nucleic acids. These are a class of sugars that either have an aldehyde group or can isomerize into having one. This proceeds by the Maillard reaction, simplified in F1.3.3, resulting in advanced glycation endproducts (AGEs). Although AGEs are not ROS, they are often grouped with them, since they are harmful and indirectly participate in the generation of ROS in the body. By one route, AGEs that crosslink to mitochondrial proteins can lead to leaking of ROS during oxidative phosphorylation.[25]

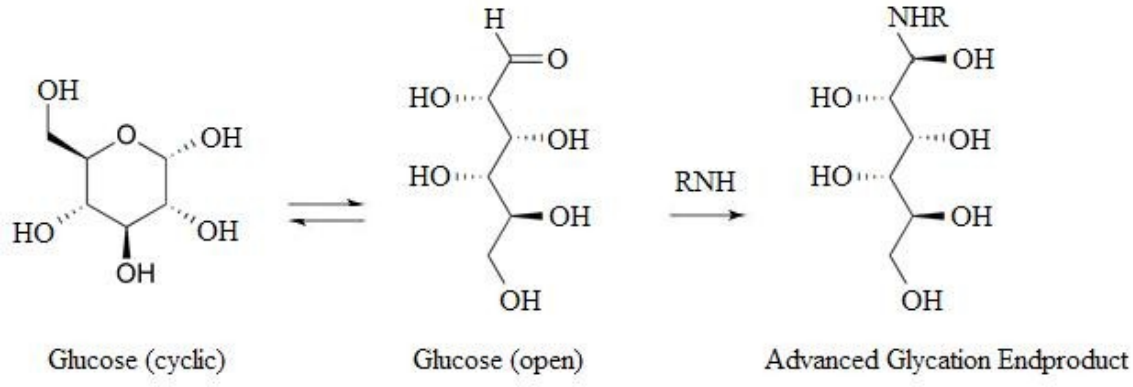


Figure 1.3.3. An example of the Maillard reaction, where reducing sugars can attach to free amino groups.

AGEs of proteins in the lens of the eye can turn them into photosensitizers, which enter excited states and participate in undesirable electron transfer reactions.[26] Damage to proteins involved in regulating DNA reduces the ability of cells to recover, making AGEs and ROS strong candidates for part of the reason why organisms age.

DNA modified by reactive molecules often have abnormal properties.[19] This is especially prominent when nucleotides have participants in base-pairing overridden by new functional groups. By interfering with replication, regulation, and causing repair mechanisms to take place which do not always replace the site of the lesion with the correct base, DNA adducts are mutagenic.

1.4 DNA Repair

Mutations and damage are of great consequence because DNA uniquely serves as a permanent copy of the cell and mitochondrial genomes. Since these alterations can impede cell function, several mechanisms for repairing damaged DNA have evolved. Repair pathways can be divided into two basic classes: 1) direct reversal of the chemical reaction that caused the damage and 2) removal of the damaged base followed by its replacement with newly synthesized DNA.[27]

Only a few types of DNA damage are repaired using direct reversal. Two examples are the repair of pyrimidine dimers that can result from exposure to UV light and alkylated guanine residues that have been modified by methyl or ethyl groups added to the O^6 position on the purine ring.[27] Although direct reversal is an efficient way of dealing with a specific type of DNA damage, excision repair is a more general mechanism for fixing a wide variety of chemical alterations to DNA.

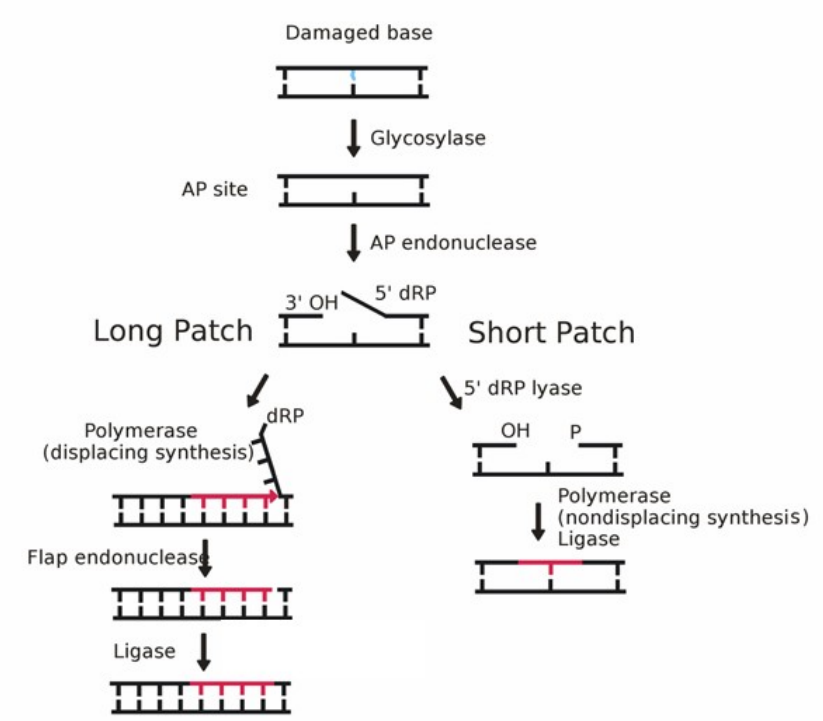


Figure 1.4.1. The general mechanism of base-excision repair (BER) as it proceeds through the long and short patch pathways.[27]

In excision repair, the damage is recognized and removed, either as free bases or nucleotides. The resulting gap is then filled in with newly synthesized DNA, using the complementary strand as a template. There are three different types of excision repair mechanisms: 1) base-excision repair, 2) nucleotide-excision repair and 3) mismatch repair.

Base-excision repair (BER) arises when damage to single bases is recognized and they are removed from the DNA molecule. This is the mechanism by which the 8-oxo-dA lesion studied in this thesis is repaired.[28]

The removal of a damaged base is catalyzed by DNA glycosylase. DNA glycosylase is an enzyme that cleaves the bond linking the base to the deoxyribose of the DNA backbone. This will yield a free base as well as an apyrimidinic or apurinic (AP)

site, which is a sugar without a base.[27] These sites are repaired by AP endonuclease, which cleaves the molecule adjacent to the AP site. The remaining deoxyribose moiety can then be removed, thus leaving a single nucleotide gap. DNA polymerase and ligase fill the gap via the long or short patch as shown in F1.4.1. The choice between long or short patch is thought to be influenced by the type of lesion, the cell cycle stage and whether the cell is terminally differentiated or actively dividing.[29] Some lesions, such as oxidized or reduced AP sites, cannot be cleaved by the lyase, so they must be processed by the long-patch pathway.[29]

The 8-oxo-dA lesion studied in this thesis is repaired by the BER pathway. It was shown in the 1990s that eukaryotic 8-oxo-dG-DNA glycosylase (OGG1) and mammalian endonuclease VII-like protein 1 (NEIL 1) could efficiently remove 8-oxo-dA, but only when it is paired with cytosine.[30] 8-oxo-dA•dC pairs can occur when the oxidized deoxyadenosine triphosphate (dATP) is misincorporated during DNA replication.[16] It has only recently been reported that human thymine-DNA glycosylase (hTDG) is able to excise 8-oxo-dA from its pair with thymine; the form that occurs when normal dA•dT pairs are oxidized directly.[30] The results of this thesis are most relevant to this repair pathway, since the artificially-induced lesion in the dodecamer is across from a dT, not a dC. Although this thesis will attempt to describe the type of helical distortion for enzyme recognition, the excision process itself occurs at the same rate regardless of the flanking sequence.[30]

In addition to the removal of the 8-oxo-A lesion, BER is the process by which uracil is removed from the DNA molecule. Uracil can appear by two primary mechanisms: 1) it can accidentally be incorporated in the place of thymine during replication and 2) by the deamination of cytosine.[27] Other abnormal bases that can be

removed by base-excision include hypoxanthine formed by the deamination of adenine, most types of alkylated purines, and bases damaged by oxidation or ionizing radiation.[13]

Whereas BER involves single base pairs, Nucleotide-excision repair (NER) removes damaged bases as part of an oligonucleotide containing the lesion. This pathway usually repairs bulky, helix-distorting lesions, which often result from exposure to UV light or reactions with various carinogens. There are nine major proteins involved in nucleotide-excision repair in mammalian cells, largely identified from studies on individuals with xeroderma pigmentosum (XP), a rare genetic disorder that causes the affected individual to be extremely sensitive to UV light and to develop multiple skin cancers on regions of their body upon exposure.[31] Nucleotide-excision repair can be divided into two subpathways: global genomic and transcription coupled.[32] The two differ in how they recognize DNA damage, but share the same process for lesion incision, repair and ligation, as can be seen in F1.4.2.

The basic model for mammalian cells begins with the XPA protein recognizing the damaged DNA and forming complexes with the other proteins involved in the repair process.[27] These proteins include XPB and XPD, which act as helicases that unwind the damaged DNA. The binding of XPA to the lesion also recruits XPF and XPG to the repair complex,[27] which are endonucleases that cleave the DNA on the 5′ and 3′ sides of the damaged site, removing an oligonucleotide consisting of approximately thirty bases. The resulting gap is filled by DNA polymerase δ or ε and then sealed by ligase.

Mismatch repair (MR) recognizes base-pairing violations that result from insertions, deletions and misincorporations, which can arise during DNA replication and recombination. In humans, mismatched pairs are principally removed by the proofreading activity of DNA polymerase δ and ε during replication and by a complex of recruited

proteins (MutL, MutS, MutH, and PCNA) during any time of the cell cycle. RNA is proofread by aminoacyl-TRNA synthetases.[33]

When DNA polymerase recognizes an incorrect base pair, it will move backwards by one nucleotide. The 3′-5′ exonuclease activity of the enzyme allows the incorrect pair to be removed. Following excision, the polymerase will reinsert the correct base and replication will continue forwards. The ones that are missed are subject to later correction by the Mut protein system. Several Mut homologs are present in humans, classified as MutS, MutL and MutHtypes depending on their mechanism of action.[27]

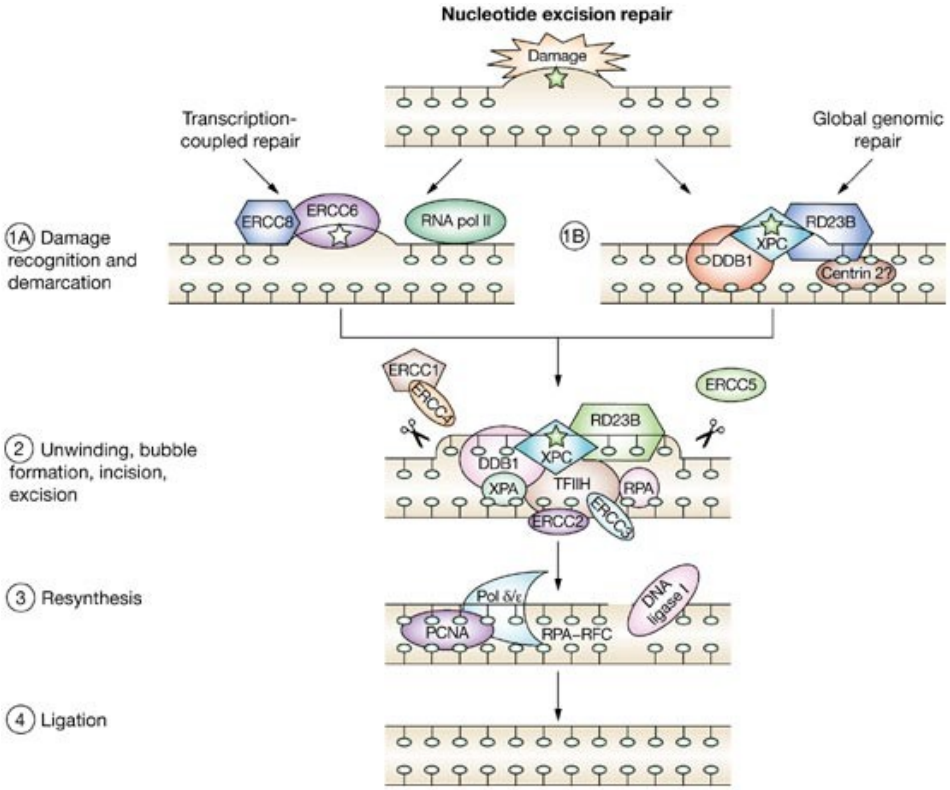


Figure 1.4.2. The two subpathways of nucleotide-excision repair, illustrating the different complexes of proteins recruited to the lesion site during transcriptionally active (1A) and inactive DNA (1B).[32]

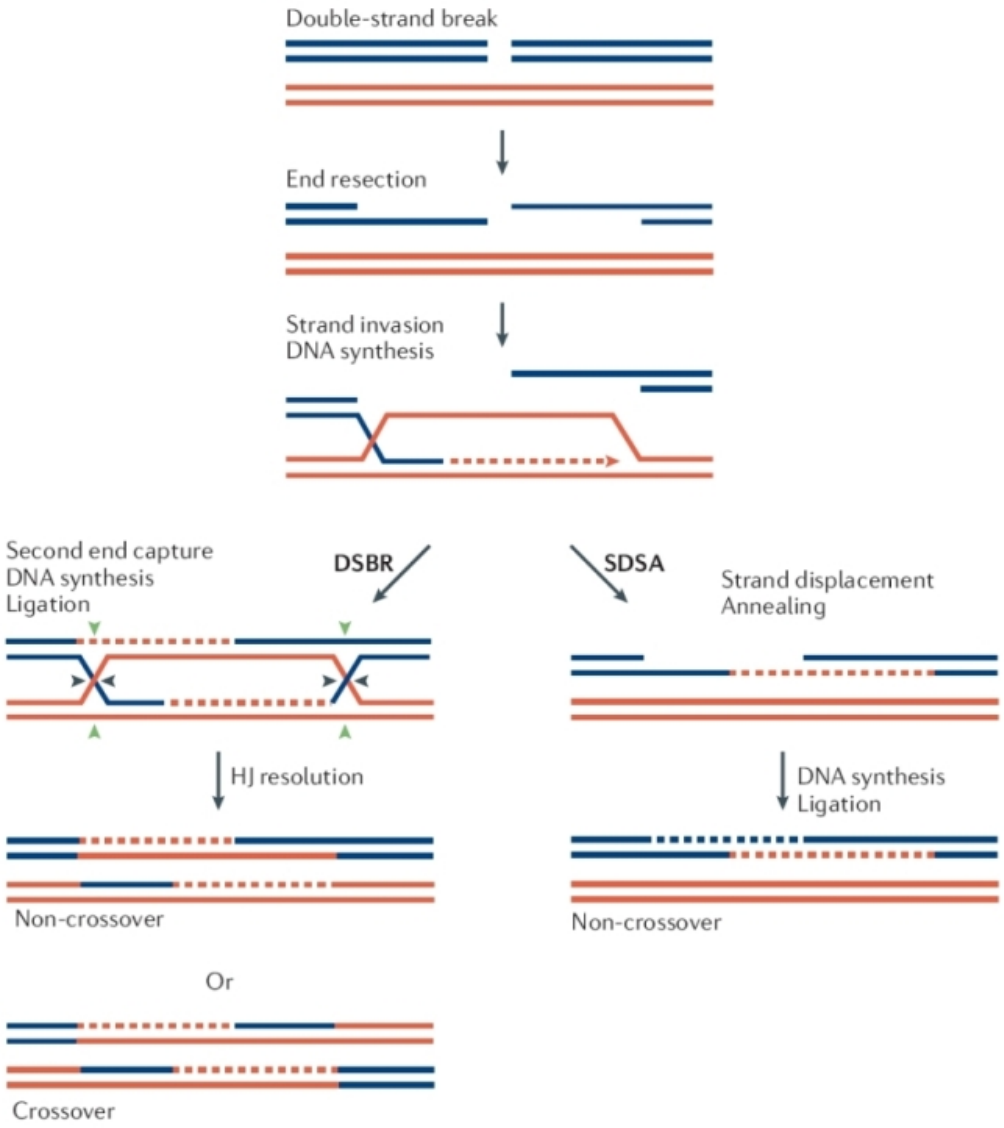


Figure 1.4.3. The double-strand break repair (DSBR) and synthesis-dependent strand annealing (SDSA) models of homologous strand recombination.[35]

Cells can initiate nucleotide-excision repair (NER) when bulky DNA adducts such as thymine dimers are generated by ultraviolet light, seen in F1.4.2. Like mismatch repair, this proceeds by two pathways. In transcription-coupled NER, the bulky lesion causes RNA polymerase to become clogged, resulting in the recruitment of a protein complex to liberate it. In the global genomic NER pathway, DNA that is not undergoing

transcription is recognized and bound by a separate complex of proteins. These two processes differ only by the stages of damage recognition; the lesion is excised and replaced by the same repair enzymes.Another type of repair process is homologous strand recombination (HR). Here, two similar or identical molecules of DNA align and one serves as the template of exchange for the other.[34] HR is most widely used by cells to accurately repair harmful breaks that occur on both strands of DNA, known as double-strand breaks. The processes can vary between organisms and cell types, but generally follow the same basic steps as F1.4.3. After a double-strand break occurs, the sections of DNA around the 5′ ends of the break are cut away by a process called resection. Next, an overhanging 3′ end of the broken section invades and links to the non-damaged DNA template, forming an X: the Holliday junction. After strand invasion, the sequence of events can follow two main pathways: 1) double-strand break repair (DSBR) or 2) synthesis-dependent strand annealing (SDSA).[35]

The DSBR pathway is unique in that the second 3′ overhang also forms a Holliday junction with the homologous chromosome. These double Holliday junctions are converted into recombination products by nicking endonucleases, which cut one DNA strand.[35] The DSBR pathway commonly results in crossover, where large segments of DNA are swapped between two strands. In the SDSA pathway, the 3′ invading strand is extended along the recipient DNA duplex by a DNA polymerase. It is released as the Holliday junction between the donor and recipient DNA molecules slides in a process called branch migration.[35] The newly synthesized 3′ end can then anneal to the other 3′ overhang in the damaged chromosome through complementary base pairing. Homologous strand recombination that occurs during DNA repair tends to result in non-

crossover products, meaning the damaged molecule is restored to its original state before the double-strand breakage.

1.5 The Dickerson Dodecamer

The Dickerson dodecamer (DDD) has been studied extensively by 1D and 2D NMR, X-ray crystallography, and modeling software. Among the many reasons for its utility, the Dickerson dodecamer was among the earliest DNA oligomers to be modeled. It adopts the B-DNA conformation, is stable, and can be cleaved by EcoRI, a restriction endonuclease whose structure has been determined.[36]

In the first paper by Richard E. Dickerson in 1981,[3] advances in DNA synthesis made it possible to support the claims of Watson and Crick with crystals large enough for single-crystal X-ray analysis, whereas the prior work in 28 years had been supported mainly by X-ray diffraction. The study generated the first working model of the B-DNA dodecamer with main chain and glycosyl conformation angles and set the stage for further studies, carrying the eponym. The first NMR study on the Dickerson dodecamer was performed in 1983 by B. R. Reid and coworkers, where the DNA sample was prepared in D_2O and most of the non-exchangeable ^1H nuclei were assigned by chemical shift via COSY and NOESY.[37] Distance measurements were attempted by the same group on a modified (hairpin) version of the dodecamer, since it could switch between hairpin and straight chain forms in solution.[38,39]

The first NMR-derived distance measurements for the Dickerson dodecamer appeared in 2000 by Ad Bax and colleagues, stabilized with phospholipid bicelles that conferred a liquid crystal phase.[40] This model, along with the original X-ray crystallographic model by Dickerson, forms the basis for structural comparison of the

Dickerson dodecamer in this laboratory, which was obtained without any stabilizing agent. This was done for two reasons. Firstly, a stabilizing agent might mask aspects of helical distortion in the 8-oxo-dA modified dodecamer. Secondly, the conformational switching between hairpin and straight chain DNA after annealing was not observed at the temperature (25 °C) and salt concentrations (25 mM PO_4^{3-}). The predominant species observed in this thesis was B-DNA with peaks suggestive of hairpins appearing inconsistently and at very low intensity.

CHAPTER 2. COMPUTATIONAL CHEMISTRY

2.1 Density Functional Theory

Density functional theory (DFT) is an exact mathematical treatment for describing the ground state properties of an electronic system. Given the number and type of atoms in a molecule, their energy-minimized bond distances, angles, and connectivity can be approximated with increasing accuracy (but to a limit) based on the number of atomic orbitals represented by basis functions. Instead of solving the time-independent Schrödinger equation forthe ground state properties, which poses a many-body problem due to electron-electron interactions, DFT treats the molecule with a single set of electron densities using the Kohn-Sham equations[41] and approximates the interaction energy between electrons by the Local Density Approximation (LDA).[42] This reductionist approach, representing many complex interactions with fewer variables defines it as a functional theory, since a functional is a lower-dimensional function of a function.

One major challenge to solving the time-independent Schrodinger equation is the many-body problem. Since the motion of the electron is strongly coupled to that of its neighbors, computing time exponentially scales with the number of interacting particles added into the system. The Kohn-Sham equations are inspired by the variational principle, which states that although a natural system may be impossibly complex, it will tend to adopt the lowest energy configuration over time. DFT works by exactly solving each electron as non-interacting entities at their ground states and applies a correction factor E_{XC} for the approximate interactions between them, named for the exchange and correlation energies. The full Kohn-Sham treatment, LDA, and Hamiltonian operation for energy is overviewed in App A.

The LDA calculates the exchange and correlation energies E_{XC} using Quantum Monte Carlo methods (beyond the scope of this text) which are most accurate for systems where there is a low level of charge density fluctuation.[42] Since the LDA poorly predicts van der Waals dispersions and gives a very poor description of hydrogen bonding, hybrid functional are an area of active research, which are based on a combination of computed (via the Hartree-Fock formalism) and empirical observations. B3P86 is the set of hybrid functionals considered by this thesis because it has been shown to produce hydrogen bond distances and energies for individual DNA base-pairs comparable to crystal structures when the solvent is simulated with sodium ions.[43] That stated, simulation of oligomeric DNA poses a continuing challenge to research in this field due to discrepancies in distances caused by base stacking and other forces not present in single base pairs.

DFT can provide very good approximations for certain problems and works best for determining structural stability, vibrational properties, and to a lesser degree with hybrid functionals, certain electronic properties.[44] DFT is an exact theory by its construction, but like other methods of computer simulation, is not a theory where one can reach an exact result or use techniques such as longer computation or increase basis set size to theoretically come up with the exact result. Even if the user completely converges in DFT, there will be some error. Convergence is where the total energy of the calculated system differs from the previous iteration by less than a set amount. The smaller this amount, the more iterations are required (and therefore longer processing time) for the program to stop the fine-tuning.

2.2 Quantum Mechanics in Computer Simulation

As described previously, DFT uses the Kohn-Sham equations to compute the charge and spin density for N non-interacting electrons, while the exchange-correlation energy (E_{XC}) is accounted for by the LDA. Before the E_{XC} can be understood, it is important to introduce the concept of spin in a quantum mechanical system.

Spin has been notoriously difficult to picture in a classical (Newtonian) sense. It is a type of angular momentum attributed to a point particle. Since a point cannot "spin", the word is a substitute for the phenomenon. For non-point particles, it takes 360 degrees for the spin of that object to return to its initial state. This is also true for a point particle with a spin of 1. However, for an electron with spin of ½, it takes two "revolutions" to return to the initial state. This behavior is strange and has no classical counterpart, except possibly a Möbius strip, which is a circular ribbon with a half-twist. If a line is drawn lengthwise on one side of the Möbius strip, it appears on the back side after one revolution, then at the starting position after a second. Translating this to 1 temporal and 3 spatial dimensions (the classical point of view), a point particle with a spin of 1 would rotate in such a way as to cut out a spherical trajectory, while one with ½ spin would produce an inverted sphere.

To demonstrate exchange energy, a detailed description of the Pauli Exclusion Principle is required, which states that paired electrons sharing an orbital must have opposite spins.[41] Quantum mechanical particles have the unusual property of exchange symmetry. In a system of two identical particles where there are two possible states, it is impossible to distinguish which particle is in which state. The particles exist in superposition for the two states because there is an equal probability for either. This superposition of wavefunctions is mathematically treated as a sum or difference of the

two states, depending on whether the states are respectively symmetric or antisymmetric. Electrons have a spin of ½, which corresponds to an antisymmetric spin state. They are classified as fermions, as opposed to bosons with integer spin (symmetric spin states). If the electrons have the same spin in the same state (i.e. a shared orbital, since spin and spatial coordinates are coupled), their superimposed wavefunctions cancel out to zero, meaning that the probability of finding them in that state is zero. Thus, the Pauli Exclusion Principle leads electrons with opposite spins to be more likely to occupy nearby positions.

The correlation energy is a much simpler behavior of interacting electrons that decreases the coulombic potential. It is generated because the coupling of motion between nearby electrons causes them to have trajectories that encounter less repulsive force than if they were independent.[41] This feature drops off dramatically with distance as well. However, some correlations are not distance-dependent and unpredictable. Another reason why DFT can never produce an exact model of a QM system is that coupling in the form of quantum entanglement can occur between natural electrons. The only way to accurately model entanglement is by knowing beforehand which electrons are entangled, which parameters are entangled, and for what duration. This requires invoking "hidden variables" that cannot be known definitively by classical computing methods.[45, 46] Despite these handicaps, DFT experiments have shown a remarkable fidelity to empirical observations.

2.3 DNA Simulation

Owing to the complexity and number of atoms involved, full DFT treatment of oligomeric DNA exceeds the current limits of computation. On the other hand, techniques based on Newtonian laws of motion such as the UFF (Universal Force Field)[47] have inaccurately portrayed some degrees of freedom such as found in helicity, leading to the development of more specific methods like CHARMM (Chemistry at Harvard Macromolecular Mechanics).[48]

Semi-empirical methods like PM6-DH+, which mix quantum mechanics with experimental data have been developed to minimize resource intensity.[49] Small regions of DNA can be given DFT treatment however, using the ONIOM multilayered model featured in Gaussian 09.[50] In this thesis, the 8-oxo-dA5 lesion of the Dickerson dodecamer was modeled in Gaussian 09 with BPV86, while the most relevant semi-empirical method available for the rest of the molecule was PM6. Distance restraints conforming to standards put forth in sections 5.3 and 5.5 were applied to individual nuclei. While PM6 is not as accurate as other methods, especially in terms of hydrogen bonding, it must be emphasized that the point of the modeling procedure was explorative, not quantitative.

Since Gaussian 09 does not include features for specifying helical and base-pairing properties, another software tool that will be involved in the analysis is 3DNA, [51] which uses a matrix-based approach to detect them. The comparison of these parameters for the 8-oxo-A and unmodified DDD sequence are provided in sections 5.5-6.

