Missouri State.
UNIVERSITY

BearWorks

Summer 2023

Impact of Sample Conditions on DNA Phosphodiester Backbone BI/BII Conformational Equilibrium Dynamics

Autumn C. Pilarski
Missouri State University, Pilarski4672@live.missouristate.edu

Follow this and additional works at: https://bearworks.missouristate.edu/theses

Part of the Biochemistry Commons, Biophysics Commons, and the Physical Chemistry Commons

Recommended Citation

IMPACT OF SAMPLE CONDITIONS ON DNA PHOSPHODIESTER BACKBONE

BI/BII CONFORMATIONAL EQUILIBRIUM DYNAMICS

A Master's Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Chemistry

By

Autumn Catherine Pilarski

August 2023

IMPACT OF SAMPLE CONDITIONS ON DNA PHOSPHODIESTER BACKBONE

BI/BII CONFORMATIONAL EQUILIBRIUM DYNAMICS

Chemistry

Missouri State University, August 2023

Master of Science

Autumn Catherine Pilarski

ABSTRACT

DNA damage, such as single base lesions and mismatches, is highly prevalent within cells. If these DNA damage events are not repaired, they could lead to mutations and thus disease and cancer. Intricate repair mechanisms are in place to fix these damage events, one such being Base Excision Repair (BER) and associated enzyme: Thymine DNA Glycosylase (TDG). The first step of this repair process, recognition of the lesion by TDG, is not well understood. The following thesis presents results to better understand the fundamental biophysical question of how a DNA lesion within a mismatch context is recognized in a million fold excess of normal DNA bases. Part one of this research involves the investigation of an 8mer non-palindromic DNA sequence containing a U:G mismatch and how this mismatch alters the phosphodiester backbone of the DNA around the lesion site. Results showed that there was a stepwise difference in the free energy of the conformational equilibrium of the DNA backbone around the lesion site, following trends that have been shown for palindromic DNA sequences. Part two investigates attempts to optimize the expression and purification conditions for Thymine DNA Glycosylase. Part three involves analyzing how DNA sample conditions influence the backbone conformational equilibrium dynamics within a dodecamer palindromic sequence containing a T:G mismatch. Of particular interest, the free energy was impacted for the lesion sequence at the phosphate 3′ to the base-pairing partner in the presence of potassium and magnesium ions and impacted to a lesser extent at the phosphate 3′ to the lesion in the presence of potassium ion. For the DNA analysis of the phosphodiester conformational equilibrium dynamics, 1-D and 2-D Nuclear Magnetic Resonance (NMR) Spectroscopy was used.

KEYWORDS: DNA, Base Excision Repair, Thymine DNA Glycosylase, 1D and 2D solution Nuclear Magnetic Resonance Spectroscopy, BI/BII, dynamics, conformations, equilibrium, sample conditions

IMPACT OF SAMPLE CONDITIONS ON DNA PHOSPHODIESTER BACKBONE

BI/BII CONFORMATIONAL EQUILIBRIUM DYNAMICS

By

Autumn Catherine Pilarski

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

August 2023

Approved:

Gary Meints, Ph.D., Thesis Committee Chair

Natasha DeVore, Ph.D., Committee Member

Gautam Bhattacharyya, Ph.D., Committee Member

Kyoungtae Kim, Ph.D., Committee Member

Julie Masterson, Ph.D., Dean of the Graduate College

In the interest of academic freedom and the principle of free speech, approval of this thesis indicates the format is acceptable and meets the academic criteria for the discipline as determined by the faculty that constitute the thesis committee. The content and views expressed in this thesis are those of the student-scholar and are not endorsed by Missouri State University, its Graduate College, or its employees.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Introduction to DNA and Uracil

Deoxyribonucleic acid (DNA) is likely the most important biomolecule to life, as it can be considered the blueprint, encoding for biological information. DNA is made of monomeric nucleotide units which include a phosphate group, a deoxyribose sugar, and a base.[1] These nucleotides are connected via a phosphodiester linkage creating a phosphodiester backbone, with the bases sitting internally within the helix.[1] DNA has three forms: A, B, and Z.[1] The B form is a right-handed double helix created by hydrogen bonding two anti-parallel strands of DNA nucleotides, with 10 residues per turn.[1] The hydrogen bonding between the two strands is created with the base pairs of pyrimidines and purines, one- and two-membered rings, respectively.[1] These base pairs are stacked on top of each other, but to create the helical nature each pair is rotated approximately 36° with respect to each other.[1] This stacking creates a minor and major groove within the helix.

The canonical base pairs of DNA contain a purine (adenine or guanine) and pyrimidine (thymine or cytosine) hydrogen bonded to each other. Adenine (A) pairs with thymine (T) via two hydrogen bonds, and guanine (G) pairs with cytosine (C) via three hydrogen bonds[1] (Figure 1). An additional base that may be observed in DNA due to error is a uracil base. Uracil is one of the four bases found in RNA, holding the same position as thymine for DNA (A base-pairs with U, C with G).[1] The structure of this base can be found in Figure 2.

To effectively communicate about DNA, researchers use a particular numbering scheme for the base and sugar protons. Below in Figure 3 is a deoxyuridine monophosphate nucleotide with the relevant protons on the deoxyribose sugar and base labeled in different colors.

1
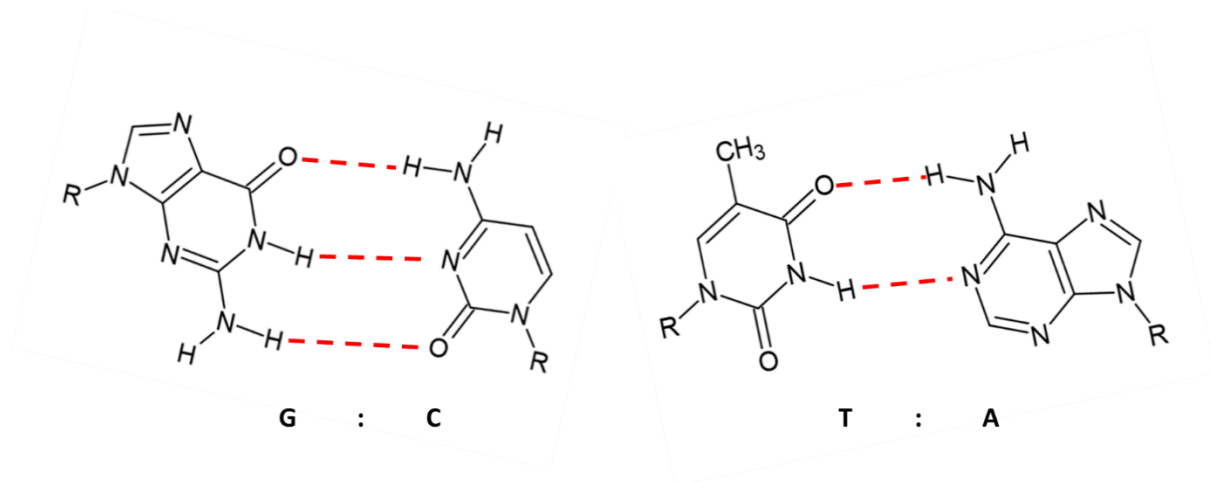
Figure 1. Canonical base pairing found in DNA. A guanine (G) base hydrogen bonded to a cytosine (C) base (left). A thymine base (T) hydrogen bonded to an adenine (A) base (right). Hydrogen bonds are indicated by the red dashed lines. "R" indicates the rest of the nucleotide.
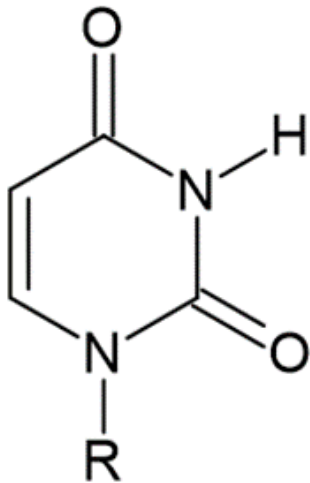
Figure 2. The structure of a uracil base. "R" indicates the rest of the nucleotide.

DNA Damage

DNA is prone to damage. It is estimated that up to 1,000,000 DNA damage events happen per cell per day.[2] These damage events can happen through a variety of mechanisms,

such as alkylating agents, reactive oxygen species, UV light, and environmental exposures.[2] A few key examples of DNA damage and how each type is repaired by the varying repair mechanisms found within cells is in Figure 4 below.
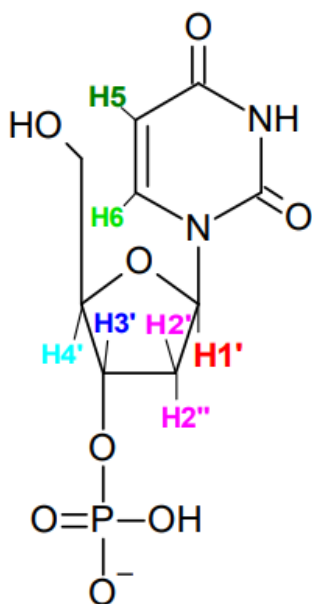
Figure 3. Numbering scheme of the protons of a deoxyuridine monophosphate nucleotide. Notice the phosphate group, deoxyribose sugar, and base, creating the nucleotide monomer.

For example, a DNA damage event can be the spontaneous deamination of a cytosine to a uracil lesion, creating a U:G mismatch within the DNA structure,[3] as can be seen from Figure 5. This process of spontaneous deamination of cytosine to uracil is estimated to occur 100-500 times per day in mammalian cells.[3] Another common process within cells is the methylation of cytosine using a methyltransferase enzyme to create 5-methylcytosine.[4] This 5-methylcytosine can then deaminate as well to form thymine, and thus create a T:G mismatch.[4] An important note to make is that 5-methylcytosine usually arises at CpG sites.[5-6] These U:G and T:G mismatches

produce C→T transition mutations.[7] Mutations within the double helix are well known to have

the potential to lead to illness, such as cancer and disease.[8]
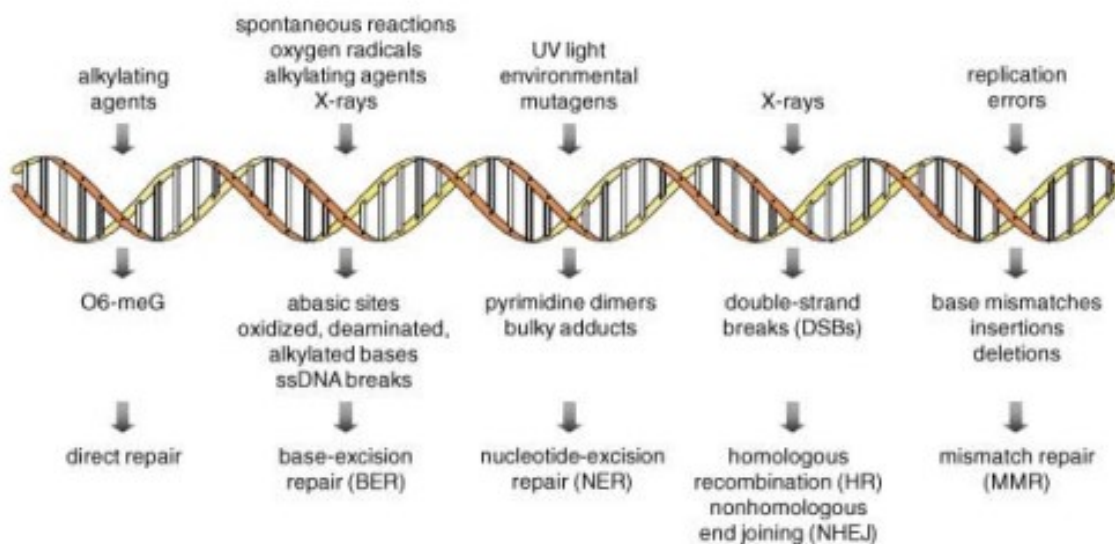
Figure 4. Examples of DNA damage and their corresponding repair mechanisms. Figure was obtained with permission from Schärer, O. D. Chemistry and Biology of DNA Repair. *Angew. Chem. Int. Ed*. **2003**, *42*, 2946-2974, published by John Wiley and Sons.
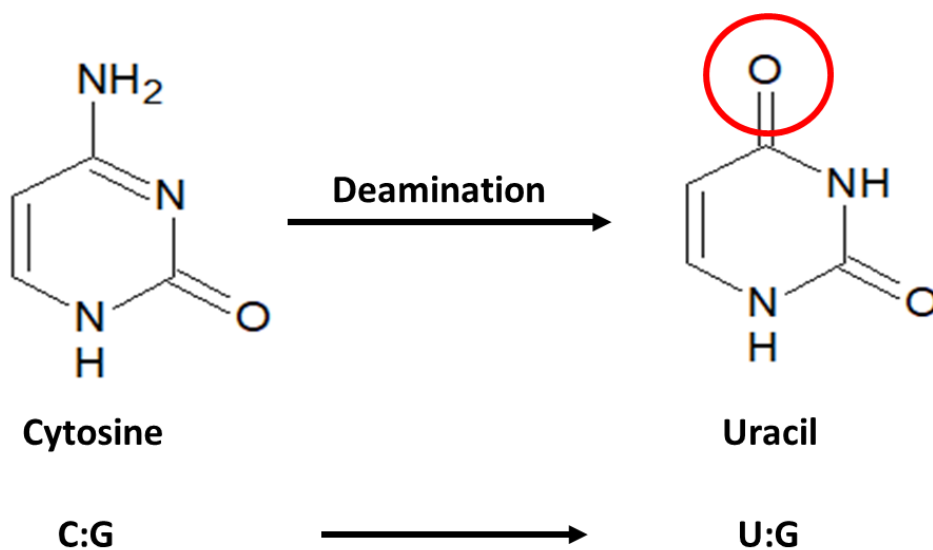
Figure 5. Visualization of cytosine deamination to a uracil. The change in the base from cytosine to uracil is indicated via the red circle.

Base Excision Repair and Thymine DNA Glycosylase

Luckily, cells have intricate repair processes to correct DNA damage events, including lesions, to prevent disease from constantly wreaking havoc on cells. One such repair process is Base Excision Repair (BER).[2,9] The general process involves the use of a DNA glycosylase which recognizes the DNA lesion and then hydrolytically cleaves the N-glycosyl bond, creating an abasic site, and the correct base incorporated by other associated enzymes.[9] BER has been found to play a critical role in epigenetic regulation through DNA demethylation.[10]

One type of DNA glycosylase that can initiate BER is Thymine DNA glycosylase (TDG). DNA glycosylases have been found to use a particular mechanism of flipping the damaged base out of the double helix after recognition.[11] The proposed flipping step of BER, where the damaged base is flipped out from the DNA helix, is shown in Figure 6,[11] and is defined as K_{flip} within the mechanism presented in Figure 7.[12]
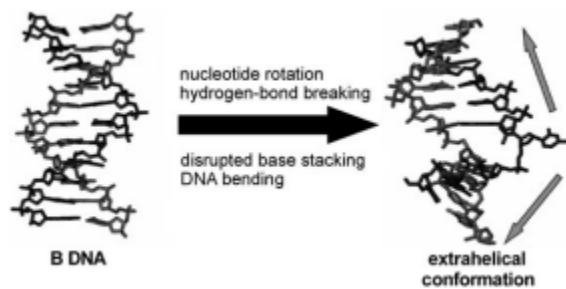
Figure 6. Nucleotide flipping step of BER. Image obtained with permission from Stivers, J. T. Extrahelical Damaged Base Recognition by DNA Glycosylase Enzymes. *Chem. Eur. J.* **2008**, *14*, 786-793, published by John Wiley and Sons.

The kinetics of TDG reacting with a substrate can be defined by the mechanism below in Figure 7.[12] Within this mechanism, E represents the enzyme (DNA glycosylase, such as TDG), S is the substrate for the enzyme (damaged DNA), ES is the enzyme-substrate complex, B

represents the cleaved base, and P represents the abasic product.[12] The steps that are of particular

importance for this research and investigation is the first step of recognition of the lesioned base

within a mismatch context and the K_{flip}, where TDG flips out the lesioned base via breaking

hydrogen bonds and rotating the nucleotide from the helical base-stacking.[11-12]

$$E + S \rightleftharpoons E{\cdot}S \overset{K_{flip}}{\rightleftharpoons} E'{\cdot}S' \overset{k_{chem}}{\longrightarrow} E'{\cdot}B{\cdot}P' \longrightarrow E + P + B$$

$$\underline{\hspace{6cm}}_{k_{max}} \qquad\qquad E{\cdot}P$$

Figure 7. Kinetic mechanism of TDG reacting with a substrate. Figure was reprinted (adapted) with permission from Dow, B. J.; Malik, S. S.; Drohat, A. C. Defining the Role of Nucleotide Flipping in Enzyme Specificity Using ^{19}F NMR. *J. Am. Chem. Soc.* **2019**, *141*, 4952-4962. Copyright 2019 American Chemical Society.

TDG is a part of the Uracil DNA Glycosylase (UDG) superfamily,[6] where there are three

other enzymes that have UDG activity within mammalian cells: UNG, SMUG1, and MBD4.[13] Of

these four, TDG has a large substrate spectrum,[5,13] with its highest affinity being for U:G

mismatches, and another important substrate being for T:G mismatches.[13] The affinity for U:G

and T:G mismatches within DNA makes it an optimal repair enzyme to study for this project

analyzing how DNA is repaired. For a visualization of TDG, Figure 8 below shows a crystal

structure of hTDG (human TDG) complexed with a DNA substrate.

In comparison to U:G and T:G mismatches, TDG does not have a high substrate affinity

for cytosine or 5-methylcytosine.[14] Research has been published showing that the specificity of

hTDG is dependent on the stability of the N-glycosidic bond, and TDG has poor affinity for

cytosine because of cytosine's very poor leaving ability.[14] It has also been found that TDG has

preferential activity for thymine within CpG:T sites, while uracil has a much smaller dependence

on this sequence context.[5] This preference of TDG activity at CpG:T sites compared to other

sites (GpG, TpG, and ApG) could be to restrict excision to thymine lesions that arise by 5-methylcytosine deamination occurring at CpG sites.[5-6]

Figure 8. Structure of hTDG complexed with a DNA substrate (broad image on the left and zoomed in image on the right). Within the structure on the right, a base can be seen flipped into the enzyme structure. Image from the RCSB PDB (RCSB.org) of PDB ID #5HF7.[15] (The source that originally published this crystal structure was Coey, C. T.; Malik, S. S.; Pidugu, L. S.; Varney, K. M.; Pozharski, E.; Drohat, A. C. Structural Basis of Damage Recognition by Thymine DNA Glycosylase: Key Roles for N-Terminal Residues. *Nucleic Acids Research.* **2016**, *44* (21), 10248-10258.)[16]