CHAPTER 3. NMR SPECTROSCOPY

3.1 Introduction to NMR

Nuclear Magnetic Resonance (NMR) spectroscopy is a powerful tool for studying molecular structures because it focuses on the spin dynamics of atomic nuclei, which tend to persist in relative isolation to the solvent and other intermolecular effects. Samples can be analyzed by a range of techniques, from small compounds in one-dimensional proton (^{1}H) or carbon-13 (^{13}C) NMR to large proteins or nucleic acids in 2-dimensional NMR.[52] NMR spectra are unique, well resolved, analytically tractable, and often highly predictable. NMR can only be practiced where nuclei possess a non-zero spin. Molecules with hydrogen are well-suited because ^{1}H has a spin of ½, at 99.9% abundance of that isotope. In contrast, carbon, nitrogen and oxygen are not easily visible by NMR, at least for their most abundant isotopes (^{12}C, ^{14}N and ^{16}O), for which samples have to be prepared specially. For DNA in solution, deuterated solvents are helpful in singling out hydrogens unable to exchange with the environment. Deuterium(which has a spin of 1) is invisible in ^{1}H NMR because it has a different resonant frequency due to the presence of an additional neutron. This technique was used for studying the Dickerson sequence in this thesis. There are a few important concepts for understanding the basics of NMR spectroscopy. These include chemical shift, J-coupling, Nuclear Overhauser Effects (NOEs) and spin diffusion.[52]

3.2 Chemical Shift

A spinning charge generates a magnetic field, which has a magnetic moment (μ) that is proportional and oriented perpendicularly to the axis of rotation. Analogous to bodies rotating in space, spin can be viewed as a type of angular momentum. These momenta point in stochastic directions in materials until an external magnetic field (B_0) is applied, conventionally aligned along the +z axis in NMR. In ^{1}H nuclei, there are two spin angular momentum states that can exist: +½ and -½.[53] The difference in energy between the two spin states depends on the strength of the external magnetic field as shown in F3.2.1. The difference is usually very small, leading to the requirement of strong magnets (~1-20 Tesla) to resolve them.

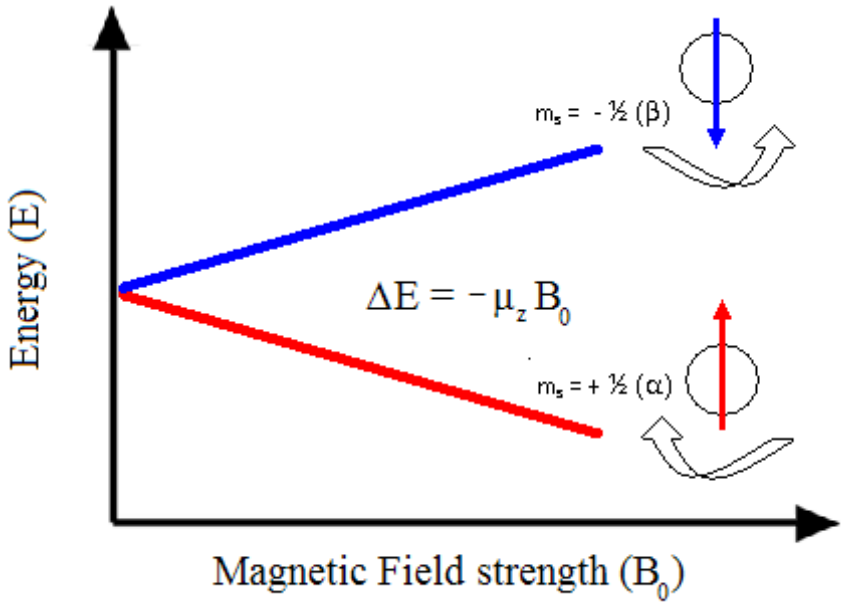


Figure 3.2.1. The spin states of ^{1}H, a magnetic dipole.[53] The two spin states have the same energy where B_0 is equal to zero. As the B_0 strength increases, the spins diverge. At any given field strength, the formula for the difference in E (eV) is provided in terms of the B_0 (T) and magnetic moment μ (eV/T).

Irradiation of the sample with a radio frequency (RF) pulse with energy corresponding to the exact spin state separation of a specific set of nuclei can cause the excitation of those nuclei to the next higher quantum energy state. As previously stated, for spin ½ nuclei such as hydrogen, the energy difference between the two spin states at a given magnetic field strength is proportional to their magnetic moment. However, even if all protons have the same magnetic moments, they do not give resonant signals at the same frequency values. The difference arises from the varying electronic environments around the nucleus of interest.[53]

Electrons around a nucleus can move in response to an applied magnetic field, creating localized magnetic fields that oppose the stronger applied field B_0. These fields can "shield" the proton from B_0 so that it must be increased to achieve resonance.[53] Such increments, known as chemical shift, are usually very small, but carry important information. Since shielding is a result of the local magnetic environment, the chemical shift provides information on the structure of the molecule.

The location of the chemical shift signal is dependent on the strength of the external magnetic field as well as the resonance frequency. In order for comparison between data sets run at different field strengths, chemical shifts are reported in parts per million (ppm) relative to a standard, as per the following equation:

$$\delta\ (ppm) = \ 10^6 \cdot \frac{v - v_0}{v_0}$$

Here, v is the signal frequency in Hertz (Hz) and v_0 is the signal frequency of the reference compound.[52] Important factors influencing chemical shift include electron density, electronegativity of neighboring groups, and anisotropic effects induced by the magnetic field. As an example, a nucleus in the vicinity of a more electronegative atom

will experience a reduction in electron density. It will thus be de-shielded and the signal will appear further downfield (higher ppm). F3.2.2 shows the chemical shifts of protons for common functional groups.

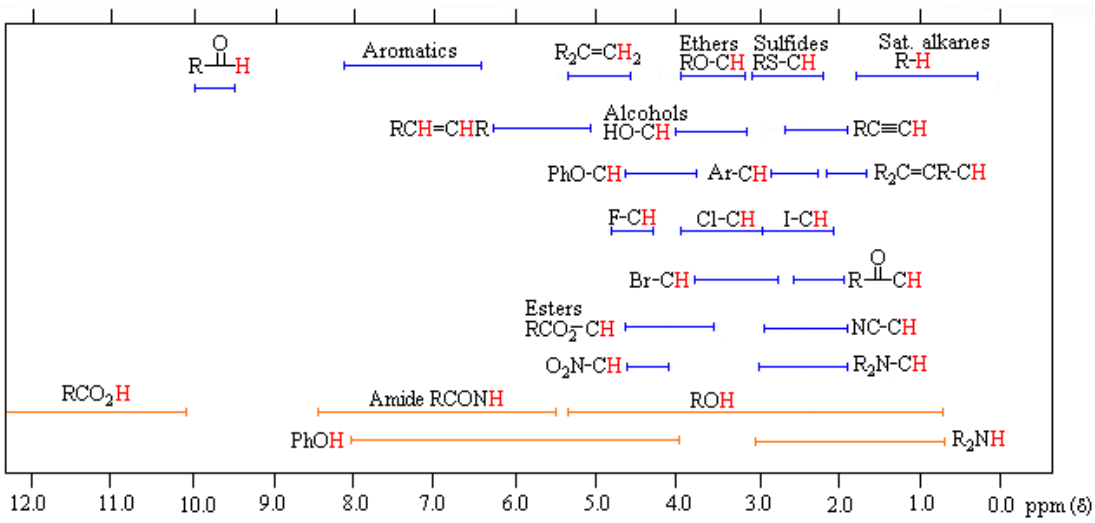


Figure 3.2.2. Chart showing the general distribution of proton chemical shifts associated with different functional groups.[53]

When comparing molecular derivatives by NMR, a peak that shifts upfield (smaller ppm) reflects a shielded nuclear environment. This can be caused by a greater electron density, which protects the nucleus from the external magnetic field. Conversely, a downfield (more positive ppm) indicates deshielding, such as when electron density is removed by that nucleus moving closer to an electron-withdrawing group. Other deshielding effects, which spread density away from the nucleus, include increased π-electron components of bonding (e.g. alkenes) and electron delocalization of aromatic compounds.

3.3 Scalar Coupling

Scalar coupling or J-coupling refers to through-bond interactions between two nuclei (typically represented as A and X) with a non-zero spin that cause NMR signals to split into multiple peaks.[52] The interaction is indirect and mediated by the electrons of the two nuclei: one spin perturbs the spins of the shared electrons, which will then perturb the second spin. Since scalar coupling is only propagated through bond orbitals, it can be used to tell which atoms are connected. For example, if two ½ spins are scalar-coupled, the spectrum of each will be a doublet with the separation between the two lines called the coupling constant J.[52] Coupling constants are denoted ${}^{n}J_{AX}$, where A and X are the interacting nuclei and n is the number of covalent interceding bonds. One-bond coupling (${}^{1}J$) are an order of magnitude larger the two- and three-bond couplings (${}^{2}J$ and ${}^{3}J$ respectively).[52] Splitting patterns for spin ½ nuclei are symmetrically distributed on both sides of the chemical shift. The central lines always have a stronger intensity than the outer lines. Some commonly observed patterns are shown in F3.3.1: a doublet has two equal intensity signals, while a triplet has three signals with an intensity ratio of 1:2:1 and a quartet has four signals with an intensity ratio of 1:3:3:1.

Figure 3.3.1. Examples of J-coupling patterns commonly observed in proton NMR spectra.[53]

The splitting patterns of Figure3.3.1 display ideal or "First-Order" arrangement of lines.[53] This is usually observed if the spin-coupled nuclei have very different chemical shifts ($\Delta\nu$ is large compared to J_{AX}). However, if the coupled nuclei have similar chemical shifts, the splitting patterns can be distorted. This is known as second-order behavior. If the chemical shifts are the same, then signal splitting disappears entirely and a singlet is observed. The splitting pattern for a given nucleus can be predicted by the n+1 rule, where n is the number of neighboring spin-coupled nuclei with the same or similar coupling constants. On the other hand, if a nucleus is spin-coupled to two or more sets of neighboring nuclei by different J_{AX} values, the n+1 rule cannot predict the entire splitting pattern. In this case more complex splitting patterns will be observed.

3.4 One-Dimensional NMR Spectroscopy

Measurement of a one-dimensional (1D) NMR experiment is carried out in two stages: preparation and detection as shown in F3.4.1. During preparation, the sample must be set to a defined spin state. Once it does, a radio frequency signal is transmitted with sufficient power (~50W) and time period (~few microseconds). In the simplest case, this is a 90^o pulse, which rotates the equilibrium magnetization M_z onto the Y-axis (M_y).[53] After this pulse, each spin precesses with its own Larmor frequency around the Z-axis and induces a signal in the receiver coil. The signal then decays, with spin-lattice (t_1) relaxation in the Z plane and spin-spin (t_2) in the XY planes. The entire process, (i.e. the variation of magnetization with time) is known as the free induction decay (FID).

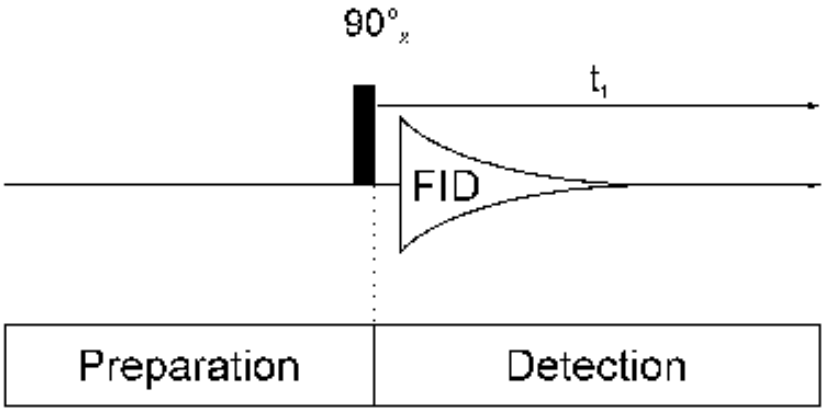


Figure 3.4.1. Schematic diagram of how a 1D NMR experiment is carried out and the creation of a FID.[53]

The information in the FID signal is incomprehensible in its native form, so it must be transformed from a time domain to a frequency domain. This is done using a mathematical procedure called a Fourier Transform (FT), which gives a spectrum of intensity versus frequency and is simplified in F3.4.2. Identifiable peaks appear, which allow for sample analysis.

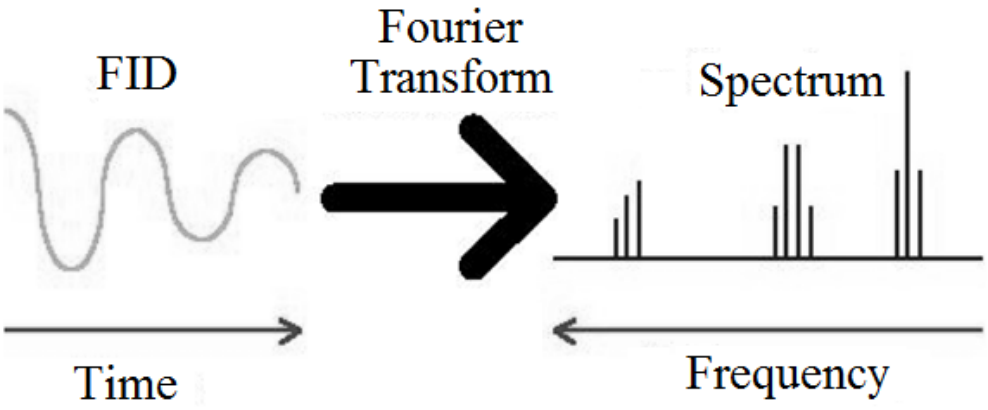


Figure 3.4.2. Fourier transform takes the FID signal produced by an NMR instrument and turns it into the NMR spectrum with assignable peaks.[53]

3.5 Two Dimensional NMR Spectroscopy

Two-dimensional (2D) NMR was first introduced in the early 1970s and it has since been widely used to correlate the resonance frequencies of several nuclei.[52] In contrast to 1D NMR, 2D NMR has four stages as demonstrated in F3.5.1. The preparation stage is similar to a 1D experiment. The key step in a 2D NMR experiment is the mixing. This allows the magnetizations or coherences to exchange through multiple types of interactions, such as scalar coupling or the nuclear Overhauser effect (NOE).[52] Finally, the signal is sampled during the detection period (t_2).

Although two frequency labeling periods are present, the signal is indirectly detected during t_1 due to the "memory" of the spins.[52] This generic 2D NMR scheme can be used to generate a homonuclear spectrum (correlates ^{1}H with ^{1}H frequencies) as well as a heteronuclear spectrum (correlates ^{1}H with ^{15}N frequencies or ^{1}H with ^{13}C frequencies). There are a number of variants conceived from each basic type but only nuclear Overhauser effect spectroscopy (NOESY) and total correlation spectroscopy (TOCSY) will be discussed in this thesis.

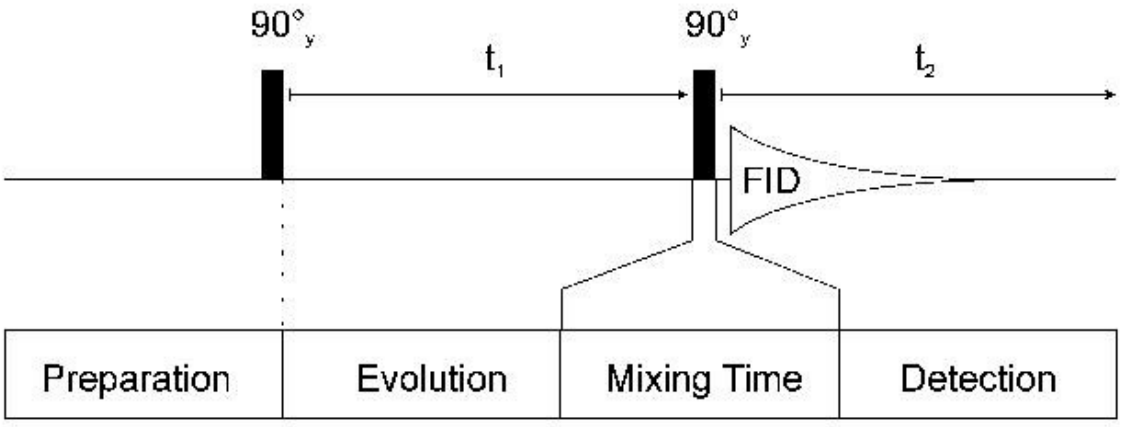


Figure 3.5.1. Sketch of a 2D NMR experiment, depicting the RF pulses at 90° to one another. [53]

3.6 Nuclear Overhauser Effect Spectroscopy (NOESY)

The NOESY signal is a measurement of the equilibrium state after perturbation of spin magnetization through space. This is possible where the spins of nuclei are near enough in space to exhibit dipolar coupling. When the resonance of one is perturbed by a magnetic field, a change in resonance affects the other, termed the NOE. Since these are through-space interactions, they are not mediated through J-coupling, which is a through-bond phenomenon.[54] The signal intensity can be calculated in the following equation:

$$r_{a,b} = r_{ref} \sqrt[6]{\frac{I_{ref}}{I_{a,b}}}$$

As can be seen in the equation, the NOE intensity (I) is inversely proportional to the 6th power of distance (r) between two nuclei (a and b). This equation is particularly useful because it allows for internuclear distances to be calculated from NOE intensities if a reference NOE and distance are known.

3.7 Spin Diffusion

Signal intensity increases linearly with mixing time when two nuclei are involved in the spin relaxation pathway. However, when other nuclei offer additional routes for relaxation, the trend is not linear.[55] To validate internuclear distances derived by NOESY, a build-up curve must be produced, which plots signal intensity over mixing time. This process is termed spin diffusion and limits the accuracy of internuclear distances because the signal contains no information about how many atoms were involved, the amount of relaxation each contributes, and the type of transition made. To understand these types of transitions, which may impart a positive or negative alteration to the NOE signal, a more detailed description of the effect is required.

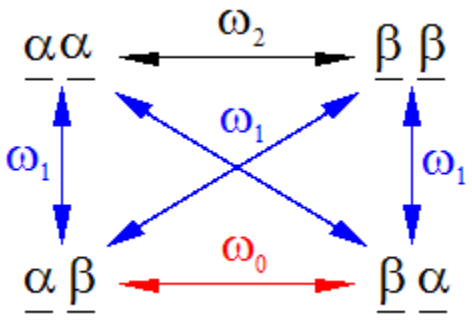


Figure 3.7.1. Two coupled nuclear spins and the spin transitions they can make. ω_0 transitions yield negative NOE signals, ω_1 have no change, and ω_2 signals generate positive ones.

For two proximal protons, each possessing a spin up (α) or spin down (β), four possible energy states are available: αα, αβ, βα, and ββ. Once the magnetic field is applied at the start of the experiment, these spins become saturated (equal in population). Signal is produced when the spins are allowed to relax to their equilibrium state. As can be seen in F3.7.1, there are 12 possible spin transitions, represented by ω. The overall sign and magnitude of the NOE (including spin diffusion) depends on which spin transitions are involved, as can be seen in the Solomon equation below.[55] The gyromagnetic ratios (γ) of the two nuclei in 1H NMR are the same.

$$\varepsilon_{NOE} = \frac{\gamma_1}{\gamma_2} \cdot \frac{\omega_2 - \omega_0}{\omega_0 + 2\omega_1 + \omega_2}$$

Positive NOEs result when the transition increases the population difference between spin states, such as when αα→ββ and ββ→αα occurs. Negative NOEs appear with αβ→βα and βα→αβ transitions because, while the rest of the molecule's nuclei equilibrate and produce a signal, these transitions result in a net decrease in the population difference.

It is important to point out that the dominant relaxation pathway depends highly on molecular motion.[55] Spin transitions occur when the frequency of the difference between energy levels matches the frequency of molecular tumbling. More mobile molecules tend to generate positive NOEs because they have fast relaxation times, which favor ω_2 transitions. Since the NOESY spectra obtained for the Dickerson sequences were in D_2O, all have positive peaks. Spin diffusion, on the other hand, is less predictable and it is essentially impossible to determine the sign and magnitude of the several transitions involved. Additional nuclei can increase or decrease the signal in spite of the solvent.

Owing to its erratic nature, only nuclei which exhibit minimal spin diffusion can produce signals that correlate with distance. For this reason, only nuclear distances derived from linear build-up curves can be included in the Dickerson model as geometric constraints.

CHAPTER 4. MATERIALS AND METHODS

4.1 DNA Preparation

Two DNA sequences were purchased from SynGen™: the Dickerson dodecamer, ds(CGCGAATTCGCG) and the dodecamer with an 8-oxo-deoxyadenosine substituted at the 5th position, ds(CGCGOA_5ATTCGCG).[56] For a detailed description of the 8-oxo-dA5 lesion, refer to F1.3.1. The two dodecamers are represented in F4.1.1.

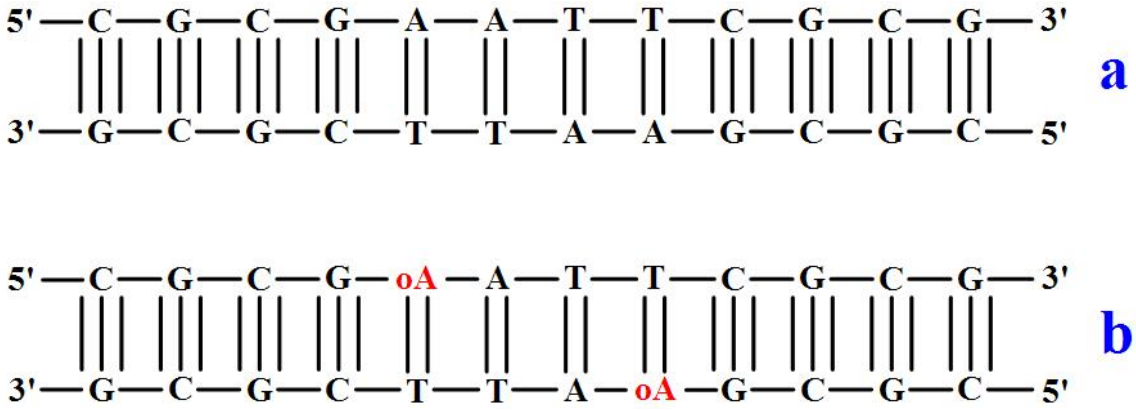


Figure 4.1.1. The Dickerson dodecamer without (a) and with (b) the 8-oxo-dA5 lesion.

Dry DNA samples were dissolved in 25 mM phosphate buffer using a stock phosphate buffer solution of 100 mM with 1% sodium azide (for antimicrobial protection), and adjusted to a pH of 7.4. To ensure all DNA was double stranded, samples were heated in a water bath at 90 °C for 15 minutes, then cooled for at least three hours. From there, the pH was adjusted to 7.4 again and the samples were lyophilized and redissolved in D_2O to remove residual H_2O. The process of lyophilization, redisollution in D_2O, and pH adjustment was performed three more times. All final DNA samples were at 5 mM.

4.2 NMR Studies

All ^{1}HNOESY experiments were performed on 5 mM DNA samples with a Varian INOVA 400 MHz NMR at Missouri State University. With the exception of the 25 °C study for the build-up curves, sample temperatures were kept at 10 °C. The spectral width was set to capture between 9.50 ppm and -0.50 ppm. For estimating internuclear distances, the mixing times were set to 50, 75, 100, and 200 ms with a relaxation time of 3s. The ^{1}HNOESY operated at 128 scans per increment and 128 increments. Since it is highly dependent with temperature, an internal reference to the water peak was taken from Gottlieb et al.,[57] which defined the center of the spectrum at $\delta = 5.051 - 0.0111T$. As per this equation, 10 °C spectra were referenced to 4.94 ppm, while those of 25 °C were to 4.77 ppm. Assignment resolution was 0.01 ppm. Sample integrity was validated by the sharp signal:noise ratio of the 1D presaturation spectrum, shown in F4.2.1.

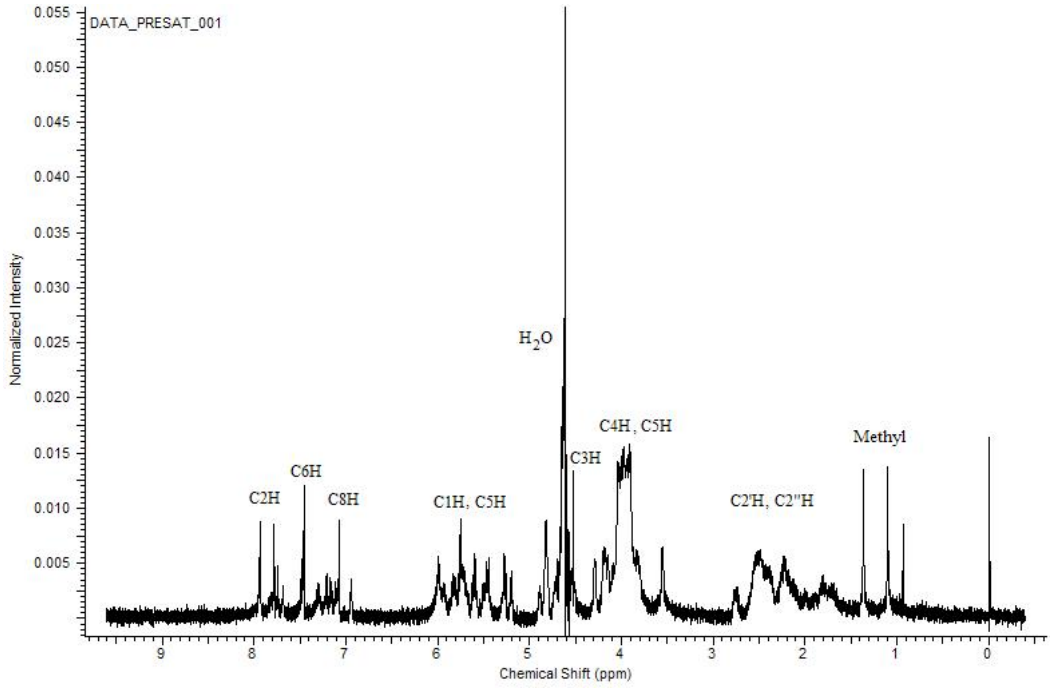


Figure 4.2.1.The 1D ^{1}H NMR spectrum of the Dickerson dodecamer. A general placement of peaks is indicated for the non-exchangeable hydrogens.[37]

4.3 Modeling Software

Varian FID files were assigned for intensities in ACDLabs NMR Processor v12.01-39104 and MestReNova 10.0.2-15465.[59,60] Once the NOE intensities translated into distance measurements, those possessing linear build-up curves were used as initial distance restraints for modeling in Gaussian 09,[60] using an ONIOM multilayered model. The 8-oxo-dA5 base was set to the B3P86 functional under the 6-31G (d,p) basis set, while the rest of the molecule ran under PM6. CHARMM PARNAH1ER1, while ideal, was not available in this version. Water was simulated using IEFPCM. All default coordinates not restricted by the NMR-derived distances were taken from the 1DUF model on the RCSB Protein Data Bank by Tjandra, et al.[40] Helical parameters were found by processing the Gaussian-derived DNA model files in the online 3DNA software package.[51]

4.4 ^{1}H NOESY Assignment of the Dickerson and 8-oxo-dA5 Dodecamers

This section will take the form of a walkthrough with images for clarity because the assignment of an NOE is easily made confusing by rote directions. In addition, the jargon inherent in the naming of peaks can take some practice. Extra attention has been paid to lowering these burdens by focusing on the general concept.

F4.4.1 portrays an annotated NOESY spectrum zoomed in to where the T7 Base-CH_3 interaction would be found for the Dickerson dodecamer. Previously, chemical shifts have all been assigned by 1D NMR and it is now a matter of finding their intensities in 2D. If no peaks appear in an expected region, those nuclei are too far apart for their NOEs to be detected.