This enzyme is believed to also have important contacts with the phosphodiester backbone of DNA,[6] which could play a role within the recognition process under investigation. A few key phosphate contacts between the DNA with a flipped base and TDG are below in Figure 9. Within this figure, the orange dots represent the phosphates within the phosphodiester backbone, and the blocks with A, T, C, and G represent bases hydrogen bonded to each other. The red pentagon is the flipped base. The arrows indicate where a multitude of amino acid residues within TDG interact via hydrogen bonds.[6] As can be seen, these arrows mostly point to areas on the phosphodiester backbone. Residues N157, S200, S273, and R275 are just a few of the residues of TDG that interact with the phosphodiester backbone of the DNA sequence through hydrogen bonding when a lesion is flipped out of the DNA helix.[6] As can be seen, the

residues interacting with the phosphodiester backbone are not positively charged even though the DNA backbone is negatively charged. Thus the interactions are not necessarily electrostatic, which would be a strong interaction, and are more likely hydrogen bonding, a weaker interaction.
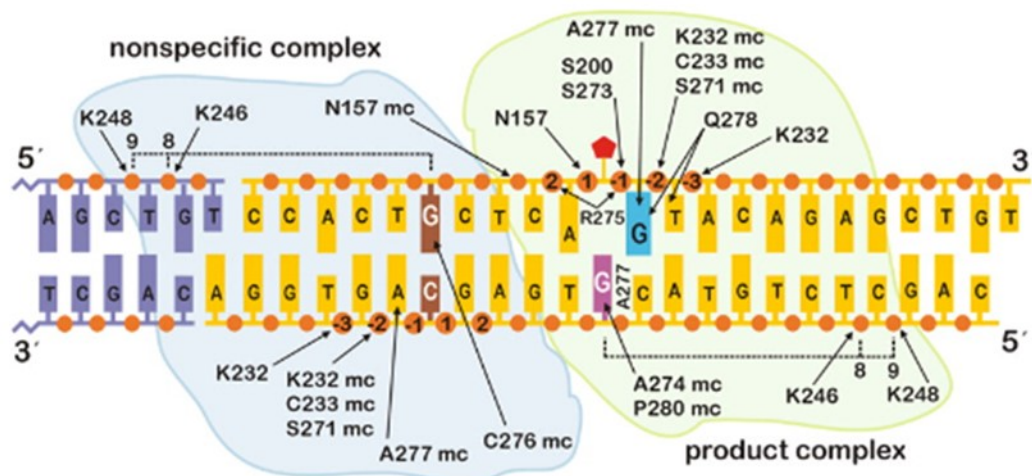
Figure 9. TDG Enzyme – DNA complex interactions. Figure was obtained (adapted) with permission from Maiti, A.; Morgan, M. T.; Pozharski, E.; Drohat, A. C. Crystal Structure of Human Thymine DNA Glyosylase Bound to DNA Elucidates Sequence-Specific Mismatch Recognition. *PNAS*. **2008**, *105* (26), 8890-8895. Copyright (2008) National Academy of Science, U.S.A.

BI/BII Backbone Conformations

Though there is a significant amount of research that has been completed about BER and TDG, the question of how TDG recognizes a lesion remains unanswered. This question is made even more difficult knowing that there is a million-fold excess of undamaged base-pairs.[6,14] Some articles have published about specifically how TDG interacts with the phosphodiester backbone of DNA,[6,16] as discussed in the previous section. Coey et al. published a study in 2016 investigating the importance of N-terminal residues in enhanced biochemical activity, posing at the end of their article that this could potentially be due to interactions of these residues with the

DNA backbone.[16] Because of the interactions of TDG with the phosphodiester backbone of DNA, looking to the backbone to potentially answer the question of lesion recognition is an interesting point of investigation.

Due to the potential importance of the DNA phosphodiester backbone with DNA's interaction with TDG, it is plausible the conformational states of the DNA backbone may play a role in the recognition process of BER. Within B-DNA, there are two primary conformations of the phosphodiester backbone: BI and BII.[17-18] These conformations, as can be seen in Figure 10, are defined by a set of dihedral angles: ε and ζ.[17-18] BI and BII are in equilibrium with each other, as DNA is an incredibly dynamic molecule. The BI/BII conformational equilibrium has been proposed to have a role in DNA-protein recognition, interaction, and binding.[19-21] This interconversion can be described as a two well potential energy diagram as shown in Figure 11.[22] As can be seen from this figure, BII is slightly higher in energy than BI. This difference in energy is defined by the Gibbs free energy, ΔG.
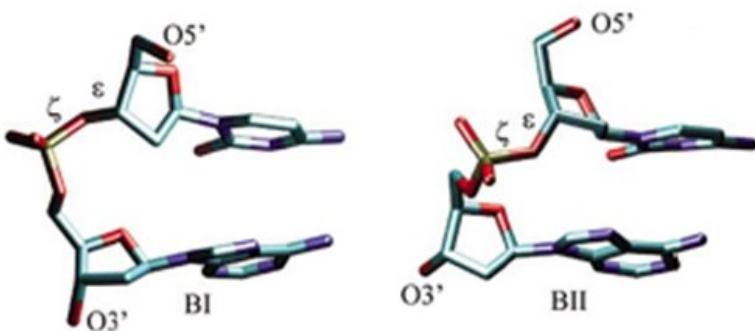
Figure 10. BI and BII conformations of DNA. Figure was obtained (adapted) with permission from Teletchéa, S.; Hartmann, B.; Kozelka, J. Discrimination Between BI and BII Conformational Substates of B-DNA Based on Sugar-base Interproton Distances. *Journal of Biomolecular Structure and Dynamics.* **2004**, *21* (4), 489-494., **© copyright # 2004**, reprinted by permission of Informa UK Limited, trading as Taylor & Taylor & Francis Group, http://www.tandfonline.com
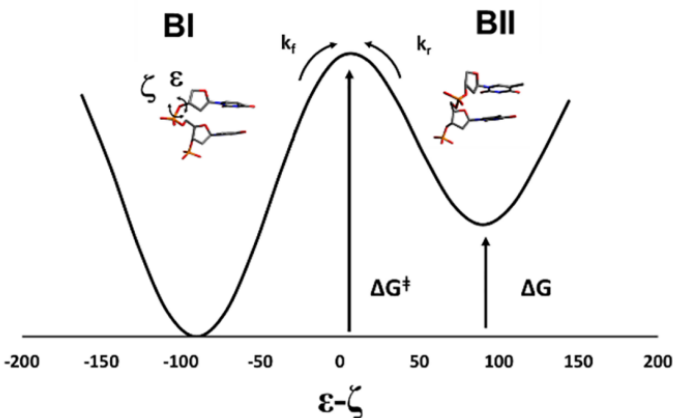
Figure 11. Two well potential energy schematic of the BI and BII equilibrium interconversion. Figure was reprinted (adapted) with permission from Westwood, M. N.; Ljunggren, K. D.; Boyd, B.; Becker, J.; Dwyer, T. J.; Meints, G. A. Single-Base Lesions and Mismatches Alter the Backbone Conformational Dynamics in DNA. *Biochemistry*. **2021**, *60*, 873-885. Copyright 2021 American Chemical Society.

Nuclear Magnetic Resonance and DNA

A significant amount of research has been completed analyzing biomolecules, such as DNA, with Nuclear Magnetic Resonance (NMR) Spectroscopy. NMR takes advantage of some nuclei, such as ^{1}H, ^{15}N, ^{19}F, and ^{31}P, having a property called spin and thus a magnetic moment.[23] NMR was first discovered in 1945 and through decades of development, several Nobel Prizes have been awarded concerning this technique including: one in chemistry to Richard Ernst in 1991 for his work on 2D NMR[23-24] and another in physiology or medicine in 2003 to Paul Lauterbur and Sir Peter Mansfield for the application of NMR concepts to biology and medicine with magnetic resonance imaging (MRI).[23,25] Biomolecular NMR is an incredibly useful tool to study the structure and dynamics of biomolecules. 2D NMR in particular, such as ^{1}H-^{1}H NOESY (Nuclear Overhauser Effect Spectroscopy) and ^{1}H-^{31}P HSQC (Heteronuclear Single-Quantum Coherence Spectroscopy) experiments are useful in studying the structure of DNA. A ^{1}H-^{1}H NOESY experiment produces peaks for two nuclei interacting within 5 Å of each other in

space.[23] This NOESY experiment allows for the assignment of protons of a DNA sequence, as has been described extensively in the literature, particularly for the well-known Dickerson-Drew Dodecamer palindromic DNA sequence.[26] In addition to assignment of the protons, the phosphorus of the DNA backbone can be assigned using a ^{1}H-^{31}P HSQC experiment.

When analyzing dynamic properties of a biomolecule using NMR, it is important to consider the time-scale on which an exchange occurs. For example, the BI/BII interconversion is an equilibrium process as presented in the previous section. If a conversion between two conformational states is considered a slow exchange, two peaks would be present on an NMR spectrum representing the two conformational states.[23] If the exchange is within the fast exchange regime, then a single peak results on the NMR spectrum resulting from a weighted average, which represents the population distribution of the conformational states.[23] The interconversion between BI and BII is considered to be within the fast exchange regime with a time-scale of $\gtrsim 10^4\,\mathrm{Hz}$[22] and thus the phosphorus peaks show on a spectrum representing a weighted average between the two conformational states. The interconversion of the BI/BII conformational states has been shown to fall within the fast-exchange and show a single peak representing the two conformations at the temperature range used within the experiments presented within this thesis and the DNA sequences that have been studied. Previous work presented by Tian et al. in 2009 analyzed the BI/BII conformational dynamics of DNA sequences using ^{31}P NMR, finding that the referenced ^{31}P chemical shift and temperature of the experiment directly correlate to %BII conformation population.[27] This %BII equation derived is dependent on the property of the peaks in the ^{31}P NMR spectrum being a weighted average of the BI and BII conformational states.