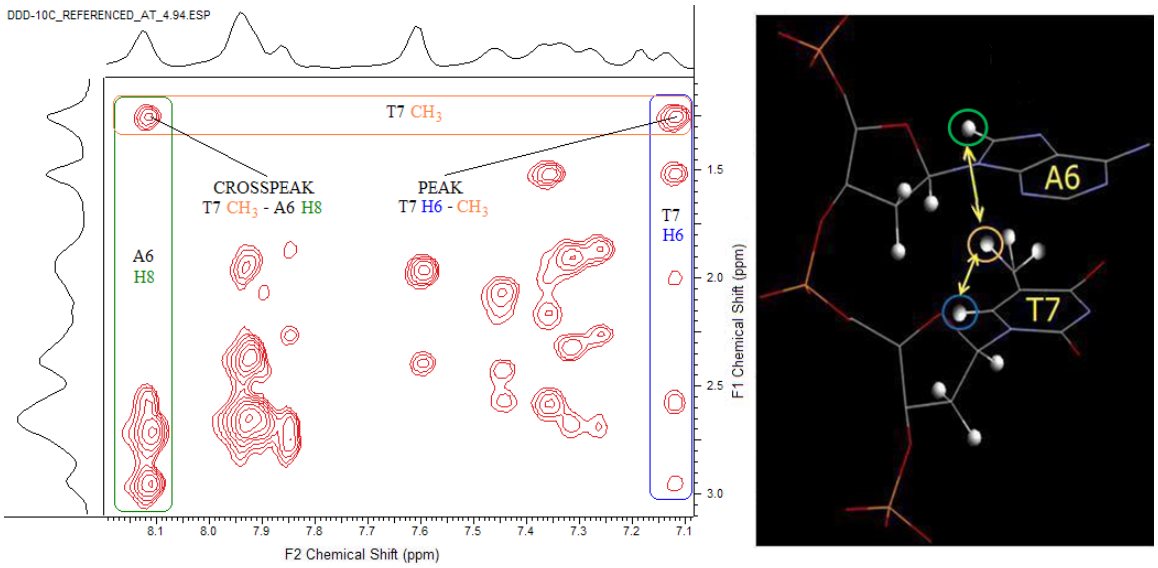


Figure 4.4.1. Example assignment of the T7 Base-CH_3peak and T7 CH_3- A6 Base crosspeak. Peaks represent interactions on the same residue, while crosspeaks are for separate residues. The spectrum is zoomed in to the region where the T7 CH_3, T7 H6, and A6 H8 peaks intersect in terms of their chemical shifts (left). Also indicated is a graphical manifestation of the through-space interactions between these atoms (right). NOEs associated with CH_3 groups represent the average interactions of the 3 equivalent protons.

To conceptually review, a non-degenerate NOE peak will appear in the 2D NMR spectrum where two hydrogens are close enough in space to interact by magnetic ^1H dipolar coupling strongly enough to measure. The orange box in F4.4.1 at 1.25 ppm represents the CH_3 on the thymine at position 7 on the dodecamer. Any NOEs at this chemical shift represent nuclear spin transfer between the T7 CH_3 and some other nearby hydrogen. The blue box at 7.09 ppm is where interactions appear with the T7 Base (C6H) hydrogen. Since there is an NOE at the crossroads of 1.25 ppm and 7.09 ppm where the boxes overlap, it follows that this is the T7-Base-CH_3 NOE peak, represented by the yellow arrow in the wireframe image to the right of the spectrum. All NOEs are assigned in this general way, although the existence of spurious, missing, shifted, and overlapped peaks must be considered. F4.4.2 shows an example of overlapped peaks and how to disambiguate them.

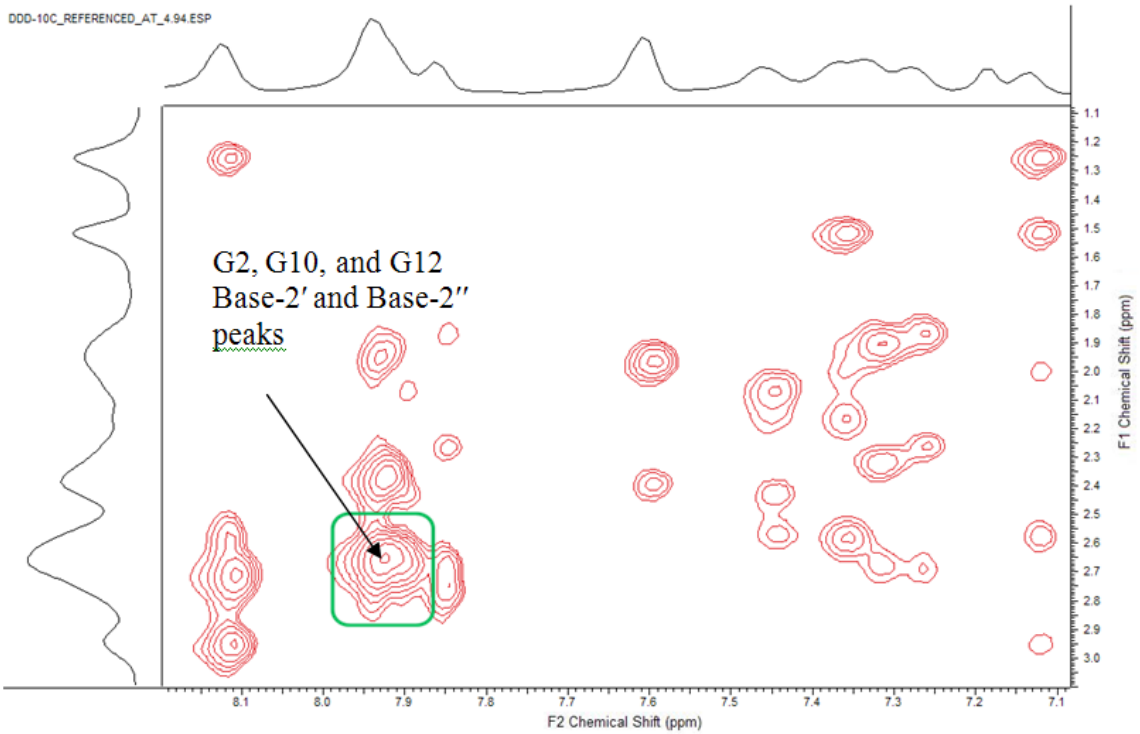


Figure 4.4.2. Peak overlap for the G2, G12, and G10 Base-2′ and Base-2″ NOEs.

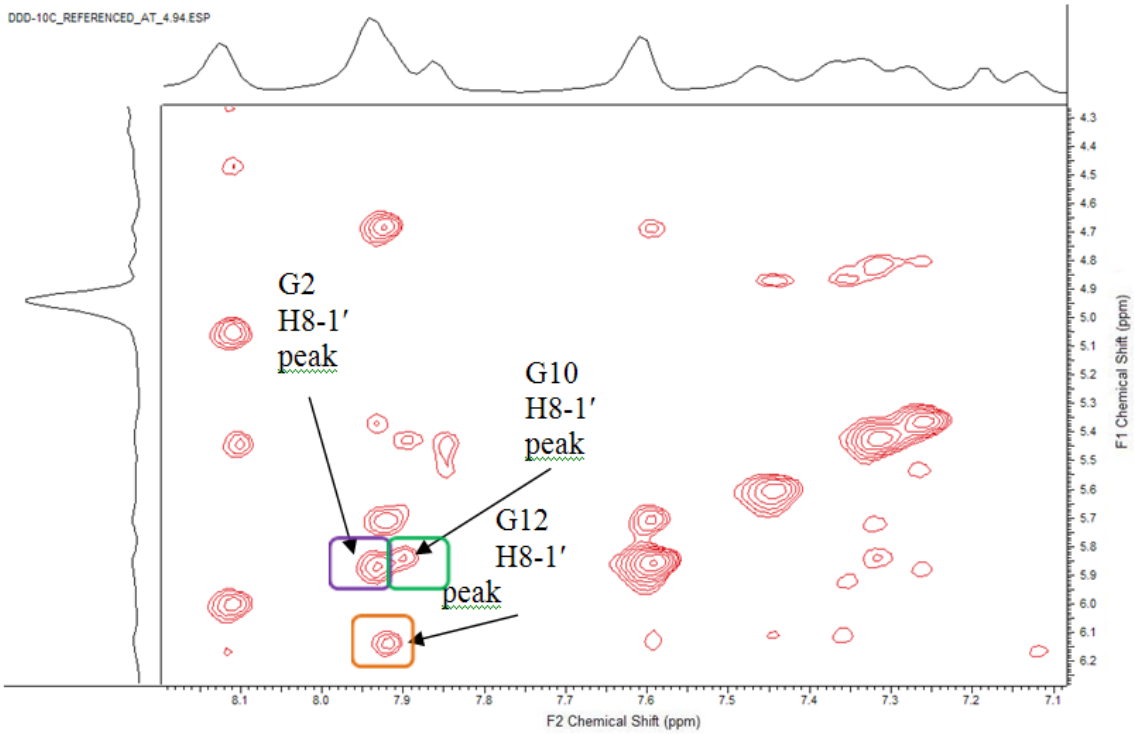


Figure 4.4.3.The H8-1′ peaks for G2, G10, and G12 do not overlap in this case, allowing for partial assignment of the overlapped H8-2′ and H8-2″ peaks because the H8 chemical shifts appear.

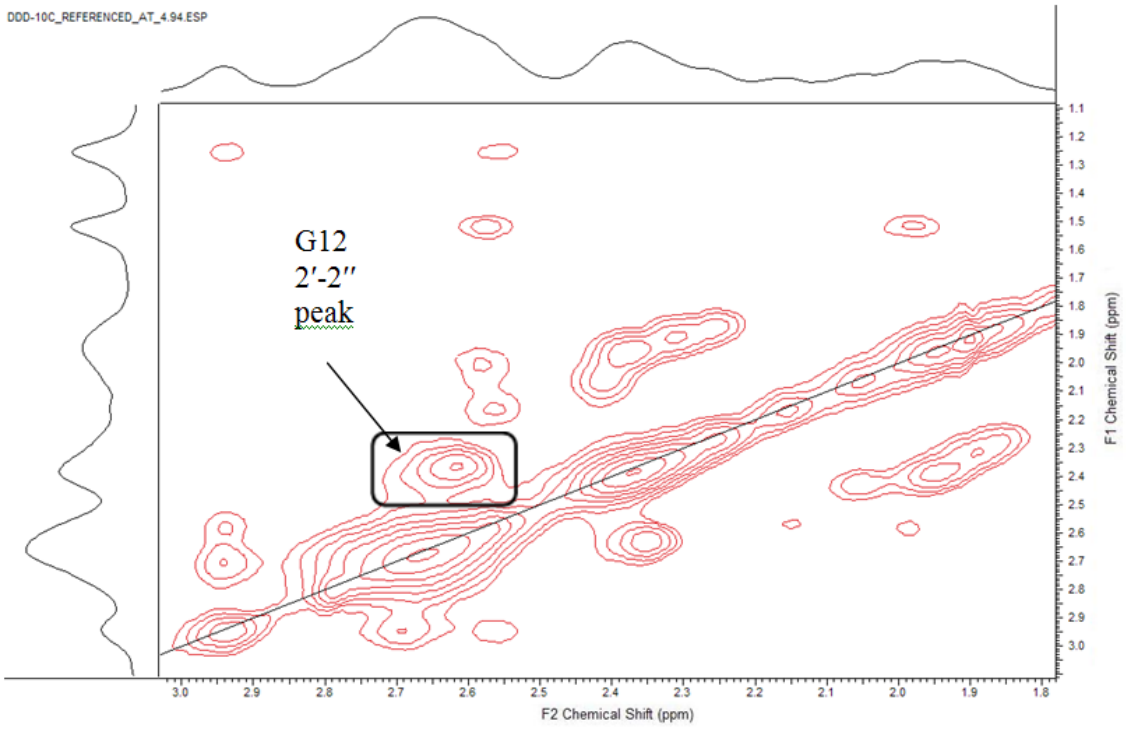


Figure 4.4.4. The 2′-2″ peak for G12, located in the 2′-2″-CH_3 region. Since the H8 chemical shift was found in Figure 4.4.3, these chemical shifts allow for the assignment of the overlapped G12 H8-2′ and H8-2″ peaks in Figure 4.4.2.

Shown in F4.4.2 is the base-2′-2″-CH_3 region (the same as F4.4.1) and spotlighted in the blue box is the spot where the H8-2′ and H8-2″ NOEs for G12, G10 and G2 all overlap. Disambiguation of these peaks involves looking in different regions and trying to assign related peaks. In F4.4.3 (the Base-1′-3′region), there is some clarity.

Since the H8 and 1′ interactions for G12, G10, and G2 are fully resolved, their chemical shifts can be assigned. The H8 chemical shift is of specific interest here and is half of the problem solved. Next, the 2′ and 2″ coordinates must be found.

Looking in the 2′-2″-CH_3 region, there is a peak for the 2′-2″ interaction of G12. This allows the H8-2′ and H8-2″ peaks to be separated from G2 and G10. Using this technique and looking elsewhere in the spectrum, the G2 and G10 coordinates can be

separated, but this is not always so. If every peak and crosspeak is overlapped, disambiguation is impossible.

Assignment of the 8-oxo-dA5 sample relies on the premise that nuclei closest to the lesion will be the most affected. Accordingly, they should have the greatest changes in chemical shift. The lesion also replaces the H8, meaning that there will be no NOEs associated with the base on A5.

One further example of DDD peak assignment will be covered, named after the "walking" mechanic of sequentially aligning nuclei in a NOESY region by their closeness in space. This is particularly useful for DNA, since the atoms involved are iterative over the length of the molecule, making the peaks predictable. This is demonstrated in the aromatic-1′ walk in Figure 4.4.5.

The NOESY walk operates on the principle that interacting ^{1}H nuclei produce NOEs where their chemical shifts meet on the F1 and F2 axes. This can be mapped wherever a horizontal or vertical line connects two NOEs. This becomes complicated where nuclei have the same chemical shifts, leading to peak overlap and several NOEs aligned in a row, but the NOESY walk can still be performed.

A straightforward example in the aromatic-1′ NOESY walk is where lines can be drawn between the T7 and its two neighbors T8 and A5. Since the chemical shifts for the T7 H6 and 1′ are known from the 1D NMR and they produce an NOE, drawing straight lines in the horizontal and vertical from this coordinate indicates wherever those nuclei interact with other nuclei. In F4.4.5 A, the lines run into NOEs that correspond to chemical shifts for the A6 1′ (by the vertical) and T8 H6 (horizontally). When the two NOEs are assigned based on their chemical shifts, they can serve as new focal points for

horizontal and vertical lines to be drawn. In this way, nearby nuclei generate an NOE pattern that can be “walked”, reflecting the pattern of their spatial orientation.

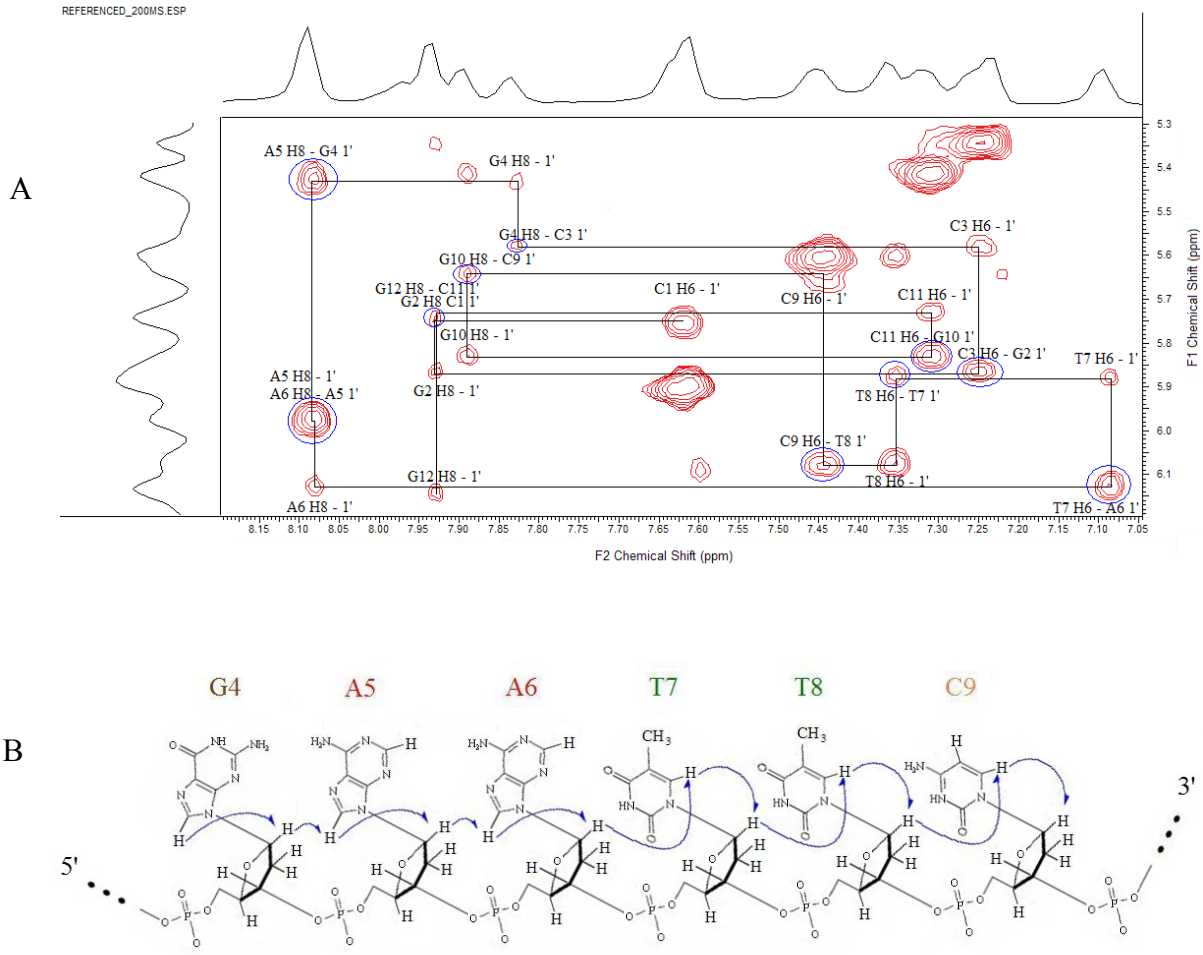


Figure 4.4.5. The Base-1′ walk. A) The walk from G1 to G12 on the NOESY spectrum. Encircled in blue are intermolecular NOEs. B) The walk as imagined geometrically from G4-C9, showing inter- and intramolecular NOEs in blue.

CHAPTER 5. RESULTS AND DISCUSSION

5.1 Chemical Shift Assignment

T5.1.1-.2 display the chemical shift assignment at 10 °C for each oligomer, referenced to 4.94 ppm. Changes in chemical shift are shown in T5.3. The G4, A5, A6, and T7 show prominent differences, so they will be discussed individually in the following sections.

Table 5.1.1. Assignment of the chemical shifts (ppm) for non-exchangeable ^{1}H nuclei in the Dickerson dodecamer at 10 °C, referenced to 4.94 ppm.

Residue	Base	H5/Me	2′	2″	1′	3′
C1	7.60	5.86	1.97	2.40	5.71	4.70
G2	7.93	N/A	2.63	2.68	5.88	4.96
C3	7.27	5.37	1.88	2.26	5.53	4.81
G4	7.86	N/A	2.68	2.75	5.45	5.02
A5	8.11	N/A	2.71	2.95	5.98	5.05
A6	8.11	N/A	2.58	2.95	6.14	5.00
T7	7.13	1.25	2.00	2.57	5.91	4.84
T8	7.37	1.52	2.17	2.57	6.12	4.87
C9	7.44	5.61	2.07	2.43	5.65	4.87
G10	7.90	N/A	2.63	2.68	5.84	4.96
C11	7.32	5.42	1.91	2.32	5.72	4.82
G12	7.93	N/A	2.37	2.6	6.14	4.68

Table 5.1.2. Assignment of the chemical shifts (ppm) for non-exchangeable ^{1}H nuclei in the 8-oxo-dA5 Dickerson dodecamer at 10 °C, referenced to 4.94 ppm.

Residue	Base	H5/Me	2′	2″	1′	3′
C1	7.65	5.93	2.03	2.44	5.76	4.71
G2	7.93	N/A	2.63	2.67	5.89	5.00
C3	7.24	5.36	1.87	2.24	5.55	4.77
G4	7.74	N/A	2.51	3.03	5.96	4.96
O-A5	**N/A**	**N/A**	**2.45**	**3.40**	**5.54**	**4.97**
A6	8.24	N/A	2.46	2.84	6.21	5.00
T7	6.93	0.36	1.96	2.57	5.96	4.84
T8	7.33	1.57	2.11	2.57	6.02	4.86
C9	7.47	5.63	2.06	2.39	5.67	4.87
G10	7.97	N/A	2.63	2.67	5.89	5.00
C11	7.35	5.45	1.93	2.34	5.74	4.85
G12	7.96	N/A	2.36	2.63	6.16	4.67

Table 5.1.3. Difference in chemical shifts between the 8-oxo-dA5 and control Dickerson at 10 °C. Colors indicate the magnitude of alteration, in the order of red > orange > yellow and white representing a negligible effect.

Residue	Base	H5/Me	2′	2″	1′	3′
C1	0.05	0.07	0.06	0.04	0.05	0.01
G2	-0.01	N/A	0.00	-0.01	0.01	0.04
C3	-0.03	-0.01	-0.01	-0.02	0.02	-0.04
G4	-0.12	N/A	-0.25	0.35	0.51	-0.06
A5	Lesion	N/A	-0.27	0.45	-0.44	-0.08
A6	0.13	N/A	-0.12	-0.11	0.07	0.00
T7	-0.20	-0.89	-0.04	0.00	0.05	0.00
T8	-0.04	0.05	-0.06	0.00	-0.10	-0.01
C9	0.03	0.02	-0.01	-0.04	0.02	0.00
G10	0.07	N/A	0.00	-0.01	0.05	0.04
C11	0.03	0.03	0.02	0.02	0.02	0.03
G12	0.03	N/A	-0.01	0.01	0.02	-0.01

5.1.1 Assignment of A5

The A5 base NOE is absent in the 8-oxo-dA5 Dickerson as expected, since the lesion replaces the proton at that spot. As such, the 2′, 2″, 1′, and 3′ peaks for A5 can only be determined by their interactions with each other and with their nearest neighbors, not by alignment with the H8 (base) peak. The 2′-2″ interaction was found at the 2.45 ppm and 3.40 ppm crosspeak. The A5 2′ peak had a decrease in chemical shift, while the 2″ had an increase, even though these protons are very close in proximity. From this, it may be inferred that the electron-withdrawing effect of the nitrogenous base, which is stronger for the 2″ proton than the 2′ in the unmodified DDD, was magnified by the 8-oxo-dA5 adduct. There are also peaks at 5.54 ppm with these two representing interactions with the A5 1′ proton, which was expected to appear in the 5-6 ppm range with the rest of the 1′ protons. There is no evidence to suggest the 2′ and 2″ protons have swapped order as

compared to the Dickerson control, which does occur in the 8-oxo-dG4 DDD as seen in T5.4. Furthermore, the 2′ and 2″ peaks, which were expected to possess crosspeaks with the 3′ proton, had crosspeaks at 4.97 ppm in the expected range of 4-5 ppm.

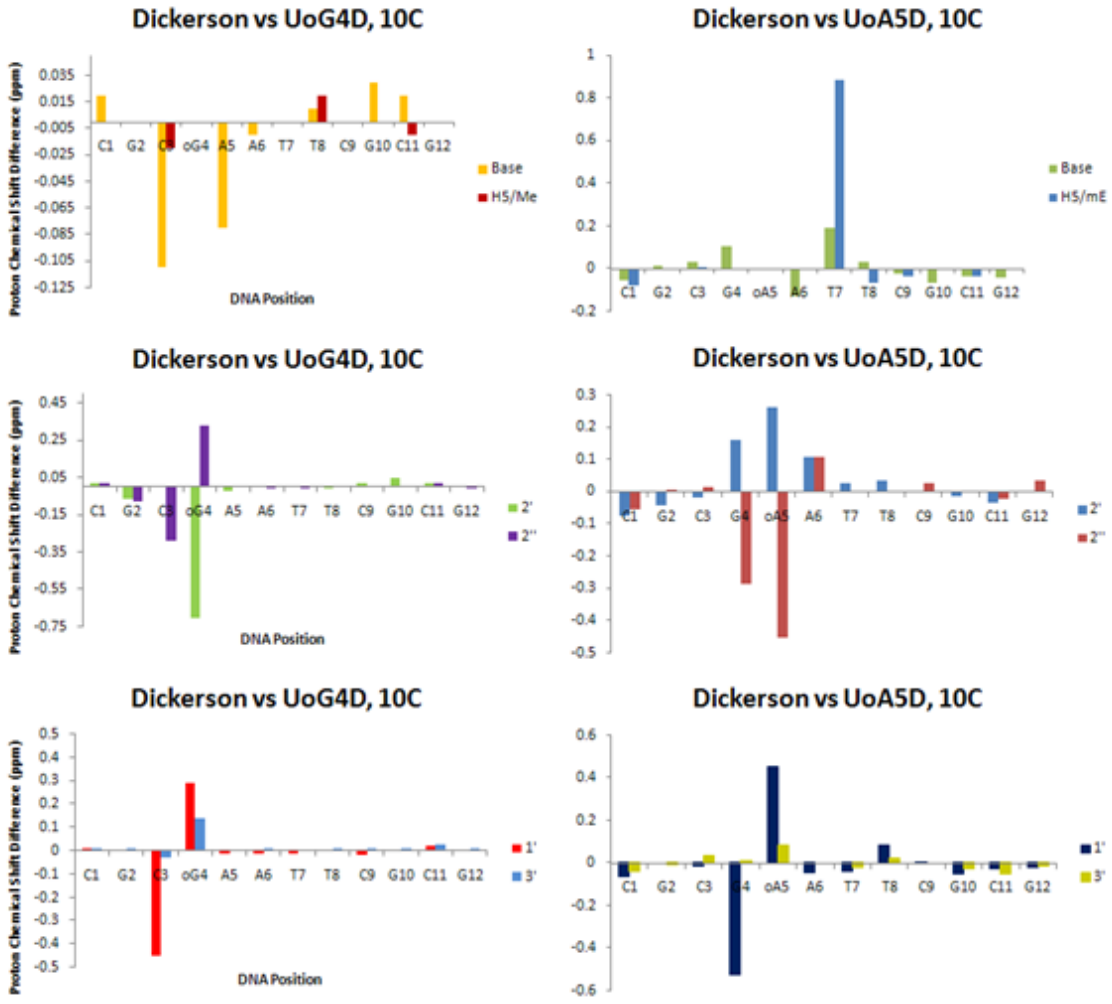


Figure 5.1.1. Differences in chemical shift between the 8-oxo-dG4 and 8-oxo-dA5 with the unmodified sequence.

The 8-oxo-dG4 and 8-oxo-dA5 share a similar NMR profile as can be seen in F5.1. For both sequences at the lesion site, there was a downfield chemical shift for the 1′ and 2′ protons and an upfield shift for the 2″ proton. Both sequences affected their 5′ neighboring residue by decreasing the chemical shift for their 1′ and 2″ protons.

Table 5.1.4. Assignment of the 8-oxo-dG4 DDD

Residue	Base	H5/Me	2′	2″	1′	3′
C1	7.62	5.90	1.97	2.40	5.73	4.70
G2	7.96	N/A	2.67	2.76	5.84	4.96
C3	7.38	5.37	1.85	2.55	5.99	4.85
O-G4	**N/A**	**N/A**	**3.36**	**2.41**	**5.15**	**4.85**
A5	8.19	N/A	2.70	2.94	6.00	5.06
A6	8.13	N/A	2.55	2.94	6.15	5.00
T7	7.13	1.23	1.97	2.57	5.92	4.83
T8	7.36	1.48	2.16	2.55	6.09	4.89
C9	7.46	5.59	2.03	2.40	5.67	4.86
G10	7.88	N/A	2.60	2.68	5.84	4.96
C11	7.32	5.43	1.88	2.30	5.69	4.80
G12	7.95	N/A	2.33	2.63	6.15	4.67

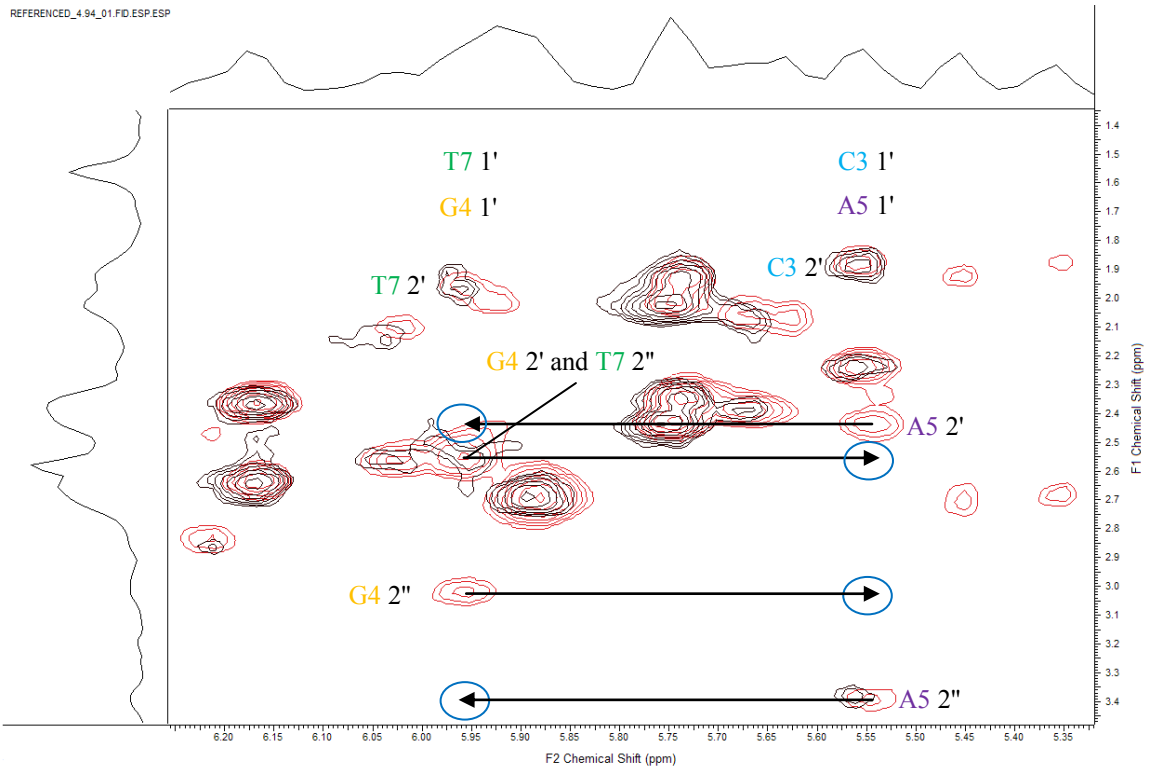


Figure 5.1.2. TOCSY (black) overlay on the NOESY (red) for the 8-oxo-dA5 DDD in the 1'-2'-2" region. Encircled in blue are missing crosspeaks between G4 and A5 that appear in the unmodified sequence. The 1' peaks for G4 and A5 are overlapped with those for the T7 and C3, respectively.

Table 5.1.5. Assignment disambiguation between G4 and A5. The first column represents a chemical shift (ppm) that could correspond to G4, A5, or A6. The second column indicates where NOEs appear at that chemical shift, providing clues about its assignment. The third column explains the disambiguation process. The fourth column identifies the residue(s) accounting for that evidence.

Assignment	Interaction	Explanation	Candidate
2.51	7.74	7.74 is the G4 base. Since the Dickerson G4 2' and 2" are at 2.68 and 2.75, it is likely that this is the G4 2'.	G4
3.03	7.74, 2.51, 5.96	3.03 is very close to the A5 2" on the Dickerson control, which is 2.95. Thus it was considered as either a crosspeak between G4 and A5 or the G4 H8 - 2". While a TOCSY was performed, no H8-2' peaks appeared for any residue, since they are separated by too many bonds. As a result, 3.03 was assigned as the G4 2" by way of there being no other candidates for the 2" position (unless the 2'/2" were overlapped).	G4
5.54	1.87, 2.24, 2.44, 3.40	1.87 is the C3 2', while 2.24 is the C3 2". If the G4 1' is here, it is not only overlapped with the C3 1', it has no G4 1'-2" peak. It seems more likely it is the A5 1' because it interacts with the A5 2' and 2'. Proof that this is the A5 1' appeared with the absence of a TOCSY peak for the supposed G4 and presence of a peak at 5.54 and 3.40, shown in Figure X.	A5
5.96	2.51, 2.98	On the Dickerson control, the G4 1' is 5.45 and the A5 1' is 5.98. By this account, 5.96 on the 8-oxo-dA5 is closest to being the A5 1', but this is misleading. 5.96 is interacting with 2.51 and 2.98, which as described above, are the G4 2' and 2". While A5 2' is also at 2.51, there is no A5 1'-2" NOE here. Thus, this must be the G4 1'.	G4
3.40	5.54, 2.45, 8.24	There is no corollary for 3.40 in the DDD. However, 5.54 is both the C3 1' and A5 1'. Judging it alone, it is unknown if this is a A5 peak or G4 crosspeak. However, since 3.40 interacts with 2.5, which is the A5 2' and 8.24, which is the A6 H8, this is likely an A5 assignment. There are no other A5 nuclei to assign, so by process of elimination, 3.40 was determined as the A5 2".	A5
2.45-2.51	6.21, 5.96	2.45 is the A6 2', which is very close to the G4 2' and the A5 2' at 2.46 and 2.51. Thus, any crosspeaks here can generate ambiguity over whether they involve G4, A5, or A6. This is a large and broad NOE. Unfortunately, with this ambiguity, the distances between these nuclei cannot be found, since there is no way to separate the intensity in this region for each.	A6, A5, and G4

Assigning G4 and A5 on the 8-oxo-dA5 sequence is challenging since both are shifted to where comparison with the unmodified DDD is not helpful. Most of their ambiguous peaks occur in the 1'-2'-2" region. Due to peak overlap and chemical shift redundancy, a NOESY was not sufficient for distinguishing through-space from through-bond peaks in this area. However, a TOCSY was useful to determine which, as shown in F5.1.2 because this method only produces peaks where nuclei interact through J-coupling, which is propagated via bonds only. A detailed description of the assignment and disambiguation process for this region is included in T5.1.5.

5.1.2 Assignment of G4

The G4 base peak shifted -0.11 from the Dickerson control, which is the same magnitude but opposite sign as A6. The G4 2′, 2″, and 1′ were shifted upfield much more than A6. Thus, it can be assumed that the 8-oxo-dA5 lesion had a greater impact on G4 than A6. Further evidence for this effect can be found in there being more crosspeaks between A5 and G4. Specifically, crosspeaks were found between the G4 H8 and the A5 2′ and 2″, while none were found with the A6 H8.

5.1.3 Assignment of A6

No new crosspeaks for A6 were observed. The base peak was shifted by +0.13 ppm, the 2′ by -0.12 ppm, and the 2″ proton by -0.11 ppm, making the effects of the 8-oxo-dA5 lesion apparent on A6, but not extreme. Whereas in the unmodified DDD, the A6 2″ had the same chemical shift as the A5 2″, this degeneracy no longer became the case with the 8-oxo-dA5 sequence, indicating that they are experiencing different electronic environments.

5.1.4 Assignment of T7

It is interesting that the T7 base changed by -0.20 ppm and the methyl by -0.89 ppm because the T7 is one residue further away from the lesion. However, considering that the dodecamer is dsDNA, it is more likely that the effects are being observed on the opposite strand, in which case the T7 is adjacent and diagonal to the lesion, being paired with A6. Why T8 is not shifted remains a mystery, since T8 directly base-pairs with A5, the lesion site. The extreme upfield change in the T7 methyl (shielding), which is greater than any change in A5 suggests several possibilities which will be noted, but are inconclusive by ^{1}H NMR alone:

1) The T7 methyl is gaining substantial electron density, shielding it. This could be possible if it was nearby an electron donating group or wedged into an area of higher electron density.

2) There is a distortion in the helix, causing the T7 methyl to be closer to A6. If this were so, then a crosspeak between one or more of the A6 and T7 protons would be observed. Indeed, crosspeaks were found between the T7 methyl and the A6 2′ and 2″. However, these are also observed for the Dickerson control. No other crosspeaks between A6 and the T7 methyl were found.

3) The damaged A5 base is extruded and twisted to where it is closer to the T7 methyl on the opposite strand. If this were so, then a crosspeak between one or more of the A5 and T7 protons would be observed. However, no crosspeaks were found for any of the A5 protons with any of the T7 protons. That stated, 5.92 ppm was the 1′ assignment for both A5 and T7, so if crosspeaks occurred there, they would not be separated from the peaks.

Since the T7 methyl is so strongly shielded, it is most likely interacting with a group with higher electron density. If this were so, then new peaks could be observed and the three protons might lose degeneracy. For the 8-oxo-dA5 DDD, a new peak was observed corresponding to the T7 methyl (0.05 ppm) interacting with the C9 base (7.16 ppm). These two are not normally in proximity, as a nucleoside separates them. It is unlikely that this is a real interaction, although 5 sample runs show that there is always a peak there. The methyl peaks did not split to remove the degeneracy on each hydrogen and no other remarkable interactions were found.

5.1.5 Missing and Additional NOEs

T5.1.6 lists the peaks and crosspeaks that are not shared in common between the 8-oxo-dA5 DDD and unmodified sequence. One explanation for why the T7-CH_3 was so strongly shifted is that it was behaving in an alternate cross-relaxation pathway. While crosspeaks consistently appear for the T7-CH_3 with C9 and G10 instead of A6, this pathway is highly improbable because several residues separate them. Furthermore, those distant residues do not have different chemical shifts from the unmodified DDD.

Table 5.1.6. Missing and additional peaks between the 8-oxo-dA5 and unmodified DDD sequences.

DDD	8-oxo-dA5	Nuclei involved	DDD	8-oxo-dA5
1.24 - 6.14	0.36 - 6.21	T7 CH_3 - A6 1′	present	absent
1.24 - 7.44	0.36 - 7.47	T7 CH_3 - C9 H6	absent	present
5.45 - 2.95	5.96 - 3.40	G4 1′ - A5 2″	present	absent
5.98 - 2.68	5.54 - 2.51	A5 1′ - G4 2′	present	absent
5.98 - 2.75	5.54 - 3.03	A5 1′ - G4 2″	present	absent
1.24 - 7.90	0.36 - 7.97	T7 CH_3 - G10 H8	absent	present

†overlapped in the unmodified sequence

††NOE appears consistently, but is an unlikely relaxation pathway due to distance

The significance of these peaks is inconclusive and requires further study. The T7-CH_3 has no other new or missing interactions to account for why it is so strongly shifted. G4 and A5 had strongly affected chemical shifts, with four missing NOEs in the 1′ - 2′ - 2″ region.

How exactly the nucleosides moved to account for the chemical shifts remains to be explained. From the ^{1}H NMR alone, only broad statements can be made. While it might be reasonable to suggest that the 8-oxo-dA5 base could possess greater electron-withdrawing capacity due to having an additional oxygen atom, it is involved in an aromatic ring system, which does not necessarily behave in that fashion. Geometric changes are also not simple to envisage when there are many atoms involved in DNA and only a few ^{1}H nuclei from these experiments. Additionally, T7, which had the most dramatic downshift in its methyl protons, had no crosspeaks and thus no NOE interactions with A5. From this, three inconclusive interpretations can be made for T7:

1) The oxidation of A5 may have produced a downstream or opposite strand change.

2) The T7 base and methyl were close enough to the 8-oxo-dA5 lesion that shielding would take place, but not close enough to participate in observable nuclear coupling.

3) The T7 base and methyl are being shielded by other atoms or a combination.

The paradoxically negligible effect on T8, which directly opposes A5, makes sense when considering that only exchangeable hydrogens are involved in Watson-Crick base pairing between A and T. Thus, the protons observed in this experiment are not in the hydrogen-bonding region. However, it is still unusual that no changes were observed when residues further away were affected.

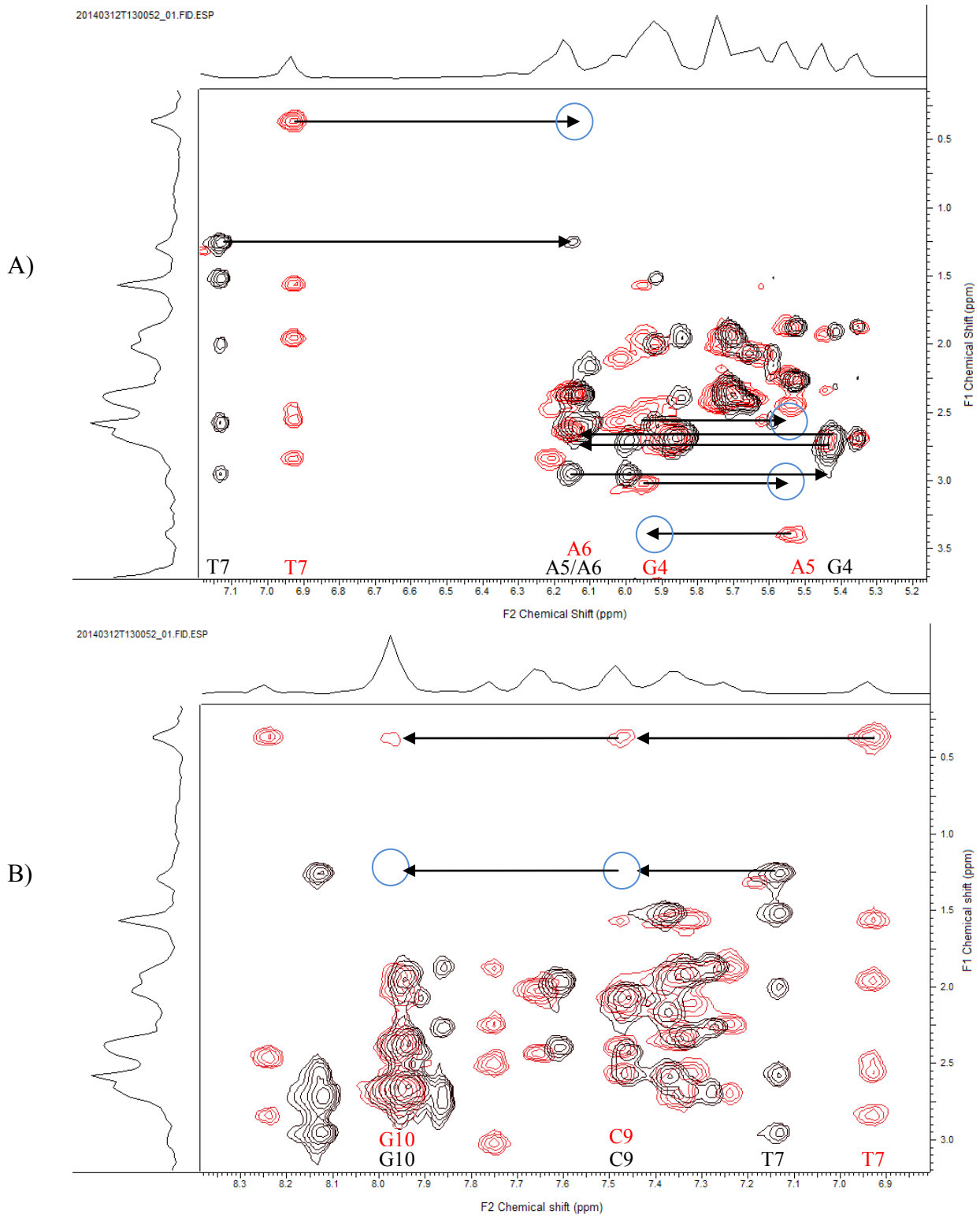


Figure 5.1.3. Overlapped spectra of the 8-oxo-A DDD (red) with the unmodified sequence (black) for the 1′-2′-2″ region (A) and Base-2′-2″ region (B). Missing crosspeaks from Table 5.6 are encircled in blue, with arrows pointing from the direction of the residue involved. Residues are indicated at the X-axis.

5.2 Temperature Studies

The assignments for the DDD and 8-oxo-dA5 were compared at 10C and 25C to survey for molecular mobility. It was hypothesized that the lesion might confer sensitivity to internal motion, which could account for the T7 methyl shielding. As can be seen in T5.7 and 5.8, no significant temperature changes were found.

Table 5.2.1. Difference in chemical shifts between10 °C and 25 °C for the 8-oxo-dA5 DDD.

8-oA5	Base	H5/Me	2′	2″	1′	3′
C1	-0.02	0.00	-0.05	-0.03	0.01	-0.04
G2	-0.03		-0.07	-0.01	-0.01	-0.01
C3	-0.02	-0.03	-0.03	-0.01	0.01	0.05
G4	-0.02		-0.05	-0.04	-0.04	0.01
A5			0.05	-0.05	0.03	-0.05
A6	-0.04		-0.01	-0.03	-0.07	-0.04
T7	-0.04	0.06	-0.02	-0.04	-0.04	-0.02
T8	0.01	0.01	0.01	-0.04	0.02	0.01
C9	-0.03	-0.01	-0.04	-0.02	0.00	-0.02
G10	0.00		-0.08	-0.02	-0.01	-0.01
C11	-0.03	-0.03	-0.04	0.00	0.02	-0.02
G12	-0.02		0.00	-0.03	-0.04	0.00

Table 5.2.2. Difference in chemical shifts between 10 °C and 25 °C for the unmodified DDD.

DDD	Base	H5/Me	2′	2″	1′	3′
C1	-0.05	-0.05	-0.08	-0.04	-0.01	-0.04
G2	-0.03	0.00	-0.04	-0.01	0.00	0.02
C3	0.00	0.00	-0.04	-0.01	0.08	-0.01
G4	-0.02	0.00	-0.02	-0.02	-0.02	-0.03
A5	-0.02	0.00	-0.02	-0.02	-0.02	-0.01
A6	-0.02	0.00	-0.01	-0.02	-0.01	-0.01
T7	-0.04	0.00	-0.03	-0.02	-0.02	-0.02
T8	-0.01	0.01	-0.02	-0.02	-0.03	-0.01
C9	-0.02	0.01	-0.02	-0.02	-0.03	0.00
G10	-0.03	0.00	-0.01	-0.02	-0.02	0.00
C11	-0.03	-0.01	-0.03	-0.02	0.05	-0.02
G12	-0.02	0.00	0.02	-0.03	-0.02	-0.02

5.3 Dickerson Build-up Curves

In this section, the build-up curves for the non-exchangeable nuclei at 25 °C will be presented for each residue in the two dodecamers (F5.3.1-14). Initially, all samples in this thesis were taken at 25 °C, but it was determined that the spectra lacked many of the distinct aromatic-1' peaks. As a result, all experiments were redone at 10 °C except these, owing to the many weeks required to perform them at multiple mixing times. While the build-up curves lack some important peaks, they nevertheless yield some nuclear distances of sufficient quality to include as modeling restraints.

Error bars represent the 5% deviation of intensity expected for the instrument. Excluded from the figures are ambiguous and heavily overlapped NOEs, build-up curves at less than a 0.8 R^2, and crosspeaks between residues, which tend to be highly variable. Only those with greater than 0.9 R^2 will be included as NMR restraints for the modeling. Additionally, a restriction of 15% error will be imposed, which was the tolerance accepted in the Tjandra et al. paper[40] between its model (1DUF) and the NMR distances cited therein. Thus, the screening for acceptable NMR restraints involved a two prong approach: having a linear build-up curve and being comparable to known distances. T5.3.1-12 list the internuclear distances for inclusion to the model.

Not all NOEs in the same DNA sequence have the same expected variability because each proton experiences a different environment. Also, while the sample conditions for the sequences were identical, many more NOEs for the 8-oxo-dA5 DDD were subject to intensity variation resulting in a lower than acceptable R^2, even if they presented as distinct, assignable peaks. Further, some peaks (particularly H6-H5 NOEs on the cytosines) were prone to splitting into tetrads with zero or negative intensity at the center, shown in F5.4.11. These phenomena were not observed at 10 °C.

The build-up curves present a large amount of data, but several key features keep the purpose in perspective. As can be expected, shorter distances (as referenced to the Tjandra et al. paper[40]) generated greater NOE magnitudes for most nuclei. This trend was only violated for some short distances at fast mixing times. It is possible that spin diffusion for these nuclei involves multiple and variable relaxation pathways, which can confer a large positive or negative component to intensity.

For short nuclear distances (less than 0.8Å), the build-up curve loses correlation between intensity and distance. This is a direct result of spin diffusion. Since the magnitude of the NOE is proportional to $1/r^6$ and these distances are so short, the effects of spin diffusion are multiplied considerably as mixing time increases. On the other hand, signal to noise plays a significant role in reducing efficacy at shorter mixing times. This dilemma mostly makes short distances unfit for inclusion.

For long distances, there are fewer nuclei to be involved in the spin-lattice relaxation pathway. Spin diffusion plays a very minor role, if any. However, the intensity of the signal is much lower and some distances can never be resolved, even with a stronger magnetic field-strength NMR because the NOE effect sharply diminishes with distance by $1/r^6$. Crosspeaks (peaks between nuclei of different residues) tend to represent long distances and while they are reproducible, their intensities are more subject to variation than intrastrand peaks, making them unfit for inclusion in the build-up curves.

Depicting the errors caused by spin diffusion is complicated because the R-squared is also affected by the standard 5% error of the instrument. While longer distances may suffer little from spin diffusion error, the R-squared can indicate poor linearity. Spin diffusion correlates best with R-squared at short distances where this error

is very small. As a result, the most accurate NOEs will appear in a goldilocks zone that balances long and short distances.

A further confounding factor is that the slopes of the build-up curves should ideally be the same, but all ^{1}H nuclei suffer spin diffusion to some extent, even if they are sampled in the linear range. Additionally, not all build-up curves have the same number of nuclear distances nor do they plot all non-exchangeable nuclei in the residue. Even if the NOE can be assigned and distinguished, peak broadening and overlap, especially in the guanosine residues, can make it impossible to clarify signal intensity under the chosen experimental parameters. For these reasons, the distances are compared to the best known literature values (T5.3.1-12).

It is important to note that for dsDNA, the natural conformations of the two strands are not identical. Indeed, the original paper by Dickerson et al. described the sequence as being remarkably asymmetrical.[3] As such, all NOEs are assumed to represent two distances, one for each strand, and that the distance ascertained is the average. The more asymmetrical the two strands, the greater the deviation. This deviation amounts to another unknowable error for modeling by ^{1}H NMR because the NOEs from each strand cannot be distinguished. This was a necessary sacrifice however, because nuclear DNA is double-stranded when not undergoing replication and transcription.