Thesis Objective

The phosphorus chemical shift value from an NMR experiment directly relates to %BII of the DNA backbone at particular phosphate steps.[27] Previous research has presented the potential implications of BI/BII interconversion with protein interaction with DNA,[19-21] as well as the interaction TDG has with the phosphodiester backbone of DNA.[6,16] The question arises whether the dynamic switching between BI and BII and the energetics associated with this interconversion could be a way by which TDG recognizes a lesion.

This has already been investigated for palindromic DNA, which found an energetic perturbation associated with this equilibrium around the lesion and base-pairing partner.[22] For the following thesis, this investigation on the energetic perturbations with the phosphodiester backbone conformational equilibrium was expanded to non-palindromic DNA containing a U:G mismatch to investigate sequence context dependence. In addition, the work used a 2′-fluoro-substituted uracil sequence to analyze the phosphate backbone to eventually complex with TDG. For this project, TDG would need to be expressed and purified and thus attempts on the optimization of the protocol is presented within this thesis. Lastly, prior research analyzed DNA within 25 mM phosphate buffer conditions.[22] The impact of sample conditions on the energetics between this BI/BII interconversion of DNA containing a lesion is unknown. The most complete part of this thesis investigates the influence of varying salt, pH, and crowding conditions on the energetics of the BI/BII conformational equilibrium of palindromic DNA containing canonical bases and DNA containing a T:G mismatch.

METHODS

DNA Preparation

The DNA sequences within this study were ordered from Integrated DNA Technologies (IDT), Coralville, IA. The DNA strands were resuspended in HPLC grade H_2O and quantitatively transferred to a Falcon tube. 1.0 mL of 25 mM phosphate buffer (pH 7.4-7.5) was added to the Falcon tube. The buffer components varied for the sample conditions project, and are listed in the first subsection below. 25 mM phosphate buffer was used unless otherwise noted.

To anneal the DNA strands, the mixture in the above falcon tube was heated in a hot water bath at 90°C for approximately 15 minutes. After heating, the sample was cooled to room temperature on the benchtop. The sample, once cooled, was frozen in liquid nitrogen and put onto the lyophilzer to dry overnight.

The dried DNA was resuspended in 1 mL of 99.9% Deuterium Oxide (Cambridge Isotope Laboratories, Inc.; Andover, MA; CAS #7789-20-0). The pH of the sample was measured using a pH probe, and adjusted to 7.35-7.45 using approximately $0.1 - 0.2$ M sodium hydroxide (NaOH) and 0.1 M phosphoric acid (H_3PO_4). After the pH was adjusted, the sample was frozen in liquid nitrogen and dried on the lyophilizer overnight. This procedure of resuspension, pH measurement and adjustment, freezing, and drying was repeated until the sample read between the 7.35-7.45 range two times consecutively without pH modification.

After the pH remained steady within this 7.4 range and NMR experiments were ready to be completed, the dried DNA was resuspended in 1 mL of 100% Deuterium Oxide (Cambridge

Isotope Laboratories, Inc.; Andover, MA; CAS #7789-20-0). The final concentration was typically 0.5-1 mM of DNA within 1 mL of D_2O.

Buffers for Sample Conditions Study. Table 1 shows the list of buffers used within the sample conditions project and their corresponding components and pH. Each buffer was made by weighing the relevant compounds listed, adding deionized (DI) water to adjust the pH to around 7.4, and then diluting to 100 mL with DI water. These buffers were used in place of the 25 mM phosphate buffer in the above procedure of DNA sample preparation within the sample conditions project presented in the Results section of this thesis.

In addition to the below sample conditions, the influence of pH and crowding agents were also investigated. For the pH conditions, the control 25 mM phosphate buffer was used with the sequences of interest. When it came time to run the pH experiments, the pH was adjusted as described in the DNA sample preparation section to a pH of around 5 for an acidic condition ($\approx$5.2 for T9 and $\approx$5.3 for DDD), and around 9 ($\approx$8.9 for DDD and T9) for a basic condition. The crowding agent samples involved DNA being prepared using Buffer 1 below (25 mM phosphate and 100 mM NaCl), and poly(ethylene glycol) – 4000 (PEG-4000) was added in a 10% w/w ratio and then a 20% w/w ratio.

DNA Sequences of Interest. The list of DNA sequences of interest within this thesis is presented in Table 2. The 8mer control, 8mer U4, and 8mer FU4 were used within the non-palindromic study presented within the first part of the Results section. These sequences are 8 bases long and the numbering begins at the 5′ end of the top strand (G1) and continues to the end of the strand (C8). The bottom strand is numbered from the 5′ end (G9) to the end of the strand (C16). The top strand is referenced in this thesis as "front-end" and the bottom strand as "back-end." As can be seen in Table 2, the 8mer FU4 is the same sequence as the 8mer U4. The only

difference between the two is that the 8mer FU4 sequence was ordered with a fluorine at the

arabino position of the sugar. This position can be seen in Figure 12.

Table 1. List of Buffer Components for the Sample Conditions Project.

Buffer #	Components	Final pH
1	25 mM phosphate 100 mM sodium chloride	7.42
2	25 mM phosphate 100 mM potassium chloride	7.41
3	25 mM phosphate 100 mM sodium chloride 5 mM magnesium chloride	7.39
4	25 mM phosphate 100 mM sodium chloride 20 mM magnesium chloride	7.39
5	25 mM phosphate 200 mM sodium chloride	7.40
6	25 mM phosphate 300 mM sodium chloride	7.40

The use of this fluorinated uracil nucleotide is to allow for complexation with TDG

without cleavage, first determined in 1997 by Schärer et al.[7] The first complexed MUG enzyme

with a substrate made use of this fluorine method, allowing for a crystal structure of a hydrolytic

DNA glycosylase-substrate complex.[28]

The Dickerson-Drew Dodecamer (DDD) and T9 were used within the sample conditions project. The numbering scheme for DDD and T9 are as described for the 8mer, but because of the palindromic nature of the DNA duplex, only the top strand is numbered (C1 through G12). It should be noted that the 8mer project uses non-palindromic DNA while the sample conditions project uses palindromic DNA.
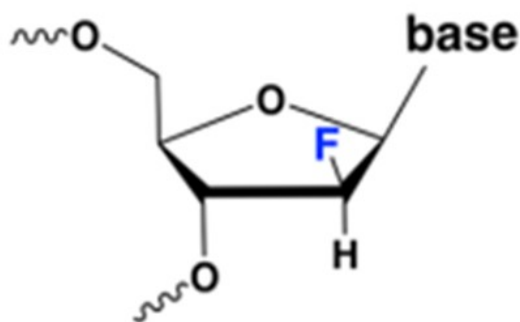
Figure 12. Image of fluorine at the arabino position of a deoxyribose molecule. Figure was reprinted (adapted) with permission from Dow, B. J.; Malik, S. S.; Drohat, A. C. Defining the Role of Nucleotide Flipping in Enzyme Specificity Using ^{19}F NMR. *J. Am. Chem. Soc.* **2019**, *141*, 4952-4962. Copyright 2019 American Chemical Society.

NMR Experiments

All NMR experiments were performed on a 400 MHz Varian Inova NMR instrument within the Department of Chemistry and Biochemistry at Missouri State University using VnmrJ software. After loading and setting the parameters of the relevant NMR experiments, the instrument was shimmed and tuned to the relevant nuclei. Experiment temperatures varied between 278 K and 298 K, depending on the project and relevance to the experiment progression. The temperature is noted for each data set and figure within this thesis. Spin was off for all 1D and 2D experiments, unless otherwise noted. The following four subsections provide

descriptions of the general parameters used for each NMR experiment. The data presented in the

Results section followed these parameters unless otherwise noted.

Table 2. DNA Sequences of Interest.

Abbreviation	DNA Sequence
8mer control	5′—G—A—G—C—G—T—T—C—3′ ||| || ||| ||| ||| || || ||| 3′—C—T—C—G—C—A—A—G—5′
8mer U4	5′—G—A—G—U—G—T—T—C—3′ ||| || ||| || ||| || || ||| 3′—C—T—C—G—C—A—A—G—5′
8mer FU4	5′—G—A—G—U—G—T—T—C—3′ ||| || ||| || ||| || || ||| 3′—C—T—C—G—C—A—A—G—5′
Dickerson-Drew Dodecamer (DDD)	5′—C—G—C—G—A—A—T—T—C—G—C—G—3′ ||| ||| ||| ||| || || || || ||| ||| ||| ||| 3′—G—C—G—C—T—T—A—A—G—C—G—C—5′
T9	5′—C—G—C—G—A—A—T—T—T—G—C—G—3′ ||| ||| ||| || || || || || || ||| ||| ||| 3′—G—C—G—T—T—T—A—A—G—C—G—C—5′

^{1}H-^{1}H NOESY. 64 scans per t1 increment were typically used (128 when time allowed), with 200 t1 increments. The spectral width was 4001.6 Hz and complex points 2048. The offset within the channels was set to -150.00 Hz.

¹H-³¹P HSQC. 64 scans per t1 increment were typically used (128 when time allowed), with 256 t1 increments. A one-bond J1xh value of 10.0 Hz was used. The spectral width in F2 was 4001.6 with 2048 complex points and a relaxation delay of 1.00 s. The spectral width in F1 was set to 1950.0 Hz. An X pulse width of 16 µs was used with a X pulse power of 56 dB, and a X offset of 4250 Hz. The offset for ¹H was -150.00 Hz and the decouple offset for ³¹P was 4250 Hz.