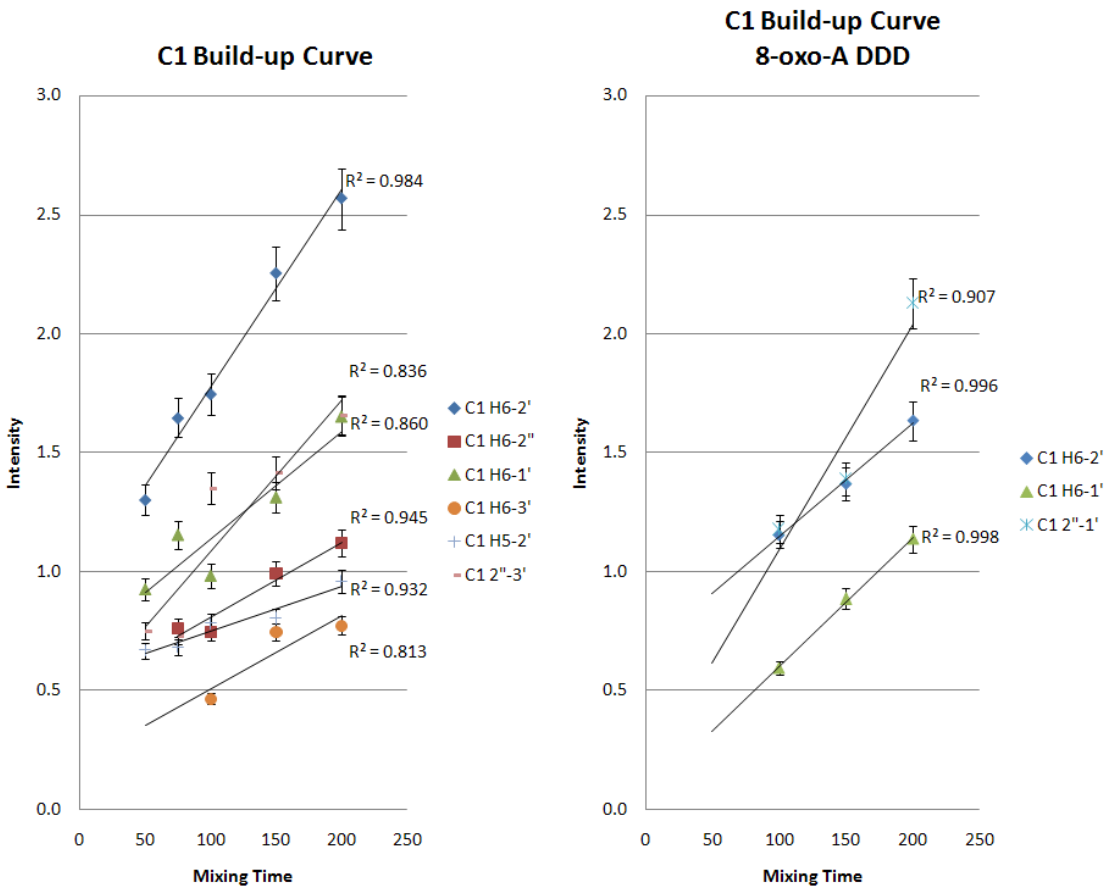


Figure 5.3.1. The build-up curves for C1 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.1. The interatomic distances for C1, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% Difference
C1 (DDD)	H6	2'	2.39	2.40	0.50%
C1 (DDD)	H6	2"	2.74	3.94	30.34%
C1 (DDD)	H6	1'	2.57	3.73	31.02%
C1 (DDD)	H6	3'	2.92	3.66	20.30%
C1 (DDD)	H5	2'	2.81	4.36	35.47%
C1 (DDD)	2"	3'	2.57	2.83	9.10%
C1 (8-oxo-A)	H6	2'	2.40	2.40	0.15%
C1 (8-oxo-A)	H6	1'	2.55	3.73	31.48%
C1 (8-oxo-A)	2"	1'	2.30	2.41	4.65%

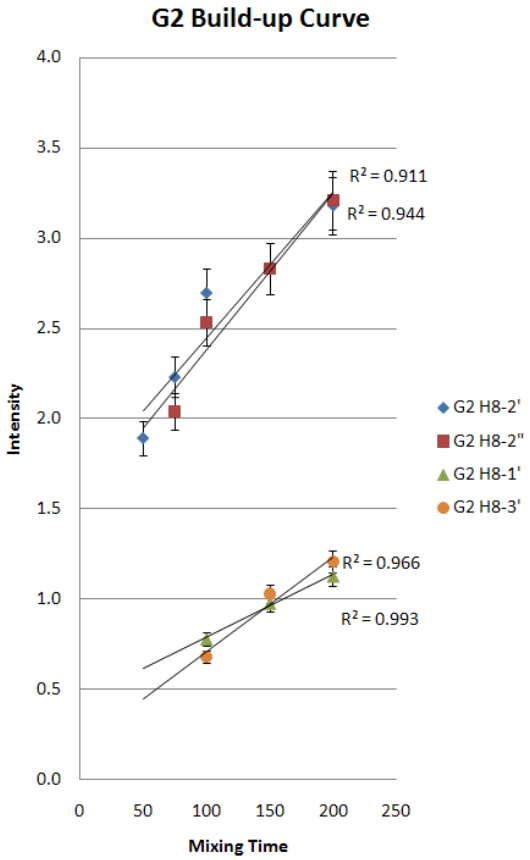


Figure 5.3.2. The build-up curves for G2 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.2. The interatomic distances for G2, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% Difference
G2 (DDD)	H8	2'	2.17	2.30	5.74%
G2 (DDD)	H8	2"	2.17	3.78	42.66%
G2 (DDD)	H8	1'	2.58	3.91	33.88%
G2 (DDD)	H8	3'	2.55	4.19	39.12%

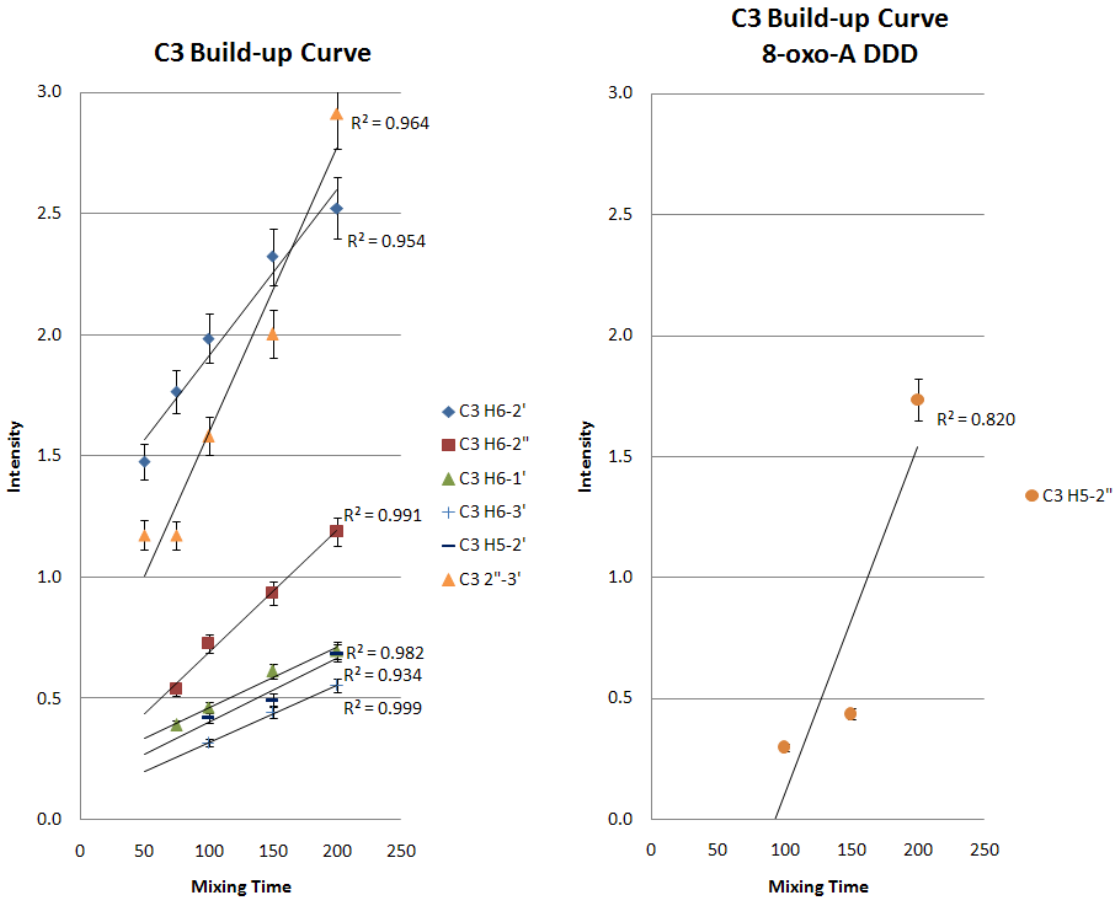


Figure 5.3.3. The build-up curves for C3 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.3. The interatomic distances for C3, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% Difference
C3 (DDD)	H6	2'	2.34	2.38	1.36%
C3 (DDD)	H6	2"	2.66	3.89	31.68%
C3 (DDD)	H6	1'	2.90	3.72	22.05%
C3 (DDD)	H6	3'	3.02	3.70	18.59%
C3 (DDD)	H5	2'	2.91	4.41	34.00%
C3 (DDD)	2"	3'	2.29	2.81	18.68%
C3 (8-oxo-A)	H5	2"	2.30	5.78	60.23%

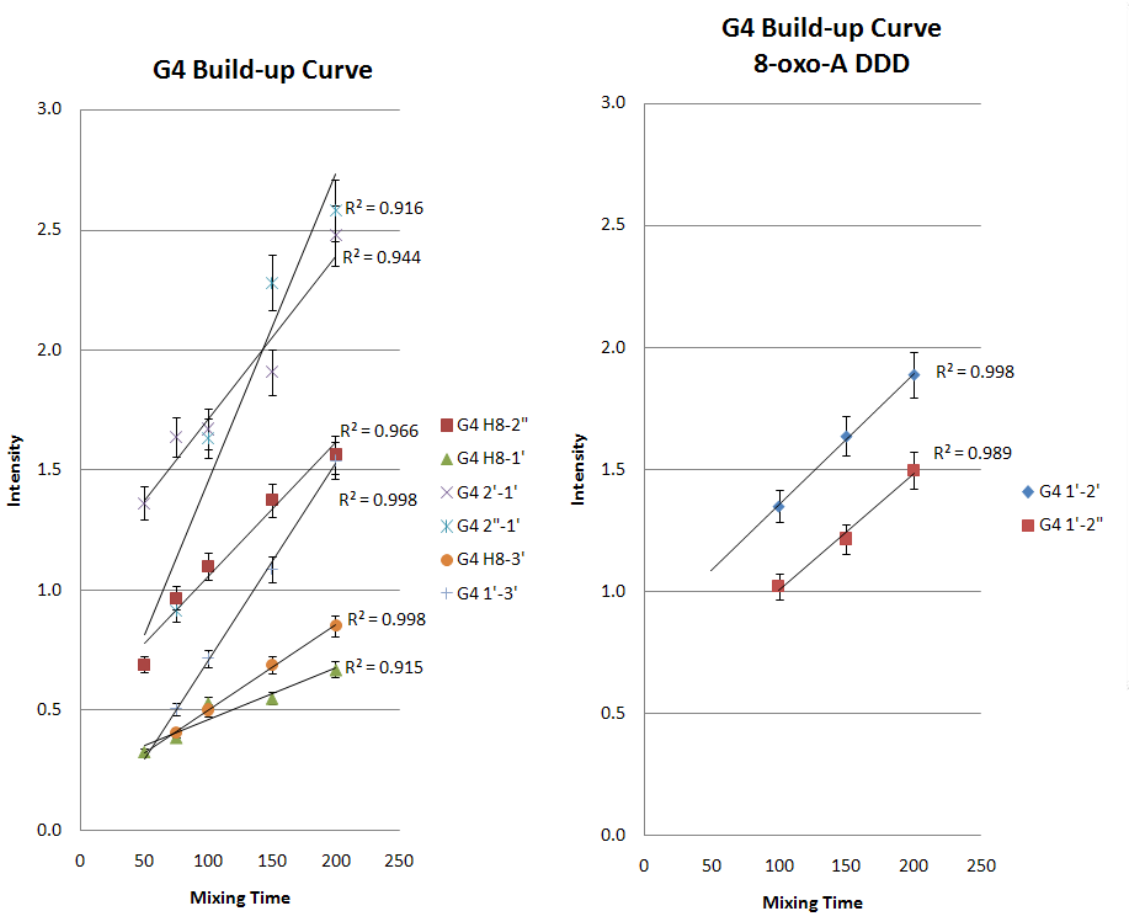


Figure 5.3.4. The build-up curves for G4 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.4. The interatomic distances for G4, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% Difference
G4 (DDD)	H8	2"	2.50	3.64	31.36%
G4 (DDD)	H8	1'	2.88	3.87	25.58%
G4 (DDD)	H8	3'	2.77	4.27	35.14%
G4 (DDD)	2'	1'	2.32	3.03	23.65%
G4 (DDD)	2"	1'	2.30	2.38	3.20%
G4 (DDD)	1'	3'	2.51	3.78	33.75%
G4 (8-oxo-A)	2'	1'	2.21	3.03	27.06%
G4 (8-oxo-A)	2"	1'	2.30	2.38	3.20%

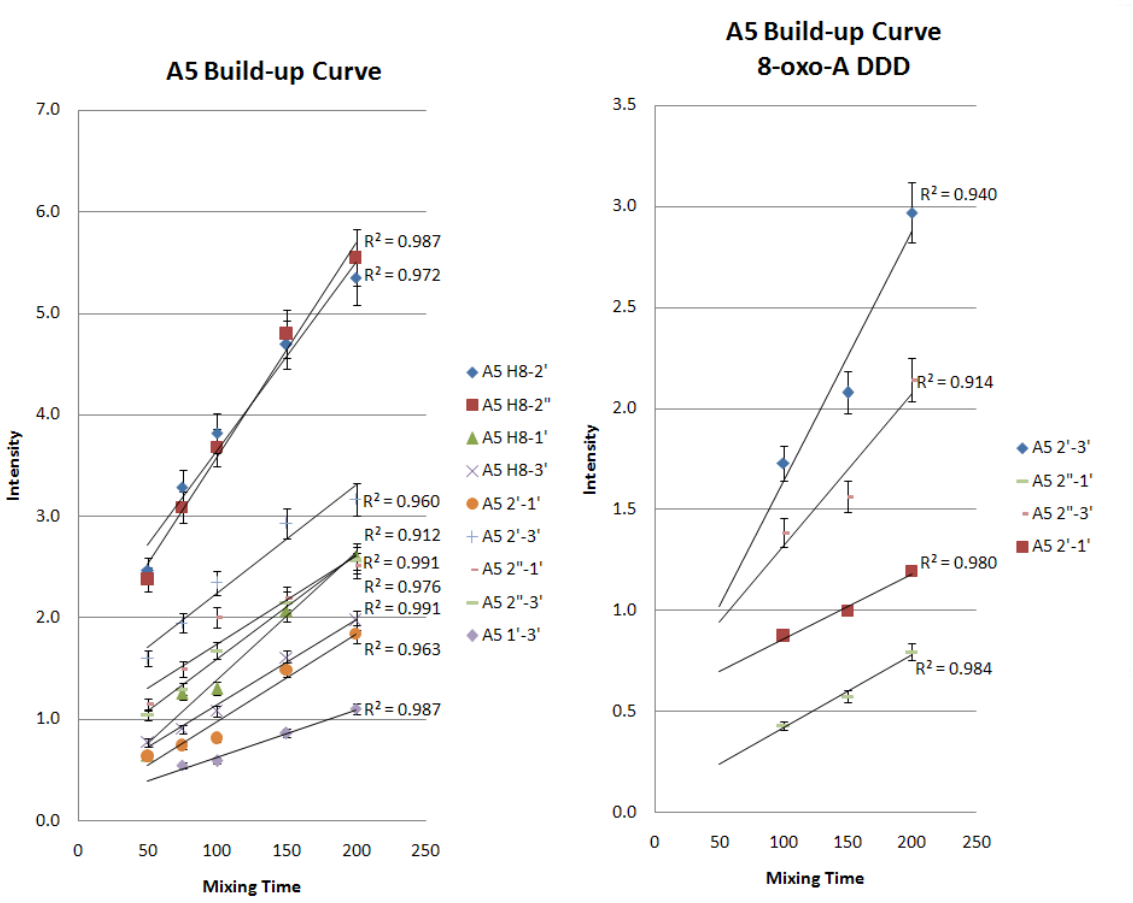


Figure 5.3.5. The build-up curves for A5 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.5. The interatomic distances for A5, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% Difference
A5 (DDD)	H8	2'	2.03	2.47	18.06%
A5 (DDD)	H8	2"	2.02	3.98	49.33%
A5 (DDD)	H8	1'	2.29	3.91	41.54%
A5 (DDD)	H8	3'	2.39	4.17	42.64%
A5 (DDD)	2'	1'	2.42	3.02	19.73%
A5 (DDD)	2'	3'	2.21	2.31	4.37%
A5 (DDD)	2"	1'	2.30	2.34	1.69%
A5 (DDD)	2"	3'	2.29	2.73	16.20%
A5 (DDD)	1'	3'	2.64	3.84	31.36%
A5 (8-oxo-A)	2'	1'	2.15	3.02	28.81%
A5 (8-oxo-A)	2'	3'	1.85	2.31	20.20%
A5 (8-oxo-A)	2"	1'	2.75	2.34	17.37%
A5 (8-oxo-A)	2"	3'	1.95	2.73	28.68%

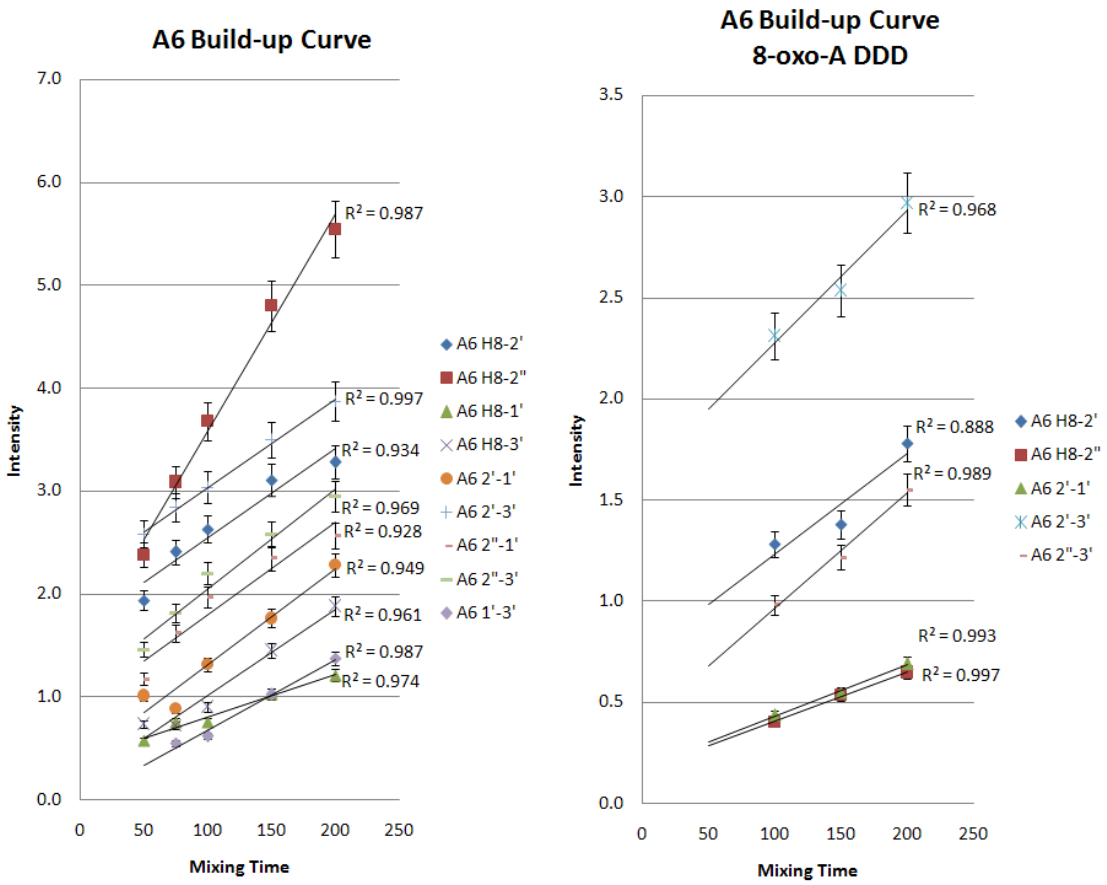


Figure 5.3.6. The build-up curves for A6 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.6. The interatomic distances for A6, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% Difference
A6 (DDD)	H8	2'	2.21	2.36	6.39%
A6 (DDD)	H8	2"	2.02	3.85	47.52%
A6 (DDD)	H8	1'	2.61	3.91	33.37%
A6 (DDD)	H8	3'	2.42	4.18	42.03%
A6 (DDD)	2'	1'	2.35	3.01	22.15%
A6 (DDD)	2'	3'	2.15	2.32	7.43%
A6 (DDD)	2"	1'	2.3	2.33	1.08%
A6 (DDD)	2"	3'	2.25	2.75	18.39%
A6 (DDD)	1'	3'	2.55	3.84	33.59%
A6 (8-oxo-A)	H8	2'	1.78	2.36	10.37%
A6 (8-oxo-A)	H8	2"	2.5	3.85	35.09%
A6 (8-oxo-A)	2'	1'	2.48	3.01	17.84%
A6 (8-oxo-A)	2'	3'	1.94	2.32	16.33%
A6 (8-oxo-A)	2"	3'	2.16	2.75	21.47%

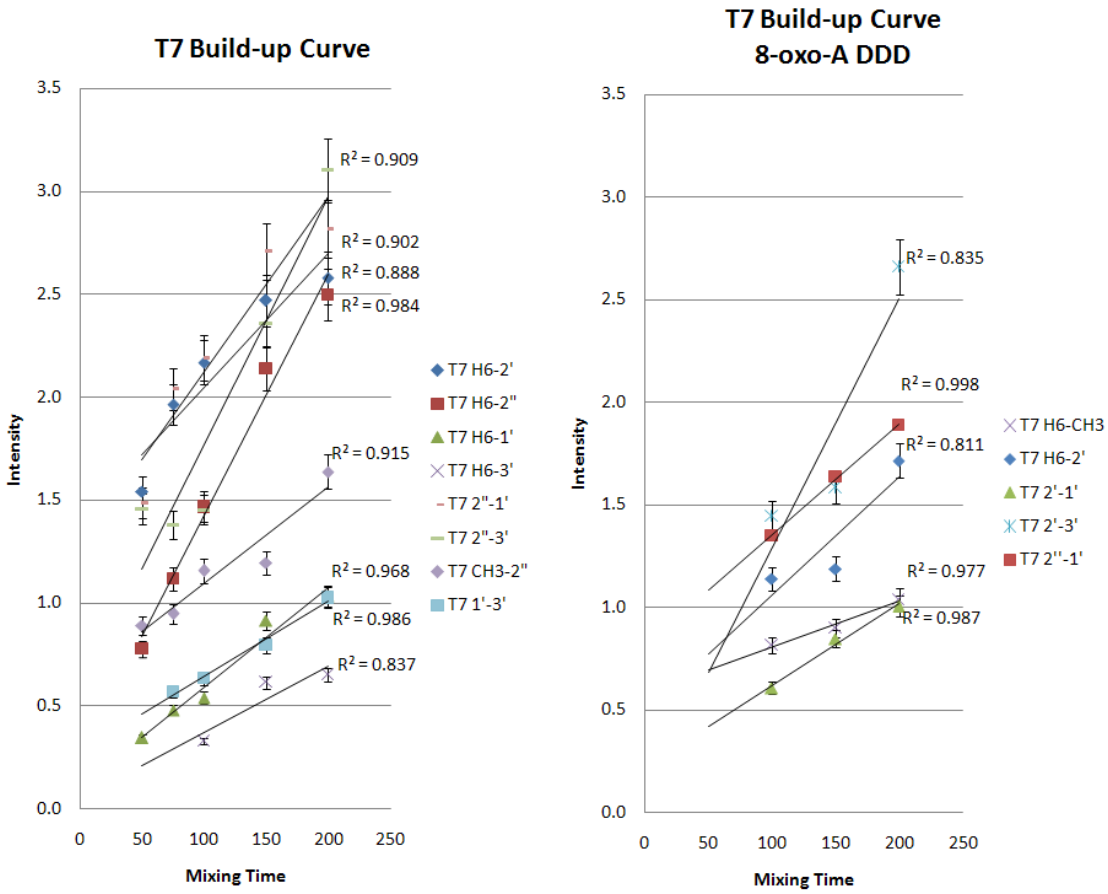


Figure 5.3.7. The build-up curves for T7 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.7. The interatomic distances for T7, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% difference
T7 (DDD)	H6	2'	2.33	2.38	1.80%
T7 (DDD)	H6	2"	2.35	3.87	39.38%
T7 (DDD)	H6	1'	2.72	3.73	27.15%
T7 (DDD)	H6	3'	2.94	3.75	21.65%
T7 (DDD)	Me	2"	2.52	4.38	42.52%
T7 (DDD)	2"	1'	2.3	2.32	1.05%
T7 (DDD)	2"	3'	2.26	2.82	19.81%
T7 (DDD)	1'	3'	2.72	3.84	29.16%
T7 (8-oxo-A)	H6	2'	2.34	2.38	1.69%
T7 (8-oxo-A)	H6	Me	2.29	2.65	13.60%
T7 (8-oxo-A)	2'	1'	2.55	3.01	15.23%
T7 (8-oxo-A)	2"	1'	2.3	2.32	1.05%
T7 (8-oxo-A)	2'	3'	2.17	2.28	4.79%

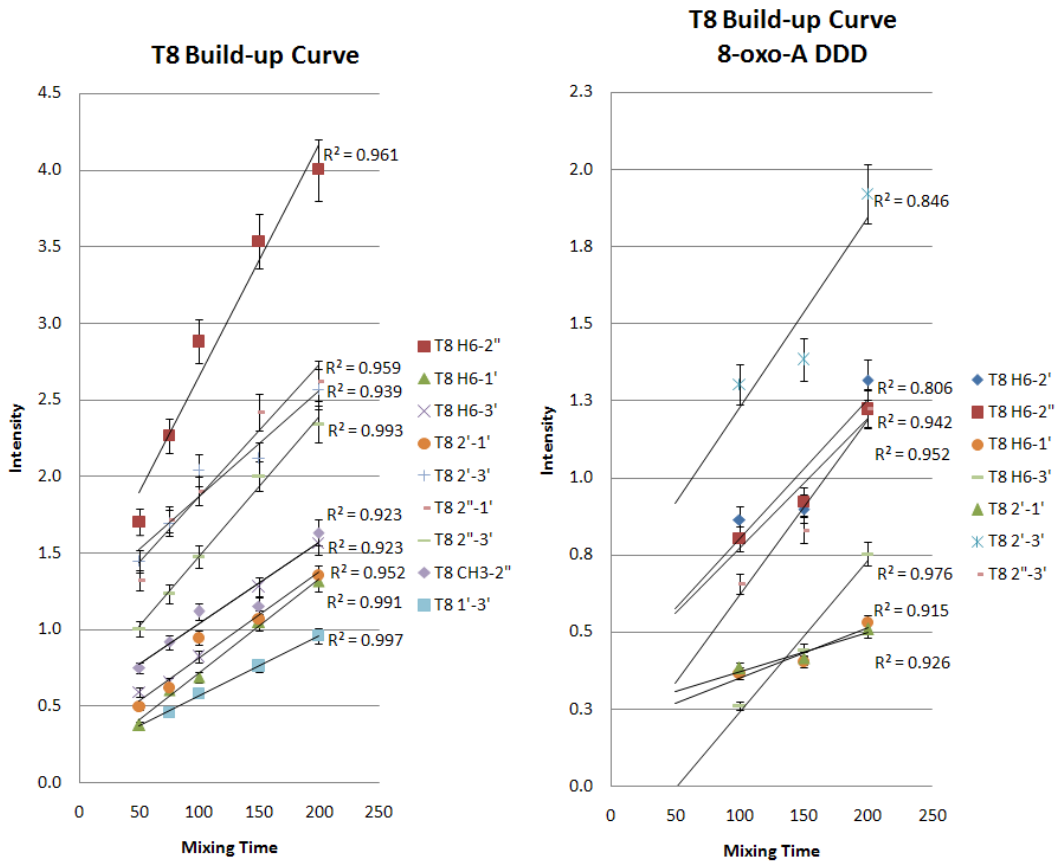


Figure 5.3.8. The build-up curves for T8 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.8. The interatomic distances for T8, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% difference
T8 (DDD)	H6	2"	2.14	3.59	40.28%
T8 (DDD)	H6	1'	2.58	3.71	30.36%
T8 (DDD)	H6	3'	2.51	3.72	32.53%
T8 (DDD)	Me	2"	2.49	4.2	40.75%
T8 (DDD)	2'	1'	2.57	3.02	14.98%
T8 (DDD)	2'	3'	2.31	2.32	0.32%
T8 (DDD)	2"	1'	2.3	2.33	1.34%
T8 (DDD)	2"	3'	2.34	2.79	15.94%
T8 (DDD)	1'	3'	2.72	3.85	29.27%
T8 (8-oxo-A)	H6	2"	2.28	3.59	36.40%
T8 (8-oxo-A)	H6	2'	2.25	2.04	10.66%
T8 (8-oxo-A)	H6	1'	2.62	3.71	29.19%
T8 (8-oxo-A)	H6	3'	2.47	3.72	33.42%
T8 (8-oxo-A)	2'	1'	2.64	3.02	12.53%
T8 (8-oxo-A)	2'	3'	2.12	2.32	8.55%
T8 (8-oxo-A)	2"	3'	2.28	2.79	18.15%

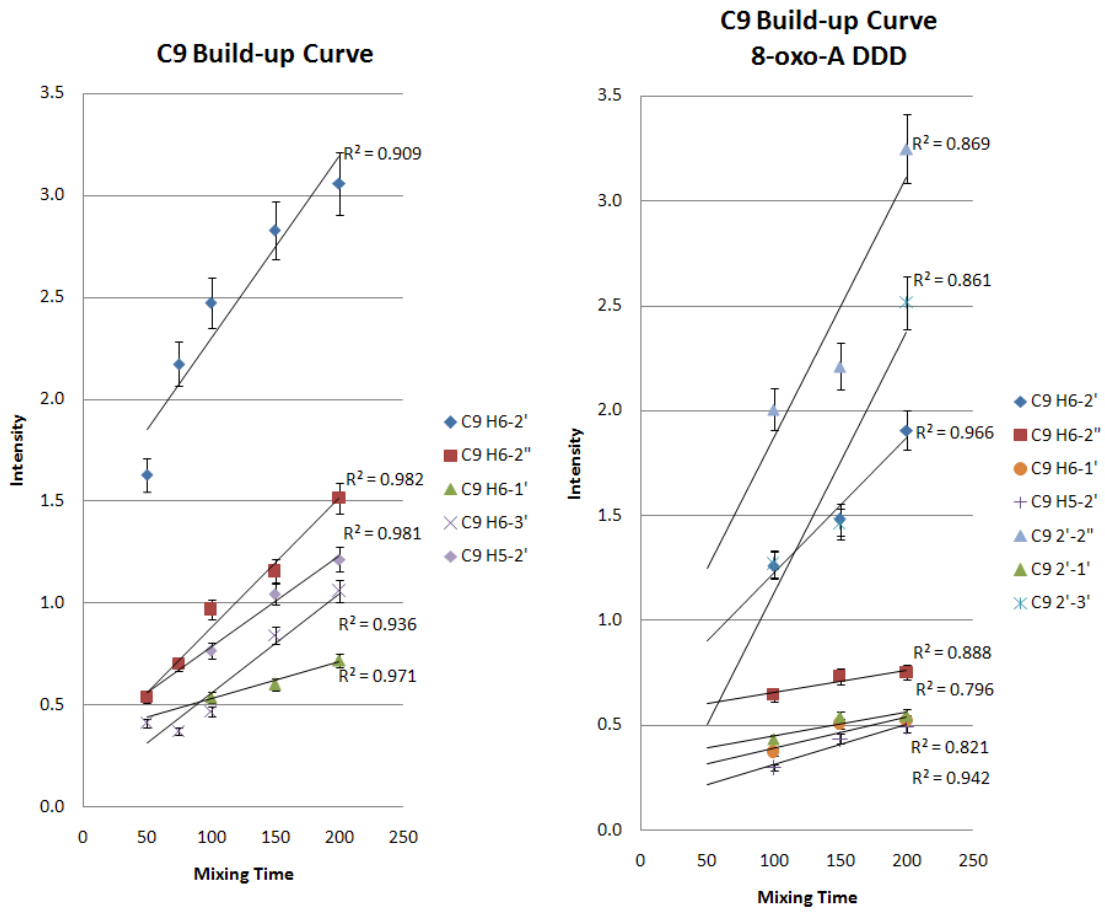


Figure 5.3.9. The build-up curves for C9 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.9. The interatomic distances for C9, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% difference
C9 (DDD)	H6	2'	2.25	2.4	6.17%
C9 (DDD)	H6	2"	2.53	3.88	34.84%
C9 (DDD)	H6	1'	2.87	3.73	23.20%
C9 (DDD)	H6	3'	2.69	3.76	28.57%
C9 (DDD)	H5	2'	2.63	4.46	41.20%
C9 (8-oxo-A)	H6	2'	2.17	2.40	9.33%
C9 (8-oxo-A)	Base	2"	2.54	3.88	34.61%
C9 (8-oxo-A)	Base	1'	2.70	3.73	27.66%
C9 (8-oxo-A)	H5	2'	2.73	4.46	38.91%
C9 (8-oxo-A)	2'	2"	1.99	1.79	11.28%
C9 (8-oxo-A)	2'	3'	2.08	2.30	9.90%
C9 (8-oxo-A)	2'	1'	2.68	3.02	11.35%

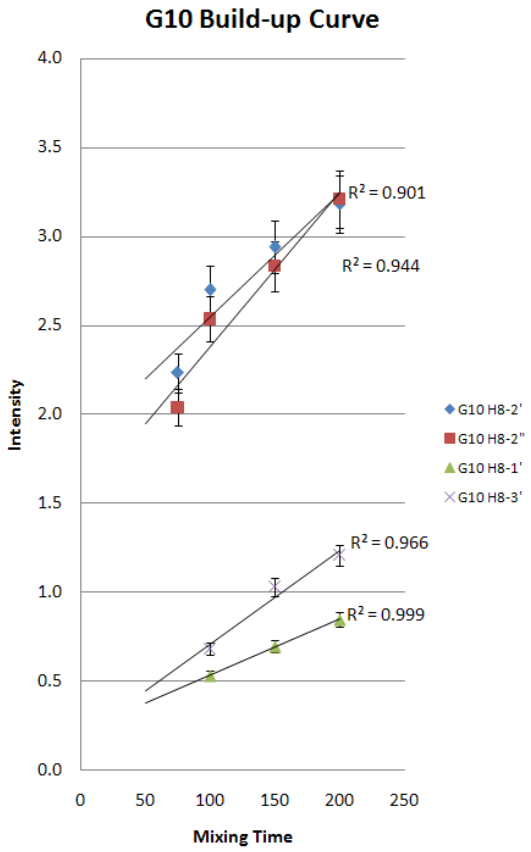


Figure 5.3.10. The build-up curves for G10 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.10. The interatomic distances for G10, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% difference
G10 (DDD)	H8	2'	2.30	2.31	0.40%
G10 (DDD)	H8	2"	2.30	3.83	39.93%
G10 (DDD)	H8	1'	2.87	3.90	26.34%
G10 (DDD)	H8	3'	2.71	4.10	34.06%

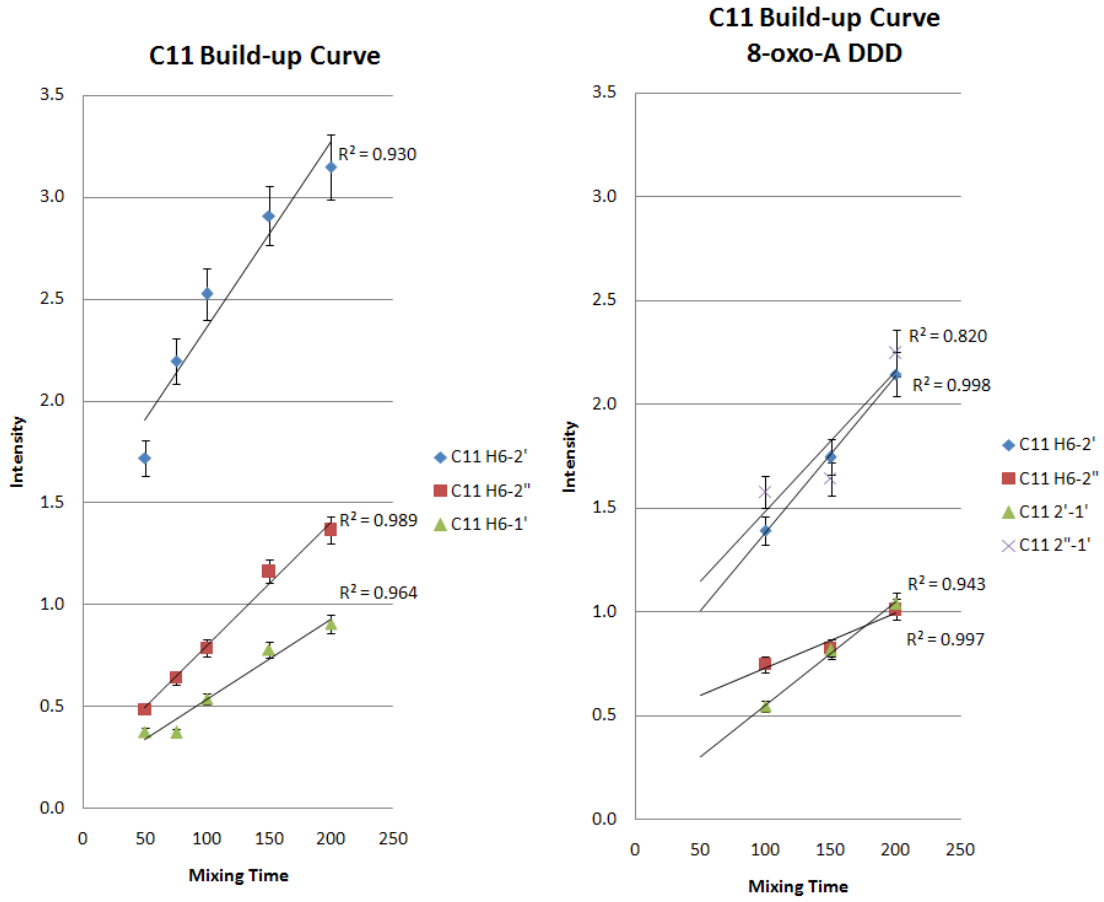


Figure 5.3.11.The build-up curves for C11 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.11. The interatomic distances for C11, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% difference
C11 (DDD)	H6	2'	2.29	2.49	7.88%
C11 (DDD)	H6	2"	2.64	3.98	33.73%
C11 (DDD)	H6	1'	2.82	3.73	24.39%
C11 (8-oxo-A)	H6	2'	2.32	2.49	6.85%
C11 (8-oxo-A)	H6	2"	2.63	3.98	33.92%
C11 (8-oxo-A)	H6	1'	2.93	3.73	21.51%
C11 (8-oxo-A)	2"	1'	2.30	2.36	2.69%

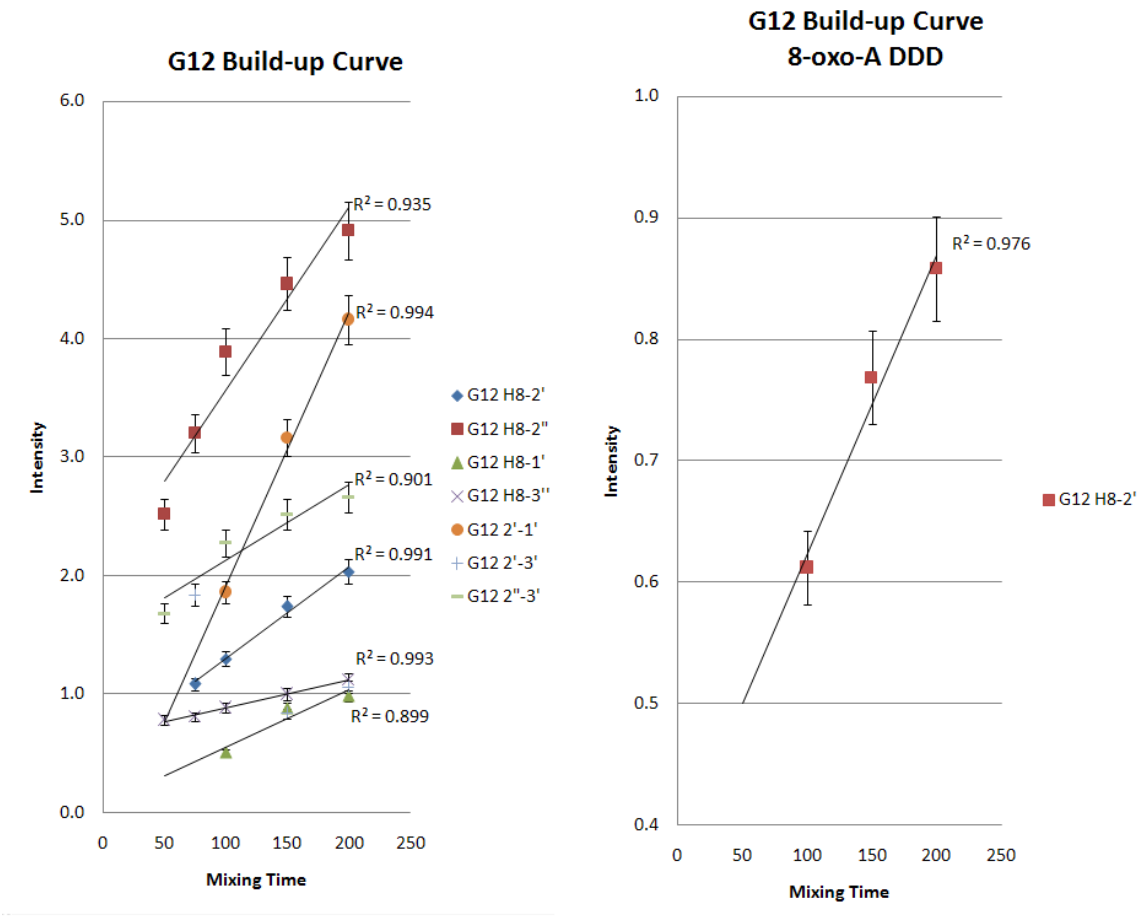


Figure 5.3.12. The build-up curves for G12 assignable NOEs. Intensity correlates with internuclear distance inasmuch as it is linear with mixing time. Linearity can be diminished by noise and spin diffusion.

Table 5.3.12. The interatomic distances for G12, with those excluded from modeling (grey). Viable DDD distances were no greater than 15% of 1DUF. For the 8-oxo-dA5 sequence, only distances with acceptable corollaries in the DDD were used.

Residue	Proton 1	Proton 2	Angstroms	Literature	% difference
G 12 (DDD)	H8	2'	2.26	4.01	43.63%
G 12 (DDD)	H8	1'	2.55	3.91	34.72%
G 12 (DDD)	H8	3'	2.50	3.88	35.59%
G 12 (DDD)	2"	1'	2.30	3.01	23.59%
G 12 (DDD)	2"	3'	2.16	2.27	4.78%
G 12 (8-oxo-A)	H8	2'	2.49	4.01	37.82%

5.4 NOESY Spectra

This section (F5.4.1-11) displays the 10 °C NOESY spectra at 200ms mixing time for reference, as a visual supplement to T5.3.1-12. While these generated the assignments, they are important for inclusion because they indicate, by way of crosspeaks, which nuclei are involved in NOE relaxation pathways.

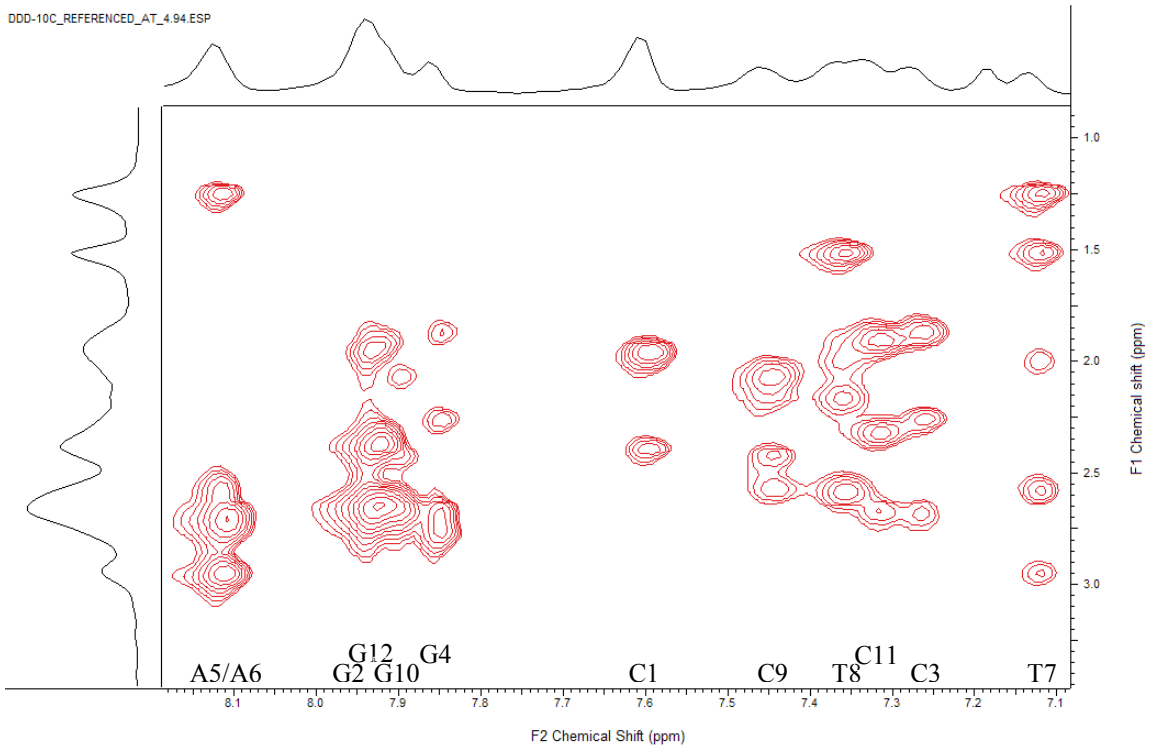


Figure 5.4.1. NOESY spectrum for the Dickerson dodecamer (DDD) at the Base-2′-2″ region.

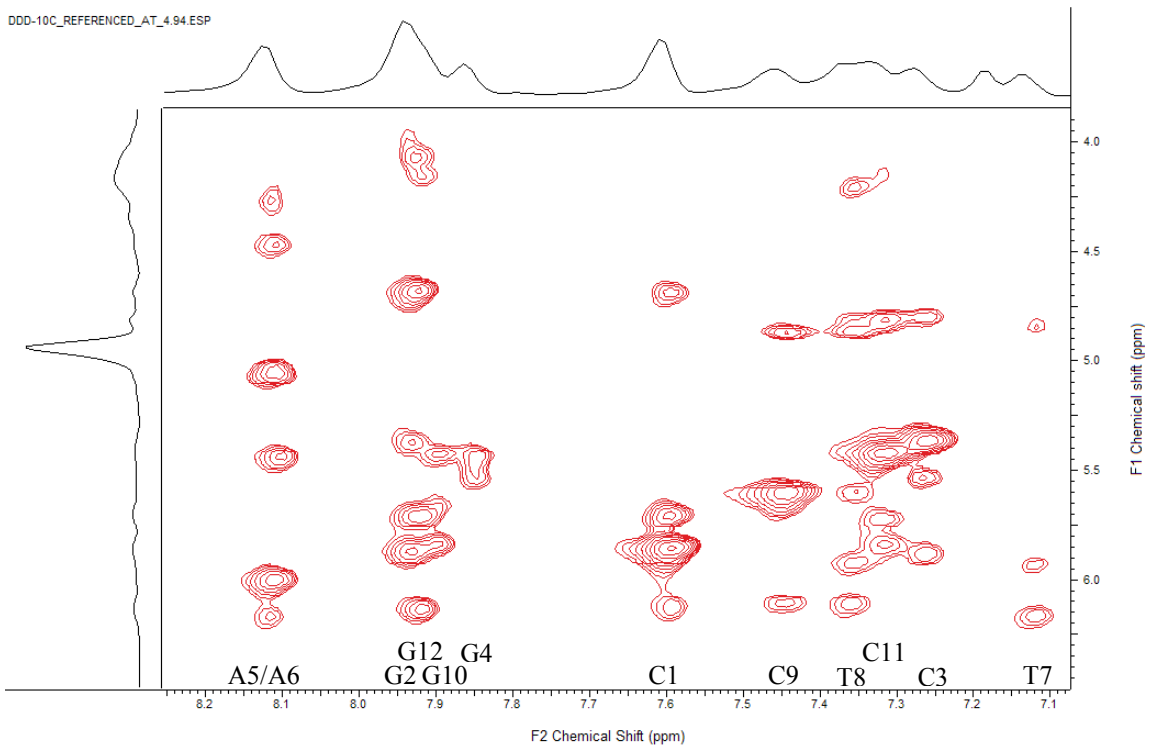


Figure 5.4.2. NOESY spectrum for the DDD at the Base-1′-3′ region.

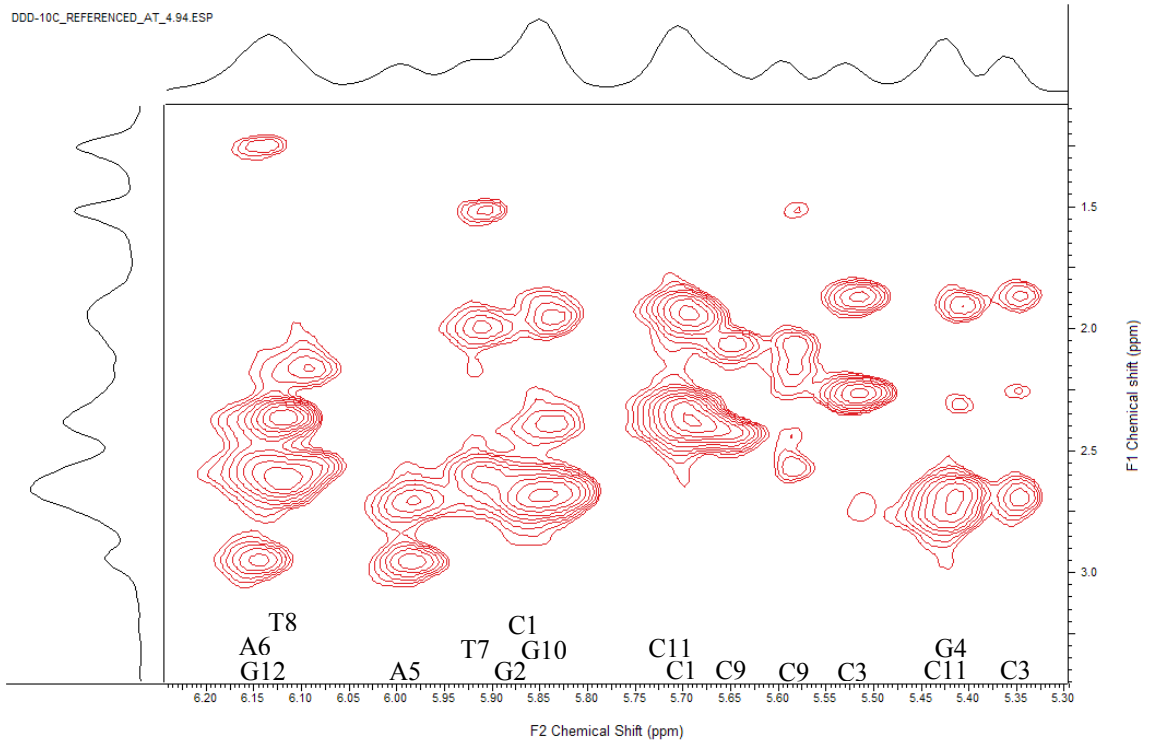


Figure 5.4.3. NOESY spectrum for the DDD at the 1′-2′-2″ region.

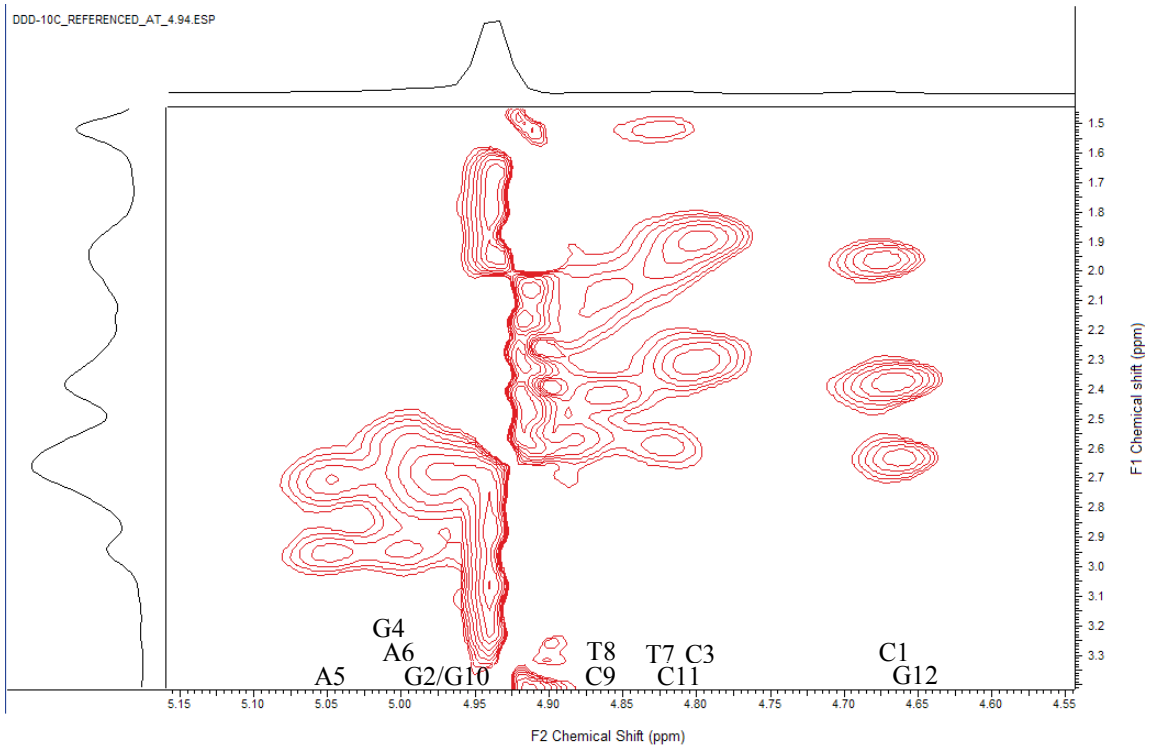


Figure 5.4.4. NOESY spectrum for the DDD at the 3′-2′-2″ region.

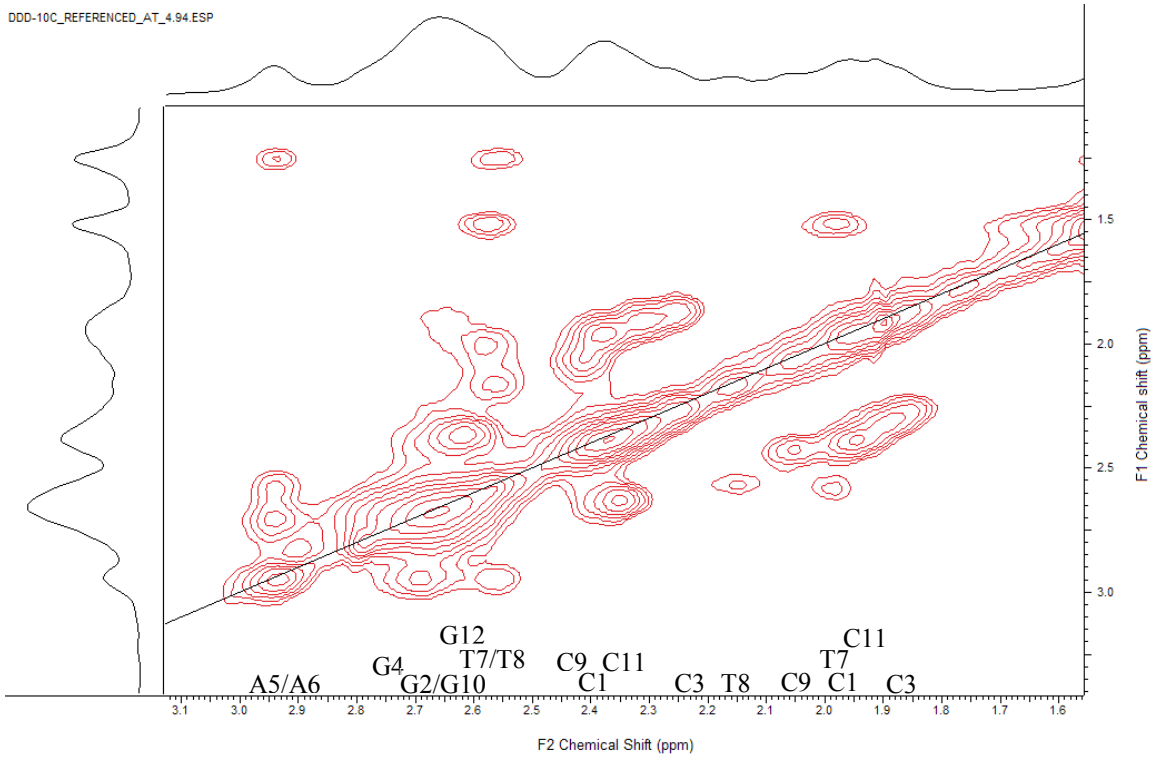


Figure 5.4.5. NOESY spectrum for the DDD at the 2′-2″ region.