1D ³¹P. 1024 scans per t1 increment were typically used (4096 when time allowed) with a pulse angle of 90 degrees. A relaxation delay of 1s was used for temperatures greater than 288K and a relaxation delay of 2s was used for experiments completed at 288K or lower. The spectral width was 6000.6 Hz with 16384 complex points. The ³¹P offset was set to 3988.00 Hz. To reference the 1D ³¹P data, a coaxial insert was used. This insert contained 85% phosphoric acid. The sample to be referenced was transferred to a thin – walled NMR tube and the insert added inside to this tube. The 1D ³¹P experiments used for the coax samples used the same parameters as the general 1D ³¹P experiments outlined above.

1D ¹⁹F. The flourine data presented in this paper has the associated parameters used. Rather than autotuning for fluorine, the instrument was manually tuned by the user.

¹H-¹H NOESY Sequential "Walk" Method for Proton Assignments

This section of the methods provides details on how a NOESY sequential "walk" is performed to assign the protons of a DNA sequence. The chemical shifts obtained from a NOESY have an error of ± 0.01 ppm due to the resolution of spectrum.

A ¹H wet 1D experiment is not feasible to do to make specific peak assignments because of the sheer magnitude of protons within unique chemical environments within these short

duplex DNA sequences. Figure 13 shows a ^{1}H wet 1D spectrum of the 8mer U4 sequence with

two regions labeled of where particular protons show up. As can be seen from Figure 13, it is not

feasible to make specific base and proton assignments based on the 1D spectrum due to peak

overlap.
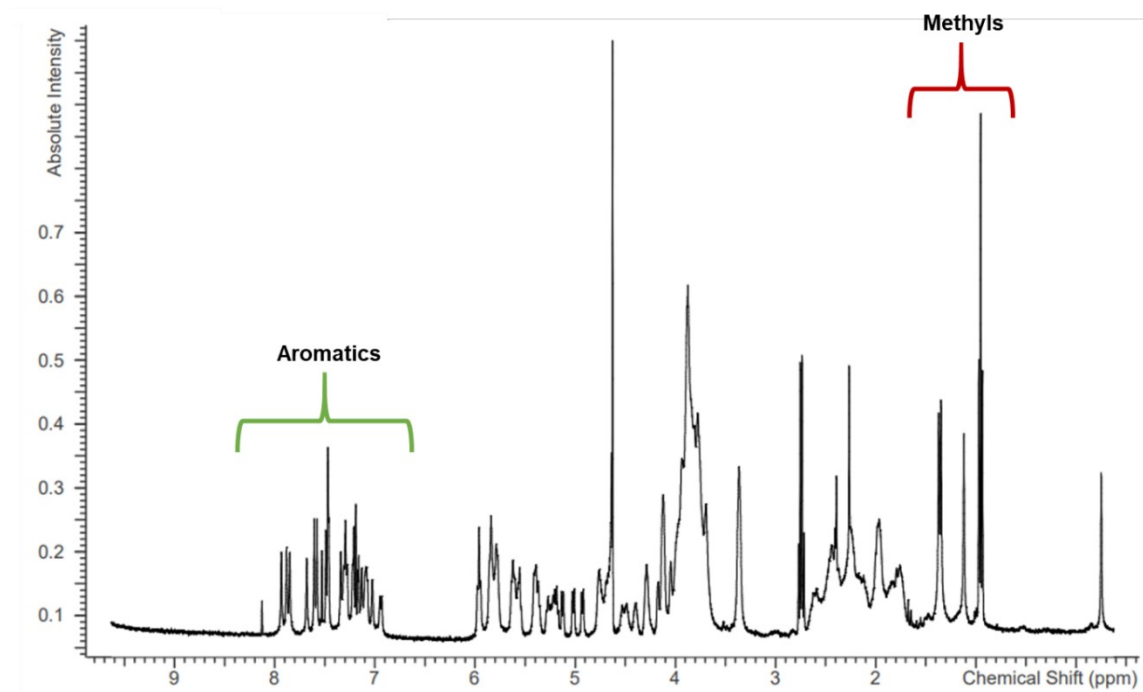
Figure 13. 8mer U4 283 K wet 1D ^{1}H spectrum with aromatic and methyl proton regions labeled.

To mitigate this issue, a 2D ^{1}H-^{1}H NOESY experiment is used instead. Each peak on a

2D NMR spectrum represents two nuclei interacting with each other. The NOESY spectrum for

a DNA sequence follows general trends in where the proton-proton interactions appear, just like

how there are certain regions within a ^{1}H wet 1D spectrum as described in Figure 13. Figure 14

provides an example of what a DNA NOESY spectrum looks like, with the proton interaction

regions of interest for this project labeled within boxes. For a ^{1}H-^{1}H NOESY, the nuclei can be

up to 5 Å apart from each other.[23]

The water peak for the NOESY spectrum is referenced according to the temperature that the experiment was performed at based on Gottlieb et al.[29] For example, at 298 K the water peak is referenced to 4.77 ppm and at 283 K it is 4.94 ppm.

The H1′, H2′, and H2″ protons are typically where a person should start in order to assign a NOESY, as these protons can be assigned using the sequential "walk" method. The NOESY walk has been compared to a sudoku puzzle, as the peaks connect in a particular way, unique to each DNA sequence and combination of bases. Within these regions, there can be general trends of where the different bases typically show up that can help in the beginning of assignments.

To assign a NOESY spectrum, it is common to begin with the thymine nucleotides within the DNA sequence. Thymine contains a methyl group as seen in Figure 1, which does not appear on any of the other bases. These methyl protons appear in a distinct area on a 1D and 2D spectrum, as shown in Figure 13, and can tell you where the thymines are located in the H2′/H2″ region and H1′ region. The thymines within the DNA sequence can then be assigned based on what they are next to.

The H2′/H2″ and H6/H8 region of the 8mer control at 283 K can be found in Figure 15. As can be seen, there is a methyl peak in Figure 15 (marked with a dark red box) that aligns with a peak to the left within the guanine region. There is only one thymine within 8mer control that has a thymine that interacts with a guanine (T6). Thus the guanine is G5. The thymine peak within the dark red box along the F1 axis is the methyl proton of T6. The value along the F2 axis is the H6 proton of T6. The value along the F2 axis of the G5 peak is the H8 proton of G5.

Following the H8 proton of the G5 base down to the H1' region can begin the walk of the H1' region. A representation of this process can be found in Figure 16. There are two peaks for G5 along this axis: one lines up to the right with a peak in the uracil/cytosine/thymine region,