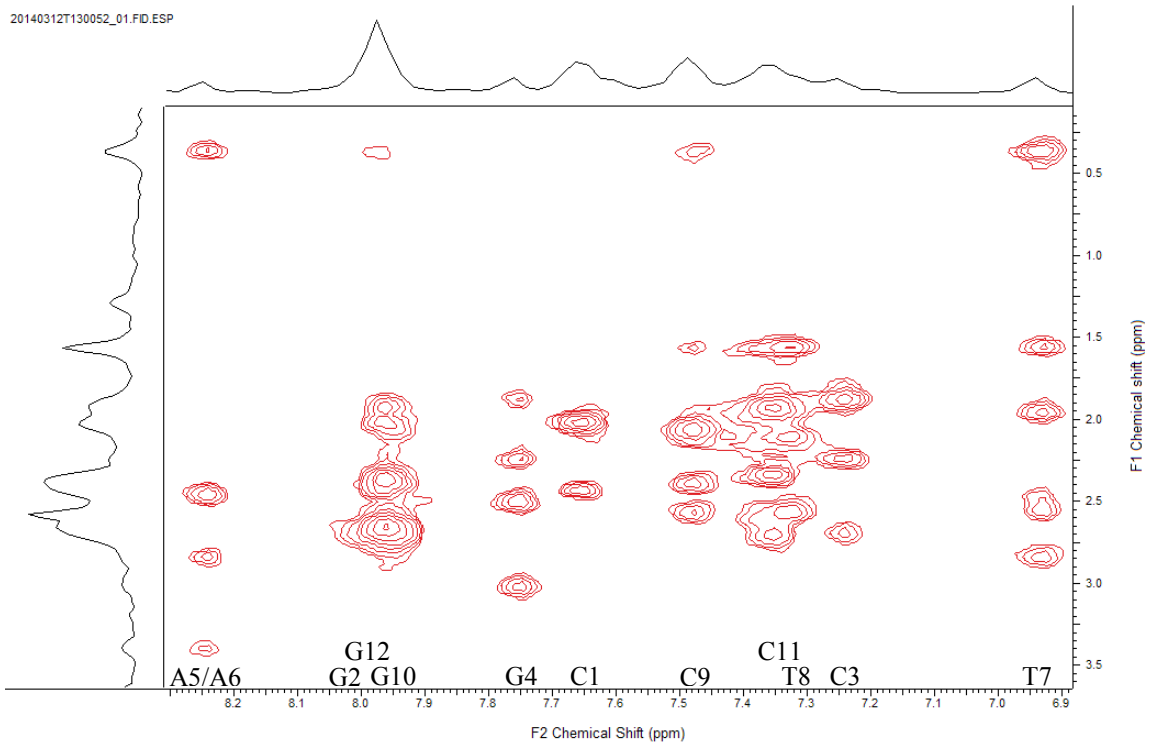


Figure 5.4.6. NOESY spectrum for the 8-oxo-dA5 DDD at the Base-2′-2″ region.

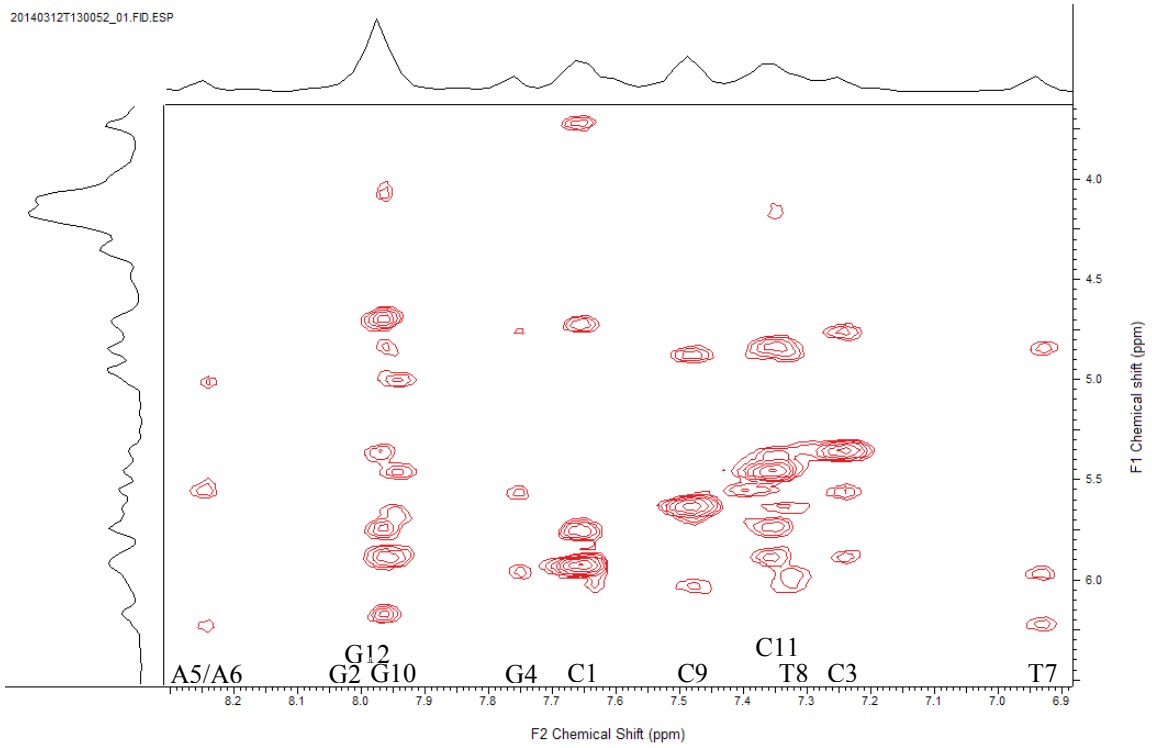


Figure 5.4.7. NOESY spectrum for the 8-oxo-dA5 DDD at the Base-1′-3′ region.

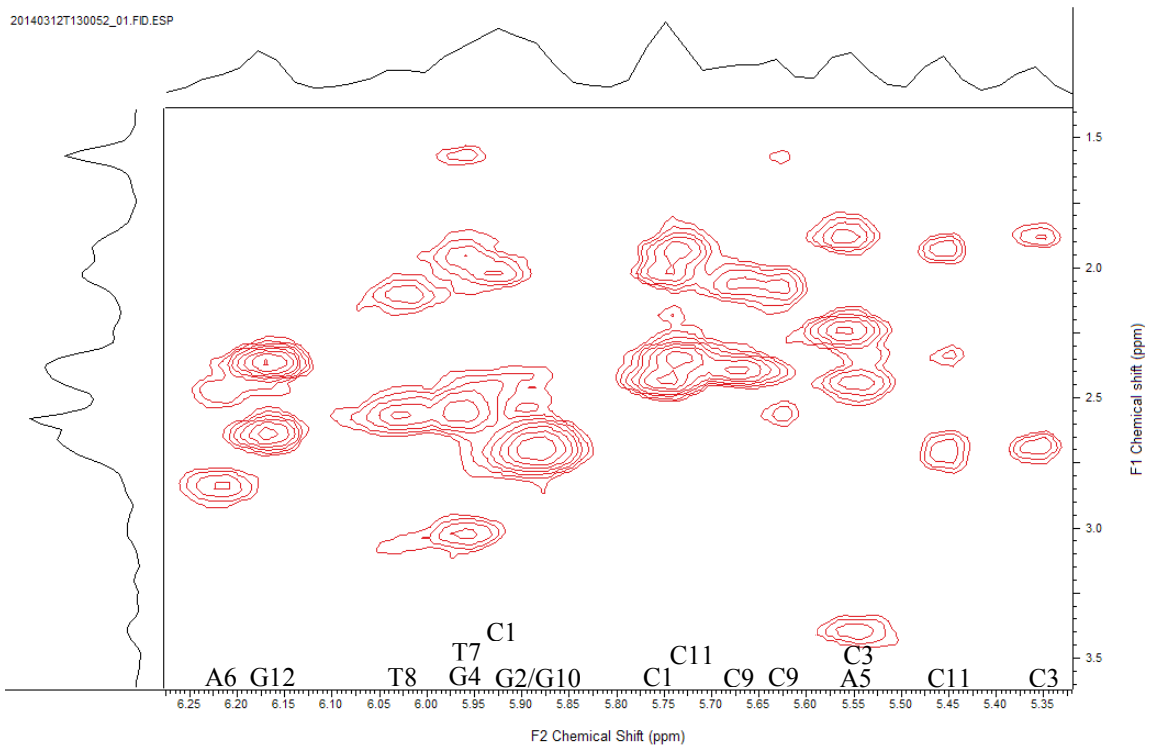


Figure 5.4.8. NOESY spectrum for the 8-oxo-dA5 DDD at the 1′-2′-2″ region.

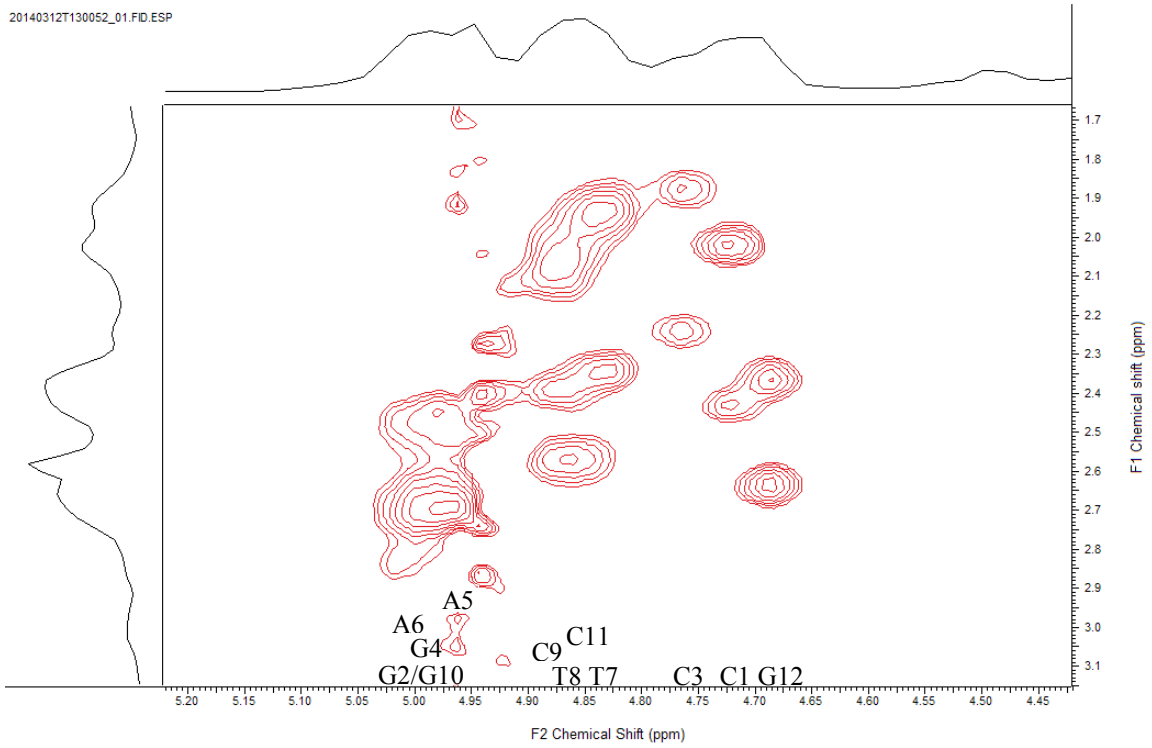


Figure 5.4.9. NOESY spectrum for the 8-oxo-dA5 DDD at the 3′-2′-2″ region.

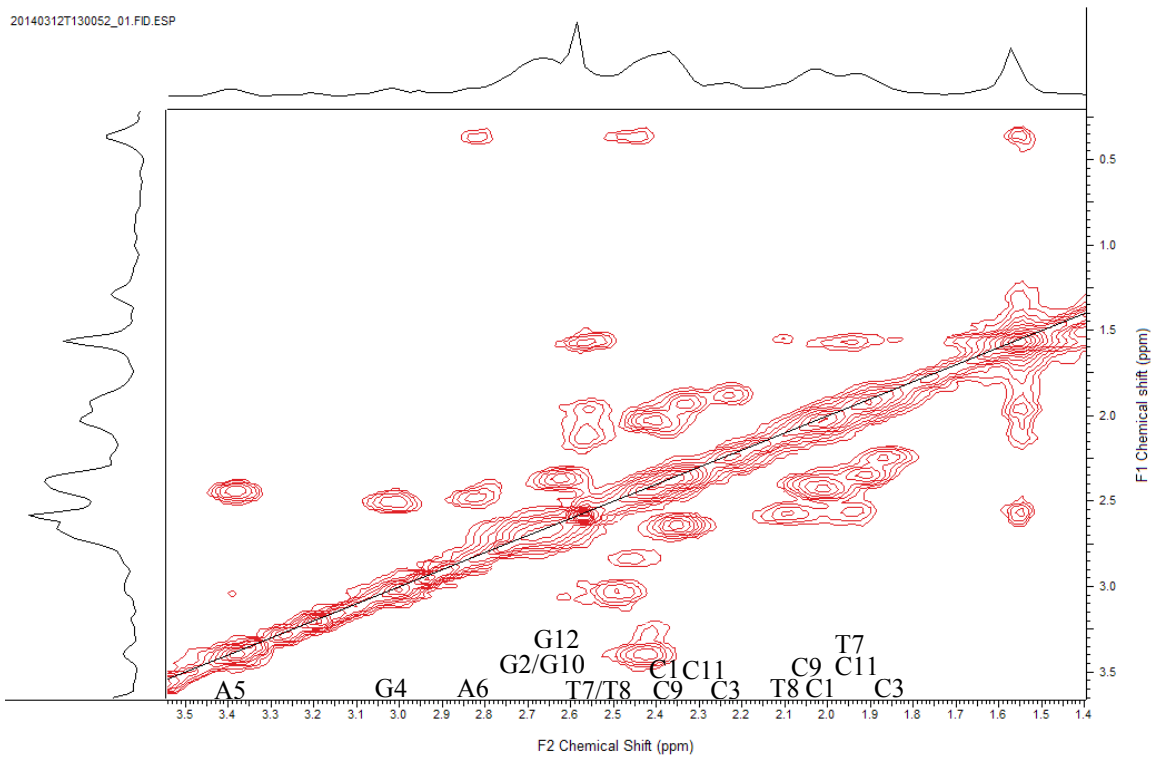


Figure 5.4.10. NOESY spectrum for the 8-oxo-dA5 DDD at the 2′-2″ region.

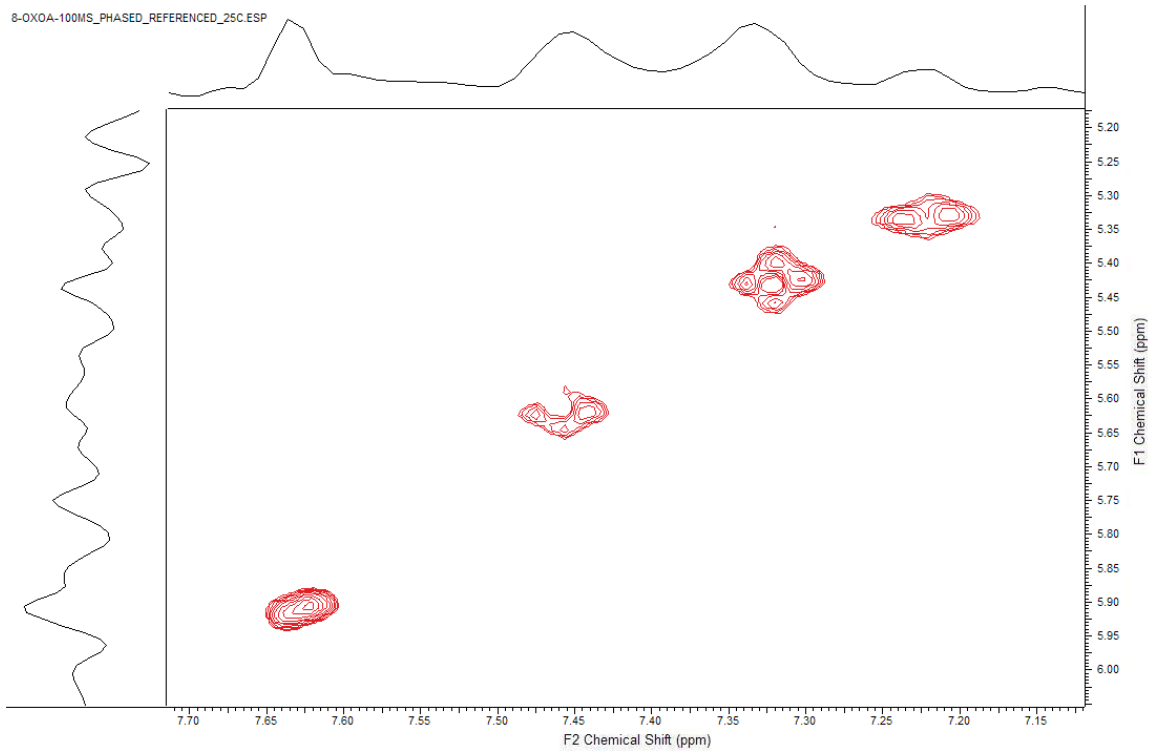
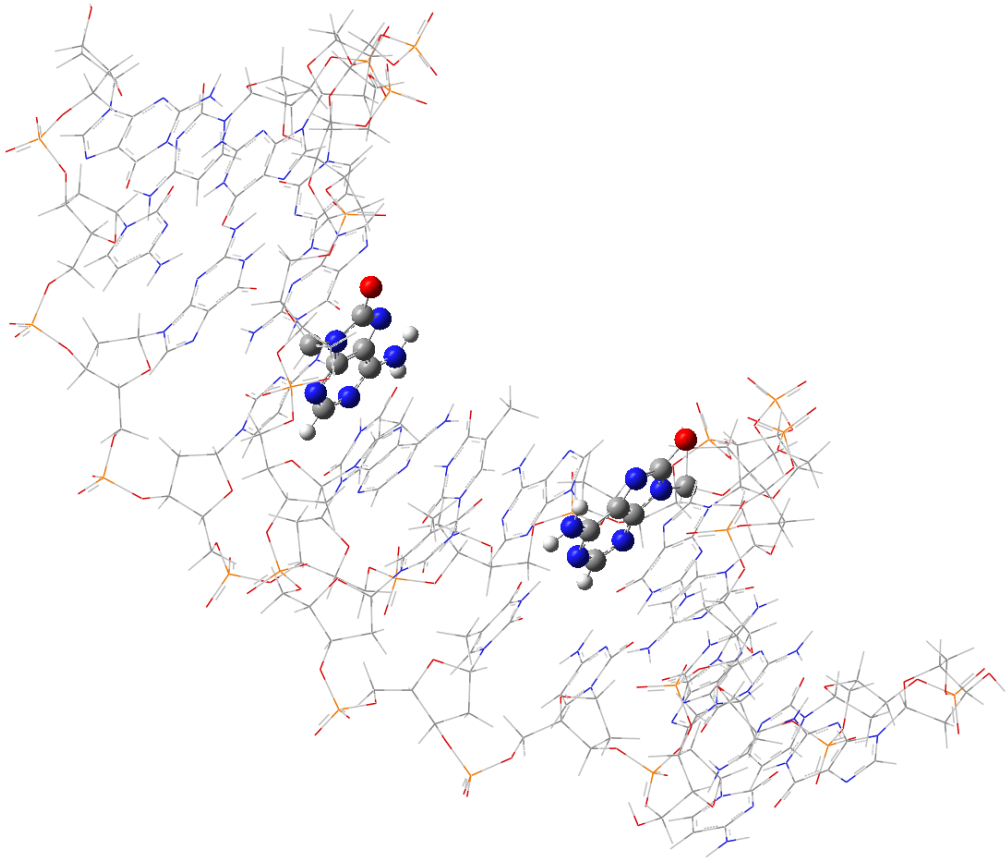


Figure 5.4.11. NOESY spectrum for the 8-oxo-dA5 DDD at the base-1′ region at 25 °C, showing the splitting of cytosine H5-H6 peaks into tetrads and absence of 1′ peaks.

5.5 Molecular Mechanics and DFT Simulation

The molecular dynamics simulation was an attempt to roughly estimate the structures of the Dickerson dodecamers, where the 8-oxo-dA5 lesion site was given DFT calculation privileges. Distance constraints were applied based on those derived from the NOE intensities in the build-up curves. NOEs were selected to participate in the modeling if the R-squared values were greater than 0.90 and deviated from previously described distances no more than 15% (DDD only). Featured in F5.5.1 is a bird's eye view of the 8-oxo-dA5 DDD, with the lesion represented as "ball and stick" over a wireframe before energy minimization. Initial attempts at molecular dynamics (MD) simulation are included in the supplemental disk and listed in App B.



F5.5.1 The 8-oxo-dA5 Dickerson dodecamer as it appears in GaussView 5.0.9 before MD simulation, with the lesion as a ball-and-stick cartoon.

While useful in principle, the molecular dynamics simulation unfortunately encountered errors for ONIOM and DFT attempts in practice. The only model of the 8-oxo-A DDD to produce an energy-minimized structure was with the UFF (included in the supplemental disk), which does not possess restrictions on internuclear distances found by ^{1}H NMR. UFF-minimized structures of DNA oligomers do not represent accurate geometries.[48] Even if the ONIOM and DFT succeeded, the model would still be intractable, considering that only nine ^{1}H internuclear distances would be empirical, out of 758 atoms in the molecule. As such, finding the helical differences for the 8-oxo-dA5 DDD is a matter for future study and only those of the unmodified DDD (1DUF) were explored.

5.6 Helicity Studies

Since damage recognition by repair enzymes is expected to be largely dependent on secondary structure, it might be clarified by comparing the degrees of freedom comprising helicity between unmodified and modified sequences. Graphic depictions of these variables are shown in F5.6.1-2.

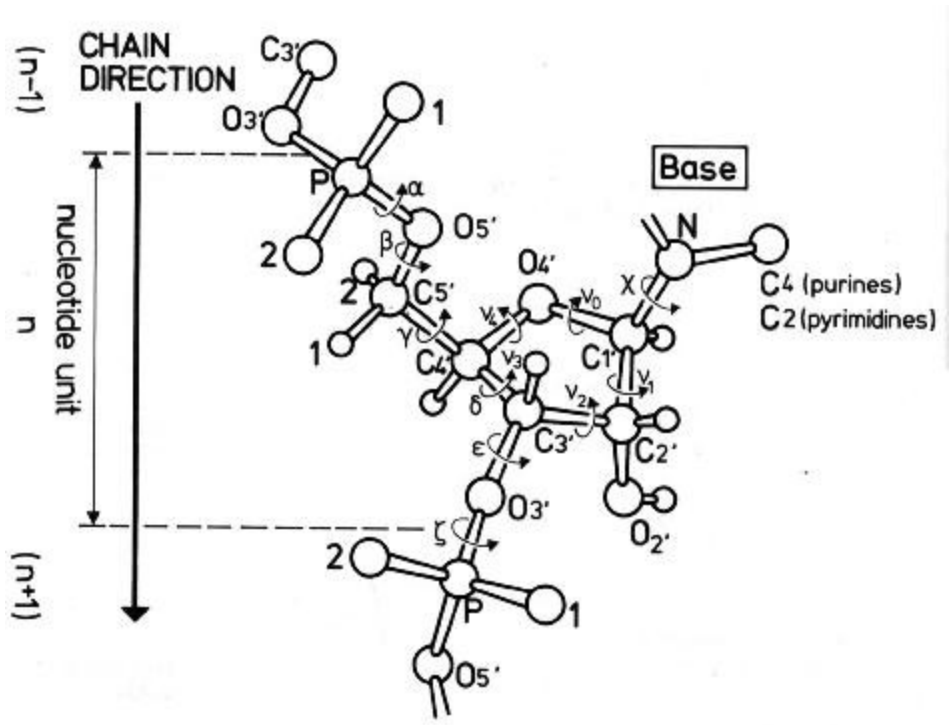


Figure 5.6.1. The 13 degrees of freedom in sugar puckering and torsion. Replicated with permission from the 3DNA tutorial.[51]

The 5 degrees of freedom for the torsion of a 5-membered ring (deoxyribose) are represented by ν_0-ν_4. Since they are constrained in a ring, their conformation is traditionally viewed as an amalgam of these torsions, called pseudorotation, which is given two values: phase angle (P) and amplitude (T_m). Two formulae have been developed for computing pseudorotation. [61,62] By comparison, the phosphate backbone and base, which undergo torsion at α-ζ and χ respectively, do not have ring constraints. Puckering and torsion of the sugar influence and are influenced by the base pairs. The helicity of base pairs can be determined by the magnitude of their 16 degrees of freedom, shown in F5.6.2.

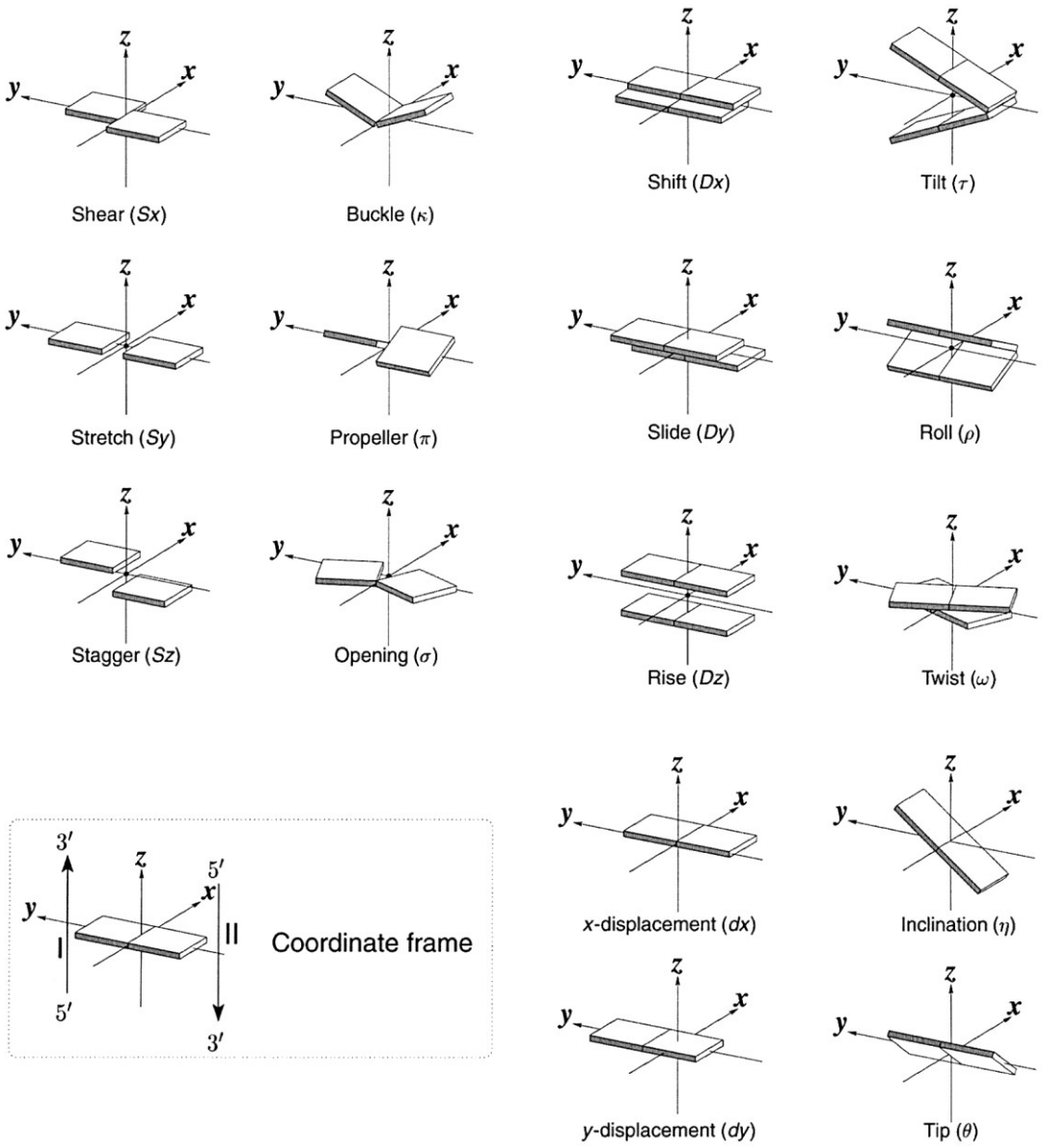


Figure 5.6.2. The 16 degrees of freedom between opposing bases in DNA. Replicated with permission the 3DNA tutorial. [51]

These 29 degrees of freedom comprise helicity, but it can be more descriptive to represent these as derived variables, such as major and minor groove width, since these are the docking sites of transcription factors and enzymes. For simplicity, only the degrees of freedom for the base pairs, major and minor grooves will be noted for the DDD in T5.1.1-2.