while the other peak lines up with a peak along the T6 H6 proton assignment (yellow box) determined in the H2′/H2″ region. Since the base on the other side of G5 is C4, a walk can be completed between the G5 H1′ proton (in the light green box) to the C4 H1′ proton in the darker green box. Figure 16 shows how this walk looks on the DNA sequence as well.
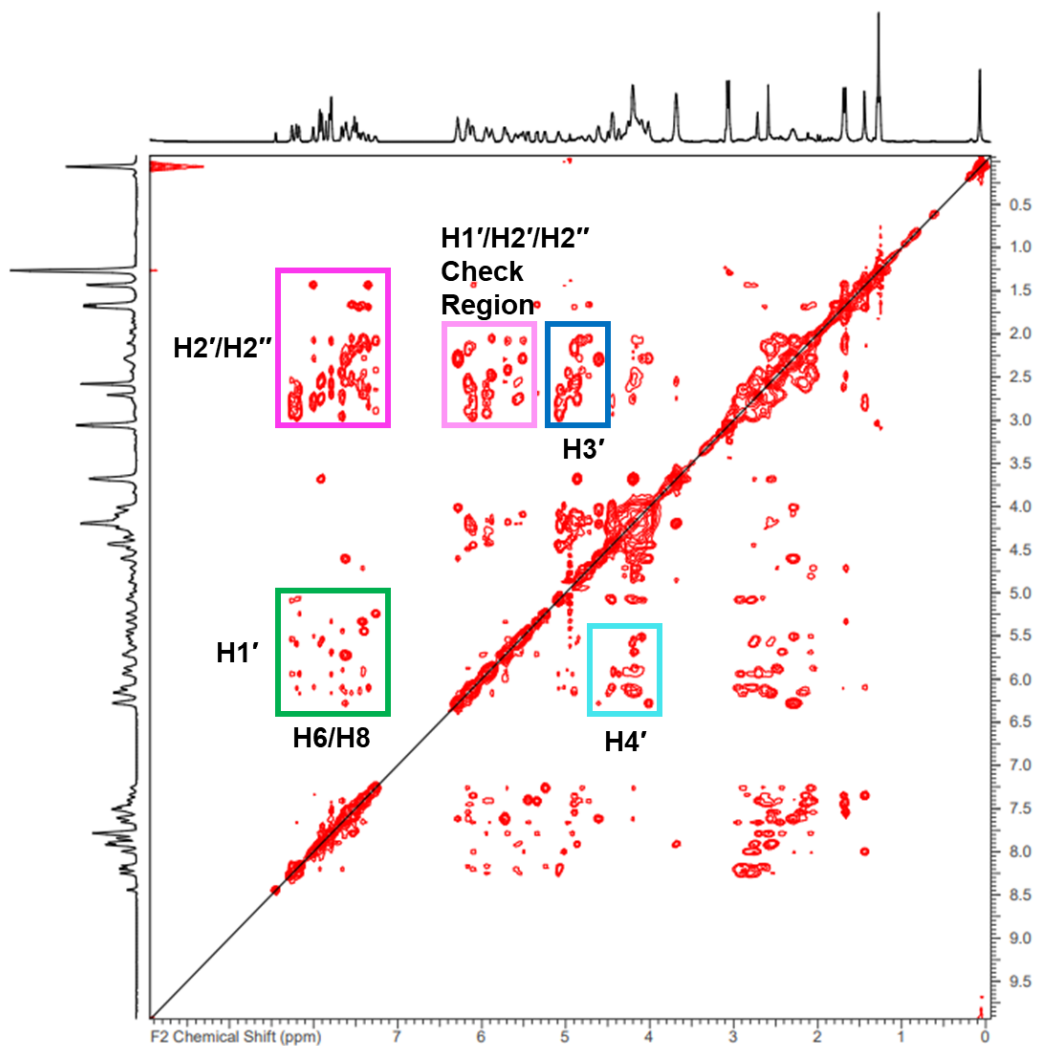
Figure 14. 8mer U4 ^{1}H-^{1}H NOESY spectrum at 283 K. The regions in boxes are the primary regions to assign the sugar and base protons of a DNA sequence. The H2′/H2″ to H6/H8 region is marked in dark pink. The H2′/H2″ to H3′ region is marked in dark blue. The H1′ to H6/H8 region is marked in green. The H1′ to H4′ region is marked in light blue. And the H1′/H2′/H2″ check region is in light pink.
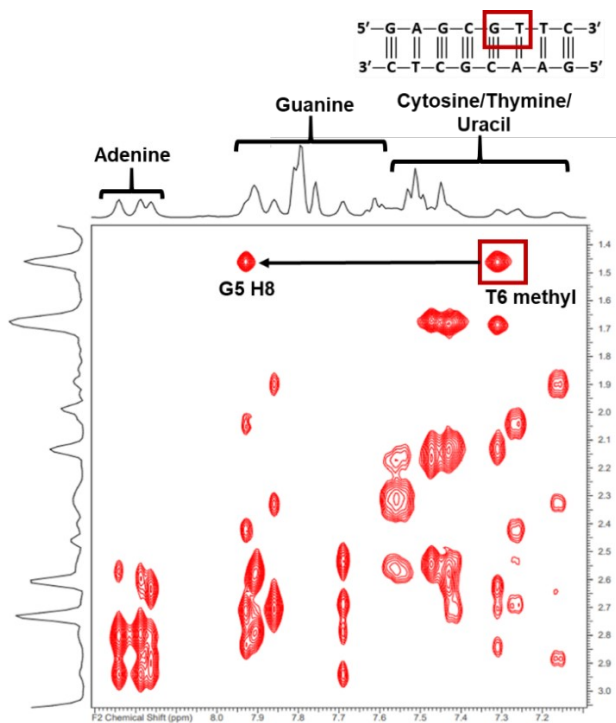
Figure 15. H2′/H2″ to H6/H8 region of the 8mer control ^{1}H-^{1}H NOESY spectrum at 283 K.
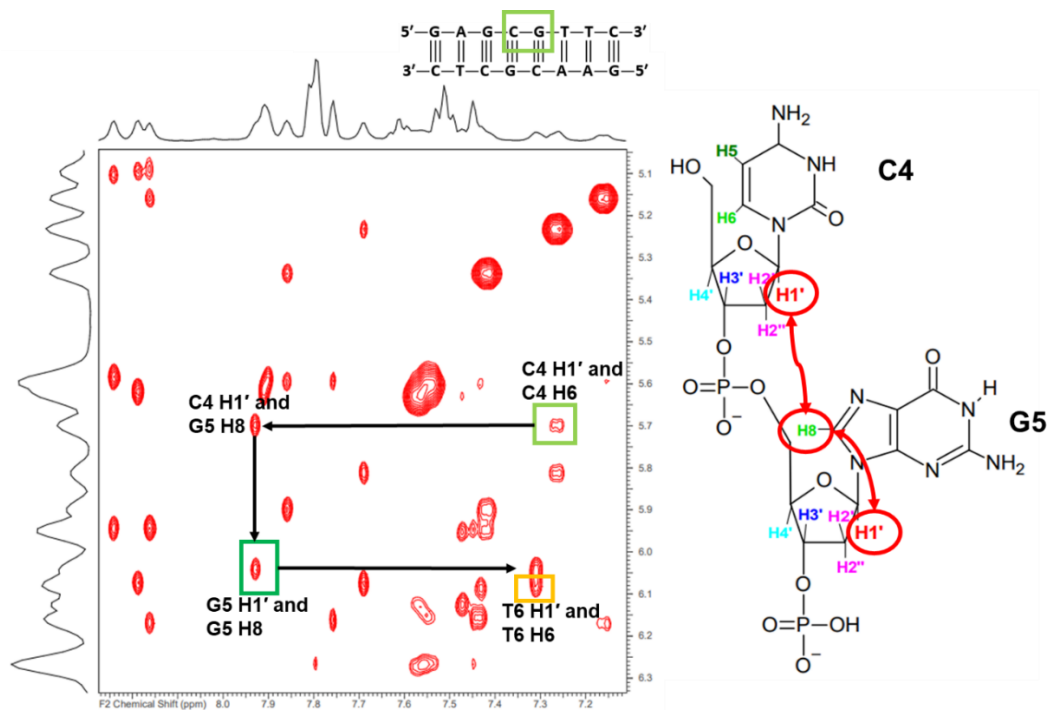
Figure 16. Example of a C-G walk in the H1′ region of the 8mer control sequence at 283 K.

This process is continued in the H1′ region to ultimately obtain a "walk" for all of the H1′ proton assignments. A representation of a completed H1′ walk for the 8mer control at 283 K can be found in Figure 17. Each side of the sequence (the front-end and back-end) have their own walks amongst each base. This is particularly true for the non-palindromic sequences as the front-end and the back-end walks on the NOESY spectrum do not connect to each other.
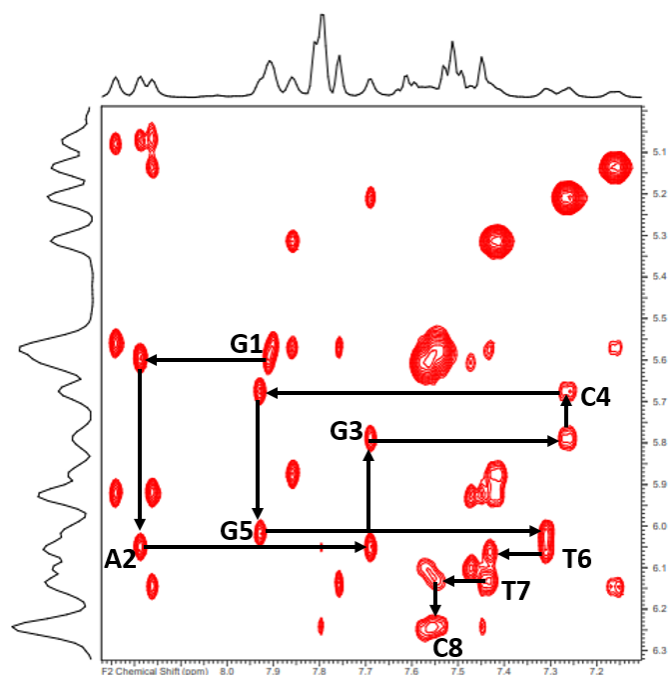
Figure 17. Representative H1′ front-end walk of the 8mer control at 283 K.

This process is repeated in the H2′/H2″ region where two sets of walks (H2″ and occassionally H2′) occur. To check that the H2′/H2″ and H1′ assignments are correct, the H1′/H2′/H2″ check region is used (shown in Figure 14 in the light pink box). This region houses the H2′ and H2″ protons along the F1 axis while the H1′ assignment of the same base is along the

F2 axis. An example of the C4 H1′/H2′/H2″ protons that were checked in this region for the 8mer control NOESY can be found in Figure 18.
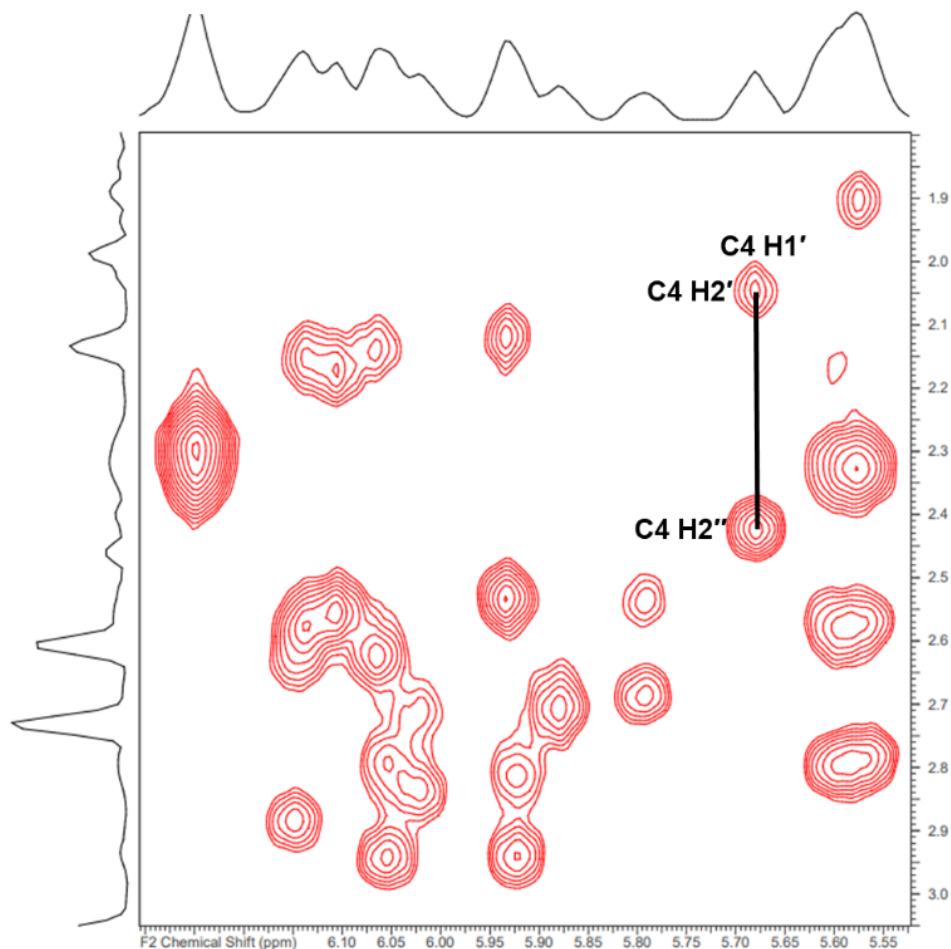
Figure 18. 8mer control H1′/H2′/H2″ check region at 283 K with the indicated H1′/H2′/H2″ interactions for the C4 base assigned from the H1′ and H2′/H2″ regions.