Table 5.1.1. Average and standard deviation of helical parameters for the 1DUF Dickerson dodecamer from the RCSB Protein Data Bank,[40] using the 3DNA software package. [51]

	Shear (Å)	Stretch (Å)	Stagger (Å)	Buckle (°)	Propeller (°)	Opening (°)
ave.	0.00	-0.39	-0.28	-0.02	-11.76	0.99
s.d.	0.38	0.16	0.50	3.33	4.49	3.03
	Shift (Å)	Slide (Å)	Rise (Å)	Tilt (°)	Roll (°)	Twist (°)
ave.	0.01	-0.46	3.41	0.02	2.32	34.45
s.d.	0.33	0.10	0.30	2.81	5.16	3.01
	X-disp (Å)	Y-disp (Å)	h-Rise (Å)	Incl (°)	Tip (°)	h-Twist (°)
ave.	-1.28	-0.01	3.32	4.41	-0.04	35.01
s.d.	1.10	0.56	0.23	9.06	5.13	2.46

Table 5.1.2. Major and minor groove distances for the 1DUF Dickerson dodecamer from the RCSB Protein Data Bank,[40] using the 3DNA software package.[51]

	Minor Groove		Major Groove	
Bases	P-P	Refined	P-P	Refined
1 CG/CG	---	---	---	---
2 GC/GC	---	---	---	---
3 CG/CG	12.4	---	19.2	---
4 GA/TC	11.7	11.6	18.1	18.0
5 AA/TT	10.7	10.6	18.0	18.0
6 AT/AT	10.2	10.2	17.5	17.5
7 TT/AA	10.7	10.6	18.0	17.9
8 TC/GA	11.6	11.6	18.0	17.9
9 CG/CG	12.4	---	19.1	---
10 GC/GC	---	---	---	---
11 CG/CG	---	---	---	---

As mentioned previously, understanding the recognition of repair enzymes to the 8-oxo-A lesion could be clarified by comparing its helical parameters to those of the unmodified DDD. This is dependent on robust model of the molecules, with accuracy honed by internuclear distances taken from several experiments, with the ^{1}H NMR values in this thesis comprising a part. As such, the data in T5.1.1-2 provide only a basis for comparison, since the model of the 8-oxo-dA5 DDD has not yet reached maturation.

CHAPTER 6 Conclusions

Several complicating factors must be understood before concluding upon NMR-derived geometries of DNA. Firstly, it takes many more experiments than described in this thesis to gather interatomic distances of the whole molecule. Only a few hydrogen nuclei were found in the linear range of the experimental parameters and of those, even fewer presented within 15% of previously published distances. As such, the benefit of adding these NMR restraints to an existing model is not impressive. Likewise, modeling the sequence with an 8-oxo-dA5 lesion based on these restraints is barely empirical. However, the process could be, presuming a fecundity of data, along with NMR-derived distances on the other elements in the molecule. Rather than focus on impact, the modeling exercise sets into motion an expandable method.

What can be concluded upon is the general landscape of the 8-oxo-dA5 DDD in the context of chemical shift assignment. The changes in select ^{1}H nuclei of G4, A5, A6, and T7 create a general boundary for the lesion, with the most shifted nuclei being the G4 1′, A5 2″ and 1′, and the T7 CH_3. Since chemical shift did not appreciably change with temperature, an internal flexibility not seen in the unmodified sequence is unlikely, although a wider range of temperatures would be more authoritative, as would more atoms. The 8-oxo-dA5 lesion showed similarities to those found in the 8-oxo-dG4. Both sequences experienced a deshielding of the 2′ and 1′ protons and shielding of the 2″. Both affected their 5′ neighboring residue by shielding effects on their 1′ and 2″ protons.

In the 25 °C samples, it is unknown what feature corresponds with the splitting of cytosine H6-1′and some other NOEs into tetrads, although it has been hypothesized to be related to spin diffusion and the mobility to change between cross-relaxation pathways.

Signal loss was another problem, leading to the disappearance of several aromatic-1′ peaks, which ultimately led to relying on 10 °C assignments. While some quality data was taken from the 25 °C acquisitions, lower temperatures are recommended. This is especially important since shorter mixing times are ideally suited for large molecules to reduce spin diffusion, an effect most prominently seen at shorter distances. The caveat is that shorter mixing times produce lower intensity spectra, which are more susceptible to noise artifacts.

Only two of the nine H-H distances ascertained to be valid in the 8-oxo-dA5 sequence deviated from the unmodified DDD by more than the tolerance limit of 15%, making them candidates for structural distortion. These were the A5 2′-3′ and the A5 2″-1′. While there was no obvious indication of changes in the 3′ hydrogen of A5 in terms of chemical shift assignment, the missing peaks between A5 and G4, along with how the 8-oxo-G4 DDD was affected at the 1′, substantiate the role of the oxopurine 1′ as a fingerprint for this form of damage.

CHAPTER 7 Future Work

The empirical relevance of the model produced in this thesis can only go so far as the number of NMR-derived distances going into it. Performing another series of build-up curves at 10 °C and at different mixing times would be ideal, with longer ^{1}H-^{1}H distances (>0.8 Å) appearing at long mixing times (>100 ms) and shorter distances suffering from less spin diffusion at short mixing times. Optimizing the sampling conditions for these two classes of internuclear distances would ensure that the conditions are suited better for both.

Additionally, heteronuclear single quantum coherence spectroscopy HSQC experiments for determining phosphate backbone parameters, along with ^{13}C NMR would fill gaps in the structure. Comparing these changes to those found in the 8-oxo-dG4 DDD would create for a more robust understanding of the helical alterations induced by oxopurines, with the ultimate goal being the detection of common elements of enzyme recognition for the repair of these lesions.

REFERENCES

(1) Avery OT, Macleod CM, McCarty M. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type III. J. Exp. Med. 1944; 79(2): 137-58.

(2) Watson JD, Crick FHC. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature. 1953; 171(4356): 737–738.

(3) Drew HR., Wing RM., Takano T, Broka C, Tanaka S, Itakura K, Dickerson RE. Structure of a B-DNA dodecamer: conformation and dynamics. Proc. Natl. Acad. Sci. USA. 1981; 78: 2179-2183

(4) Henderson E, Hardin CC, Walk SK, Tinoco I Jr., Blackburn EH. Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine·guanine base pairs. Cell. 1987; 51(6): 899-908.

(5) Hoogsteen K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. Acta Crystallographica. 1963; 16: 907–916.

(6) Alexander R. DNA comes in many forms. Gene. 1993; 135(1–2): 99–109.

(7) Breaker RR.,Gerald JF. A DNA enzyme that cleaves RNA. Chem Biol. 1994; 1 (4): 223–229.

(8) Frommer J, Appel B, Müller S. Ribozymes that can be regulated by external stimuli. Curr. Opin. Biotechnol. 2015; 31C: 35-41.

(9) Muhire1 BM, Golden M, Murrell B, Lefeuvre P, Lett J, Gray A, Poon AYF, Ngandu NK, Semegni Y, TanovEP, Monjane1AL, Harkins GW, Varsani A, Shepherd DN, and Martin DP. Evidence of pervasive biologically functional secondary-structures within the genomes of eukaryotic single-stranded DNA viruses. J. Virol. 2014; 88(4): 1972-89.

(10) Wheeler SE and Houk KN. Substituent Effects in the Benzene Dimer are Due to Direct Interactions of the Substituents with the Unsubstituted Benzene. J. Am. Chem. Soc. 2008; 130(33): 10854–10855.

(11) Karabıyık H, Sevinçek R, and Karabıyık H. π-Cooperativity effect on the base stacking interactions in DNA: is there a novel stabilization factor coupled with base pairing H-bonds? Phys. Chem. Chem. Phys. 2014; 201; 16: 15527-15538.

(12) Ageno M, Dore E, and Frontali C. The Alkaline Denaturation of DNA. Biophys. J. 1969; 9(11): 1281–1311.

(13) Wang X, Lim HJ, and Son A. Characterization of denaturation and renaturation of DNA for DNA hybridization. Environ. Health. Toxicol. 2014; 29: e2014007.

(14) Rychlik W, Spencer WJ, and Rhoads RE. Optimization of the annealing temperature for DNA amplification in vitro. Nucl. Acids Res. 1990; 18(21): 6409–6412.

(15) Wang JC. Helical repeat of DNA in solution. Proc. Natl. Acad. Sci. U. S. A. 1979; 76 (1): 200–203.

(16) Gilbert N and Allan J. Supercoiling in DNA and chromatin. Curr. Opin. Genet. Dev. 2014; 25 (100): 15–21.

(17) Siggers T, and Gordân R. Protein–DNA binding: complexities and multi-protein codes. Nucl. Acids Res. 2014; 42(4): 2099–2111.

(18) Wang S, Kool ET. Origins of the large differences in stability of DNA and RNA helices: C-5 methyl and 2'-hydroxyl effects. Biochemistry. 1995; 34(12): 4125-32.

(19) Valavanidis A1, Vlachogianni T, Fiotakis C. 8-hydroxy-2' -deoxyguanosine (8-OHdG): A critical biomarker of oxidative stress and carcinogenesis. J. Environ. Sci. Health. C. Environ. Carcinog. Ecotoxicol. Rev. 2009; 27(2):120-39.

(20) Maillard B, Ingold K U, Scaiano. Rate constants for the reactions of free radicals with oxygen in solution. J. C. J. Am. Chem. Soc. 1983; 105: 5095.

(21) Skovsen E, Snyder JW, Lambert JDC, and Ogilby PR. Lifetime and Diffusion of Singlet Oxygen in a Cell. J. Phys. Chem. B. 2005; 109: 8570.

(22) Chin KK, Trevithick-Sutton CC, McCallum J, Jockusch S, Turro NJ, Scaiano JC, Foote CS, and Garcia-Garibay MA. Quantitative Determination of Singlet Oxygen Generated by Excited State Aromatic Amino Acids, Proteins, and Immunoglobulins. J. Am. Chem. Soc. 2008; 130: 6912.

(23) Thomas DD, Liu X, Kantrow SP, and Lancaster JR. The biological lifetime of nitric oxide: Implications for the perivascular dynamics of NO and O2. Proc. Natl. Acad. Sci. 2001; 98: 355.

(24) Ayala A, Muñoz MF, Argüelles S. Lipid Peroxidation: Production, Metabolism, and Signaling Mechanisms of Malondialdehyde and 4-Hydroxy-2-Nonenal. Oxid. Med. Cell. Longev. 2014; 2014: 360438.

(25) Wang X, Yu S, Wang CY, Wang Y, Liu HX, Cui Y, Zhang LD. Advanced glycation end products induce oxidative stress and mitochondrial dysfunction in SH-SY5Y cells. In Vitro Cell Dev. Biol. Anim. 2014; 51(2): 204-9.

(26) Fuentealba D, Friguet B, Silva E. Advanced glycation endproducts induce photocrosslinking and oxidation of bovine lens proteins through type-I mechanism. Photochem. Photobiol. 2009; 85: 185.

(27) Cooper J. (2000) The Cell: A Molecular Approach (2nd Ed). Massachusetts: Sinauer Associates.

(28) Hashimoto K, Tominaga Y, Nakabeppu Y, and Moriya M. Futile short-patch DNA base excision repair of adenine:8-oxoguanine mispair. Nucl. Acids Res. 2004; 32 (19): 5928-5934.

(29) Fortini P, Dogliotti E. Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. DNA Repair. 2007; 6: 398.

(30) Saparbiev M, Talhaoui I, Couve S. 7,8-dihydro-8-oxoadenine, a highly mutagenic adduct, is repaired by Escherichia coli and human mismatch-specific uracil/thymine-DNA glycosylases. Nucl. Acids Res. 2013; 41: 912.

(31) Fuss J, Cooper P. DNA repair: dynamic defenders against cancer and aging. PLoS Biol. 2006; 4: 203.

(32) Kalluri S. Mechanisms of Disease: DNA repair defects and neurological disease. Nat. Clin. Prac. Neurol. 2007; 3: 162.

(33) Francklyn CS. DNA Polymerases and Aminoacyl-tRNA Synthetases: Shared Mechanisms for Ensuring the Fidelity of Gene Expression. Biochemistry. 2008; 47(45): 11695–11703.

(34) Alberts, B, Johnson A, Lewis J, et al. (2002). Chapter 5: DNA Replication, Repair, and Recombination. Molecular Biology of the Cell (4th ed.). New York: Garland Science.

(35) Sung P, Klein H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. Nat. Rev. Mol. Cell Biol. 2006; 7: 739.

(36) Kim, Y. C., Grable, J. C., Love, R., Greene, P. J., Rosenberg, J. M.. "Refinement of EcoRI endonuclease crystal structure: a revised protein chain tracing". Science. 1990; 249(4974): 1307–1309.

(37) Hare, D.R. Wemmer, D.E. Chou, S. Drobny, G, Reid BR. Assignment of the Non-exchangeable Proton Resonance of d(C-G-C-G-A-A-T-T-C-G-C-G) Using Two-

dimensional Nuclear Magnetic Resonance Methods. J. Mol. Biol. 1983; 171: 319-336.

(38) Hare DR, Reid BR. Three-Dimensional Structure of a DNA Hairpin in Solution: Two-Dimensional NMR Studies and Distance Geometry Calculations on d(CGCGTTTTCGCG). Biochemistry. 1986; 25: 5341-5350.

(39) Wemmer DE, Chou SH, Hare DR, Reid BR. Duplex-hairpin transitions in DNA: NMR studies of CGCGTATACGCG. Nucl. Acids Res. 1985; 13: 3755-3772.

(40) Tjandra N, Tate S, Ono A, Kainosho M, and BaxA.The NMR Structure of a DNA Dodecamer in an Aqueous Dilute Liquid Crystalline Phase. J. Am. Chem. Soc. 2000; 122: 6190 6200.

(41) Kohn W and Sham LJ. Self-Consistent Equations Including Exchange and Correlation Effects. Phys. Rev. 1965; 140: A1133.

(42) Ceperley DM and Alder BJ. Ground State of the Electron Gas by a Stochastic Method. Phys. Rev. Lett. 1980; 45(7): 566.

(43) Guerra CF, Bickelhaupt FM, Snijders JG, and Baerends EJ. Hydrogen Bonding in DNA Base Pairs: Reconciliation of Theory and Experiment.J. Am. Chem. Soc. 2000; 122: 4117 4128.

(44) Riley KE, Op't Holt BT, and Merz KM Jr. Critical Assessment of the Performance of Density Functional Methods for Several Atomic and Molecular Properties. J. Chem. Theory Comput. 2007; 3(2): 407–433.

(45) Mazzanti A and Casarini D. Recent trends in conformational analysis. WIREs Comput. Mol. Sci. 2012; 2(4): 613–641.

(46) Williams CP. (2010). Explorations in Quantum Computing. Springer Science & Business Media,. pp 319-323.

(47) UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations by A.K. Rappe, C.J. Casewit, K.S. Colwell, W.A. Goddard III, W.M. Skiff, J.Am. Chem. Soc. 1992; 114(25): 10024–10035.

(48) Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations". J Comp Chem. 1983; 4(2): 187–217.

(49) M. Korth. Third-Generation Hydrogen-Bonding Corrections for Semiempirical QM Methods and Force Fields. J. Chem. Theory Comput. 2010; 6(12): 3808-3816.

(50) Chung LW, Sameera WMC, Ramozzi R, Page AJ, Hatanaka M, Petrova GP, Harris TV, Li X, Ke Z, Liu F, Li H, DingL, andK Morokuma et al. The ONIOM Method and Its Applications. Chem. Rev. 2015; 115(12): 5678–5796.

(51) Xiang-Jun Lu & Wilma K. Olson. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. Nucl. Acids Res. 2003; 31(17): 5108-21.

(52) Marion D. An Introduction to Biological NMR Spectroscopy. Mol. Cell. Proteomics. 2013; 12(11): 3006-3025.

(53) Günther H. (2013). NMR Spectroscopy: Basic Principles, Concepts and Applications in Chemistry. Weinheim, Germany: Wiley-VCH.

(54) Morris GA and Emsley JW. (2012). Multidimensional NMR Methods for the Solution State. John Wiley & Sons.

(55) Hans J. Reich. (b) 8.2 The Nuclear Overhauser Effect. University of Wisconsin. ©2016. [Cited August 4, 2016.] https://www.chem.wisc.edu/areas/reich/nmr/08-tech-02-noe.htm

(56) Syngen Inc. San Carlos (CA); ©2000-2011 [Cited July 21, 2016.] http://www.syngendna.com/

(57) Gottlieb, H., Kotlyar, V., and Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. J. Org. Chem. 1997; 62: 7512-7515.

(58) Advanced Chemistry Development, Inc. ©1997-2010. [Cited August 4, 2016.] http://www.acdabs.com/

(59) Mestrelab Research, SL (Spain). © 2004-2016. [Cited September 2, 2016.] http://www.mestrelab.com/

(60) Gaussian 09, Revision W.09, Frisch MJ, Trucks GW, Schlegel HB,Scuseria GE, RobbMA, CheesemanJR,Scalmani G, Barone V,Mennucci B,Petersson GA,Nakatsuji H,Caricato M, Li X,Hratchian H P,Izmaylov AF, Bloino J, Zheng G,Sonnenberg JL,Hada M,Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O,Nakai H, Vreven T, Montgomery JA, Peralta JE, Ogliaro F,Bearpark M,Heyd JJ, Brothers E,Kudin KN,Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, BurantJC, IyengarSS, Tomasi J,Cossi M,Rega N,Millam JM,Klene M, Knox JE, Cross JB, Bakken V,Adamo C, Jaramillo J, Gomperts R,Stratmann RE, Yazyev O, Austin AJ,Cammi R,Pomelli C,Ochterski JW, Martin RL,Morokuma K,Zakrzewski VG,Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD,Farkas Ö,Foresman JB, Ortiz JV,Cioslowski J,and Fox DJ. Gaussian, Inc., Wallingford CT, 2009. http://www.gaussian.com/

(61) Altona C, Sundaralingam M. Conformational Analysis of the Sugar Ring in Nucleosides and Nucleotides. A New Description Using the Concept of Pseudorotation. J. Am. Chem. Soc. 1972; 94(23): 8205-12.

(62) E. Westhof, M. Sundaralingam. A method for the analysis of puckering disorder in five-membered rings: the relative mobilities of furanose and proline rings and their effects on polynucleotide and polypeptide backbone flexibility. J. Am. Chem. Soc. 1983; 105(4): 970–976.

(63) Dreizler RM, Gross EKU. (1990). Density Functional Theory: An Approach to Quantum Many-Body Problem. Springer, Berlin.

APPENDICES

Appendix A: Equations (Kohn-Sham and LDA)

Firsthand, it must be noted that the equations comprising DFT are many and beyond the scope of this thesis. A full description on the equations of DFT can be found in a volume by Dreizler and Gross[63]. Broadly speaking, DFT is a method for approximating the properties of atoms, which are quantum-mechanical (QM) systems. For a QM system, all information is contained within the wavefunction (Ψ). The Schrödinger equation is the method for solving this and works on the basis of eigenfunction transformation. In its most general sense, this can be arranged as:

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

where the Hamiltonian operator $\hat{H}$ operates on the wavefunction Ψ such that the same Ψ is returned, multiplied by a constant, the eigenvalue E. Unfortunately, for atoms with more than 2 interacting particles, it is impossible to solve the Schrödinger equation because of repeated, unpredictable interactions between them. Approximations must be made to reduce the number of variables and treat them in solvable ways. For Density Functional Theory (DFT), ϕ simplifies the eigenfunction Ψ of a many-electron molecule by treating it in terms of electron density, which is a condensed single entity as if the atom had non-interacting electrons.

Electron density is useful because according to the Hohenburg-Kohn theorem, the ground-state density of any system is a function of and determines all of its ground-state properties. For atoms, it is electron density that determines the ground-state energy. Electron density is treated in terms of the spatial coordinates x, y, and z (instead of 3N

variables for a many-electron atom), so derived values, such as bond-lengths and angles, can be rendered from this 3D energy matrix.

From a mathematical standpoint, the electron density n(r) is a single potential ν(r) acting upon the electrons in the system ϕ, which depends only on the spatial and spin coordinates of each. As a result, the one-electron equation appears as follows:

$$\left[-\frac{1}{2}\nabla^2 + \nu(r)\right]\phi(r) = E\phi(\mathrm{r})$$

where $-\frac{1}{2}\nabla^2$ is the kinetic component added to the local potential of the electron. Solutions to this equation produce orbitals (ϕ_i) representing where the electron density is for the atom or molecule, seen as

$$n(r) = \sum_{i=1}^{N} \int |\phi_i\,(r)^2|$$

This one-particle equation must represent a system of N particles and yield a ground-state energy. The Hohenberg-Kohn Theorem does this by treating the problem in terms of a functional of N, as its density n(r). This changes the ground-state energy (E) of that density to

$$E[n(r)] = F[n(r)] + \int n(\mathrm{r})\ \nu(r)\, dr$$

Since n(r) and v(r) are defined, all that remains is to define the functional F[n(r)]. Kohn and Sham divided F[n(r)]into two parts, the kinetic energy E_k functional of a non-interacting electron gas with density n(r) and the exchange-correlation energy E_{XC}. The kinetic energy component is

$$E_k[n(r)] = -\frac{1}{2}\sum_{i=1}^{N} \int \phi^*(r)\ \nabla^2\phi(r)\ dr$$

The E_{XC} component is defined by the local density approximation (LDA). Since this is computed by quantum Monte Carlo (QMC) methods, which are beyond the scope of this thesis, E_{XC} will simply be represented as

$$E_{XC}^{LDA}[n(r)] = \int \epsilon_{XC}\left(n(r)\right)n\,(r)dr$$

In addition, F[n(r)] is defined such that

$$F[n(r)] = \langle \phi | \hat{\mathrm{F}} | \phi \rangle \qquad \text{and} \qquad \hat{\mathrm{F}} = \sum_i -\frac{1}{2}\nabla_i^2 + \frac{1}{2}\sum_{i \neq j}\left|\frac{1}{r_i - r_j}\right|$$

The full Kohn-Sham equation for energy is a combination of these parts, becoming

$$E[n(r)] = -\frac{1}{2}\sum_{i=1}^{N}\int \phi^*(r)\,\nabla^2\phi(r) + \frac{1}{2}\int\int \frac{\left(n(r)\right)n\,(r')}{|r-r'|}drdr' + \int \epsilon_{XC}\left(n(r)\right)n\,(r)dr + \int n(\mathrm{r})\;v(r)\,dr$$

The extra term $\int \frac{(n(r))n\,(r\prime)}{|r-r\prime|}drdr'$ comes from normalizing the Kohn-Sham equation with respect to the electron density and rewriting it as an effective potential.

$$v_{eff}(r) = v(r) + \int \frac{n(r')}{|r-r'|} + \frac{\partial E_{XC}[n(r)]}{\delta n(r)}$$

To recapitulate, the user-created molecule is represented in DFT by the Kohn-Sham equations as a fictional system, the wavefunction ϕ of a non-interacting electron gas. From there, the observables are extracted from ϕ by calculating their expectation values, using the operations defined by the Kohn-Sham equation. Ultimately and after intensive computation, the result is an electron density solution that approximates the molecule.

Appendix B: Supplemental Material

The following tables organize all files generated by the NMR for this thesis, along with representations of that data, such as spreadsheets and files generated by modeling and graphing software. This information is included in the supplemental disk, but can also be accessed by request at Reynolds2600@live.missouristate.edu.

Mixing	Sequence	Filename
200 ms	DDD	DDD_pH7p4_longer_mix_20150904_01/data_NOESY_002/fid
200 ms	8-oxo-dA5	20151102_01/data_NOESY_001/fid
150 ms	DDD	DDD_pH7p4_longer_mix_20150904_01/data_NOESY_001/fid
150 ms	8-oxo-dA5	20151030_02/data_NOESY_001/fid
100 ms	DDD	DDD_pH7p4_multimix_20150827_01/data_NOESY_003/fid
100 ms	8-oxo-dA5	20151030_02/data_NOESY_002/fid
75 ms	DDD	DDD_pH7p4_multimix_20150827_01/data_NOESY_002/fid
50 ms	DDD	DDD_pH7p4_multimix_20150827_01/data_NOESY_001/fid

Temperature	Sequence	Filename
10 °C	DDD	s_20110429_01/data/d2o_02.fid/fid
10 °C	8-oxo-dA5	s_20140312_002/data/20140312T130052_01.fid/fid
10 °C	8-oxo-dA5	s_20140312_002/data/20140313T224102_01.fid/fid

Project Description	Software	Filename
Spreadsheet with all NMR assignments, intensities, and distances.	MS Excel 2007	Intensity Comparison Mixing Times.xlsx
DDD model generated by Tjandra et al	Gaussian/XPLOR	1DUF.pdb
DDD model generated by Dickerson et al.	Gaussian/XPLOR	1BNA.pdb
DDD model generated by the methods and restraints in this thesis.	Gaussian/XPLOR	Future work; model completed with errors: DDD_thesis.pdb
8-oxo-dA5 DDD model generated by the methods and restraints in this thesis.	Gaussian/XPLOR	Future work; model completed with errors: 8OA5DDD_thesis.pdb
8-oxo-dA5 DDD model minimized by UFF	Gaussian	8OA5DDD_UFF.log
MNOVA collection of DDD spectral images.	MestReNova 10.0.2-15465	DDD_5mm_25C.mnova
MNOVA collection of 8-oxo-dA5 DDD spectral images.	MestReNova 10.0.2-15465	8-oxo-A_5mm_25C.mnova