After confirming the H2′/H2″ assignments are correct in the check region, these protons can be followed along the horizontal to the H3′ region, as shown in the dark blue box in Figure 14. This allows for the subsequent assignments of the H3′ protons. Figure 19 presents the H3′ region for the 8mer control with how the H3′ of C4 was assigned.

Finally the H1′ protons for each base on the sequence can then be used to assign the H4′ protons, the vital assignments within the NOESY spectrum. An example of this region with the C4 H4′ assignment indicated for 8mer control at 283 K can be found in Figure 20. The H4′ proton chemical shift assignments can then be used on the ^{1}H-^{31}P HSQC spectrum.
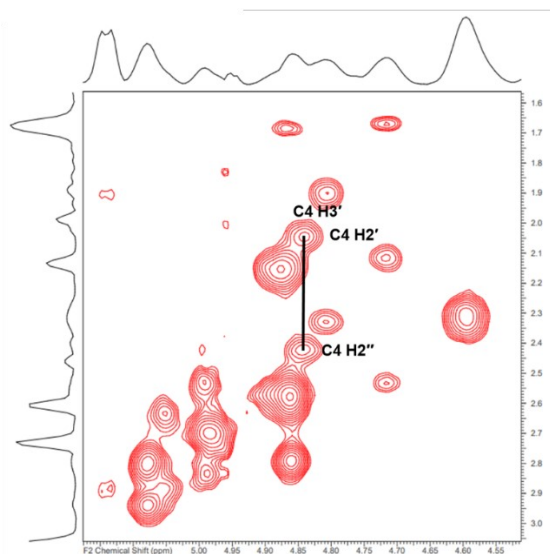
Figure 19. H3′ region of 8mer control at 283 K with the C4 H3′ assignment indicated based on its H2′ and H2″ assignments.
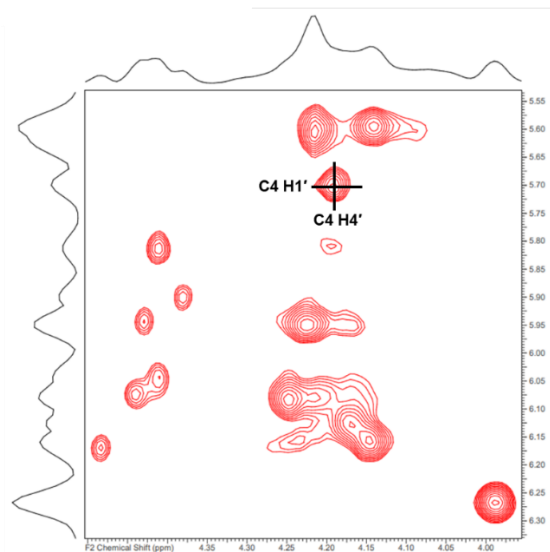
Figure 20. H4′ region of 8mer control at 283 K with the C4 H4′ assignment indicated based on its H1′ assignment.

Assignment of ^{1}H-^{31}P HSQC and ^{31}P Spectra

A ^{1}H-^{31}P HSQC contains signficantly fewer peaks than a NOESY spectrum. There are only two regions within an HSQC spectrum, as compared to a NOESY spectrum. One region is the interaction of the H3′ proton of a nucleotide and its 3′ phosphate. Due to poor resolution within this region and the difficulties that can arise in assigning the H3′ protons, the other region is relied upon more. The other region is the interaction of the H4′ proton of a nucleotide with its 5′ phosphate.

Along the vertical (F2) of the spectrum is the ^{1}H dimension. The water peak is referenced in this F2 dimension as described for a NOESY spectrum. To reference the horizontal ^{31}P dimension, a coax experiment is required. Figure 21 shows an example of a coaxial experiment for DDD in 25 mM phosphate and 100 mM KCl buffer at 298 K. The coax has a very large phosphoric acid peak that is referenced to 0.00 ppm. To the right of this peak are a set of smaller peaks which represent the 1D ^{31}P spectrum for that same sample. The chemical shift of the right most peak is recorded and used to reference the ^{1}H-^{31}P HSQC and 1D ^{31}P spectrum. On the ^{1}H-^{31}P HSQC spectrum, the top most peak is referenced to the value recorded from the coax. The most right peak of the 1D ^{31}P spectrum is referenced to the value recorded from the coax.

After properly referencing the spectrum and obtaining the referenced H4' proton chemical shift values from the NOESY as described in the "^{1}H-^{1}H NOESY Sequential 'Walk' Method for Proton Assignments" subsection within the Methods section, the phosphorus atoms within the phosphodiester backbone of the DNA sequence of interest can be assigned. An example of an assigned HSQC spectrum can be found in Figure 22. Each phosphate step is labeled as XpY, with the X being the base 5′ to the phosphate and Y the base 3′ to the phosphate.
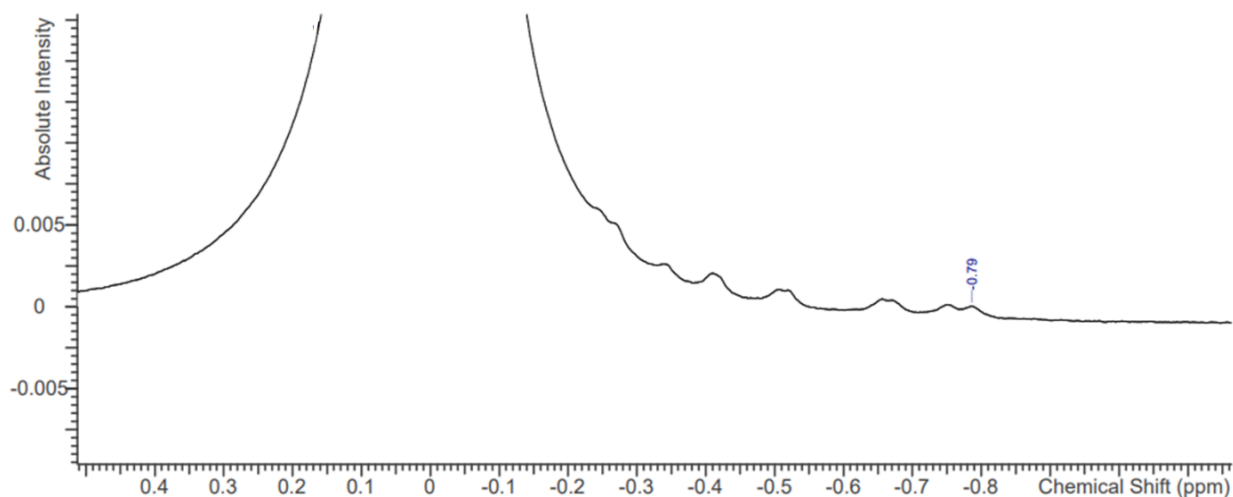
Figure 21. 298 K DDD in 25 mM phosphate and 100 mM KCl buffer coax experiment spectrum. The large peak is the phosphoric acid referenced to 0.00 ppm and the right most peak chemical shift was recorded.
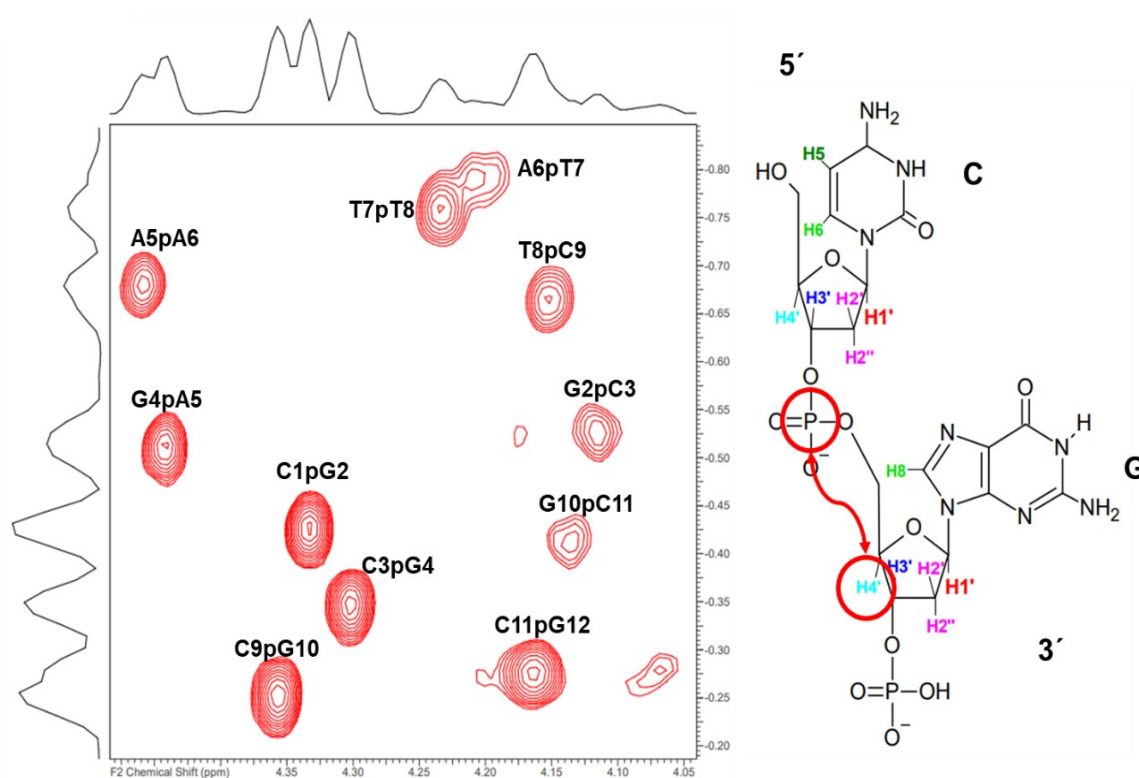
Figure 22. Representative ^{1}H-^{31}P HSQC spectrum of DDD in 25 mM phosphate and 100 mM KCl at 298 K with phosphate assignments. An example of the interaction between a phosphorus between a C and G base and the H4′ 3′ to the phosphorus is C9pG10. Notice how each peak is labeled in the XpY notation, with X representing the base 5′ to the phosphate and Y representing the base 3′ to the phosphate.

Due to the length of an HSQC experiment, over 15 hours for a 64 scan experiment at 283 K, it is more feasible to run 1D ^{31}P experiments to assign the phosphorus instead once the HSQC has been assigned. To reference a 1D ^{31}P spectrum, the value obtained from the coax is used to reference the right most peak of the 1D ^{31}P spectrum. Using the values of the HSQC, the 1D peaks can be assigned. Figure 23 shows a representative referenced 1D ^{31}P spectrum of DDD in 25 mM phosphate buffer and 100 mM KCl at 298 K. The chemical shifts from the HSQC have an error of $\pm$ 0.02 ppm due to the resolution of spectrum.
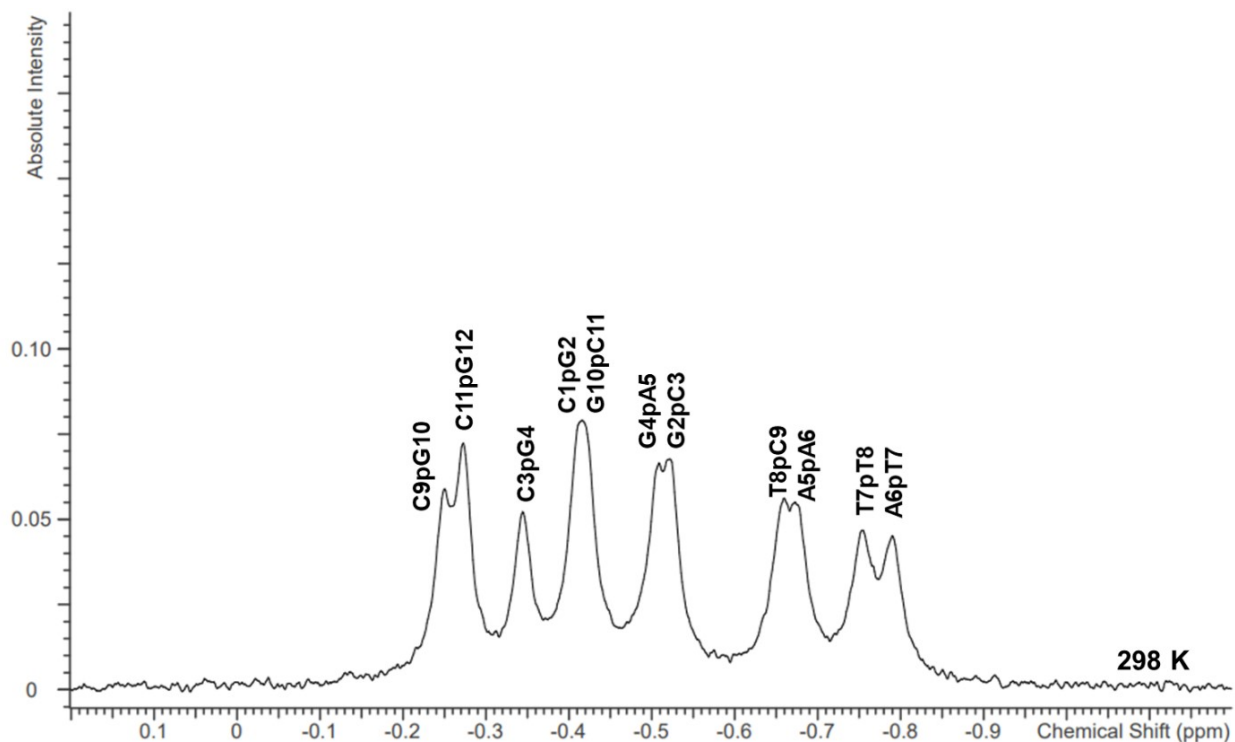
Figure 23. Representative assigned 1D ^{31}P spectrum for DDD in 25 mM phosphate and 100 mM KCl at 298 K. The right most peak was referenced via the coax experiment shown above in Figure 21.

Calculation of %BII and Energetics of the BI/BII Interconversion

Each referenced phosphorus chemical shift value from the 1D ^{31}P or ^{1}H-^{31}P HSQC spectrum can then be used to calculate %BII. As discussed in the Introduction subsection "BI/BII Conformations," the interconversion of BI and BII is considered to be within the fast exchange regime of the NMR time-scale, measured to be $\gtrsim 10^4\,\text{Hz}$[22] at the temperature range used. This results in each peak on the 1D ^{31}P spectrum being a weighted average of the two primary backbone conformations.[22,27]

The %BII can thus be calculated using Equation 1 from Westwood et al. 2021[22] below. This equation was modified from a previous equation presented by Tian et al. 2009.[27] The %BII is dependent on the referenced phosphorus chemical shift value in ppm (δP) and the temperature (T) of the experiment completed in K.

$$\text{Equation 1: } \%\text{BII} = \frac{100}{4.01875 - 0.008833\text{T}}\,\delta\text{P} + \frac{285.8 - 0.6\text{T}}{4.01875 - 0.008833\text{T}}$$

This value can then be subtracted from 100 to obtain %BI. %BII and %BI can then be used to calculate the equilibrium constant, K. The equilibrium constant can then be used to calculate the free energy, ΔG, of this conformational equilibrium as seen in Equation 2,[22] where R is ideal gas constant and T is the temperature in K. Analysis of the change in ΔG across sequences and sample conditions of this exchange process can show if there is a perturbation to the phosphodiester backbone conformational equilbrium dynamics.

$$\text{Equation 2: } \Delta\text{G} = -\text{RTln}(\text{K}) = -\text{RTln}\left(\frac{\%\text{BII}}{\%\text{BI}}\right)$$

Determined Protocol of the Expression and Purification of Thymine DNA Glycosylase

The expression and purification of Thymine DNA Glycosylase (TDG) revolved around optimization of the procedure. The most promising protocol for hTDG expression and purification that was determined in this study is presented within this section. The second subsection of the Results section details important aspects of how this protocol was determined.

TDG Plasmid Preparation. A pET28a based plasmid for *Escherichia coli* containing a TDG gene and kanamycin antibiotic-resistant gene was purchased from TWIST Bioscience. The construct this gene expressed for was TDG$^{82-308}$ with an N-terminal six histidine tag with a TEV protease cleavage site. A stock TDG plasmid was made to an end concentration of 20 ng/μL by adding autoclaved DI water to the plasmid gene tube. When not being used, the stock plasmid preparation was stored in the freezer at -20°C.

As the project went on, a plasmid preparation was completed to have more plasmid sample available. 2 mL of an overnight expression culture was used for this plasmid preparation and centrifuged at 12,000 rpm for 30 sec. The supernatant was discarded. This was repeated once more. To purify the plasmid, a ThermoScientific GeneJET Plasmid Mini Prep Kit was used. 250 μL of resuspension buffer was added to the pelleted cell from the previous step, then 250 μL of lysis solution was added. The tube was then inverted 4-6 times. 350 μL of neutralization solution was added, and the tube was inverted 4-6 times. The sample was then centrifuged for five minutes. The supernatant after centrifugation was then transferred to a Thermo Scientific GeneJET Spin Column and 500 μL of wash solution was added to the column. This column was centrifuged for 30-60 seconds. The flow through was discarded and the process of adding wash solution and centrifuging was repeated once more. The empty column was centrifuged for one minute and the column was then transferred to a new tube. 50 μL of autoclaved DI water was added to the column and incubated approximately 1.5 minutes. The sample was then centrifuged

and the resulting solution was the TDG plasmid preparation. This TDG plasmid preparation was also stored in the freezer at -20°C.

TDG Transformation Protocol. 1-2 microliters of the prepared TDG plasmid were added to BL21 competent *E. coli* cells. This mixture was put on ice for thirty minutes, then heat shocked for 30 seconds at 37°C in a hot water bath. Following the heat shock, the mixture was immediately put back on ice for two minutes. An outgrowth step was then completed by adding 500 microliters of LB or NZYB media to the mixture and incubating at 37°C for one hour. The sample was then centrifuged at 12,000 rpm for 30 seconds followed by a removal of 450 microliters of supernatant. The pellet was resuspended in the leftover solution by gently pipetting up and down. This mixture was then streaked on an LB-kanamycin agar plate and put into the incubator at 37°C for overnight. A figure of a representative plate from this transformation protocol can be found in Figure 24.
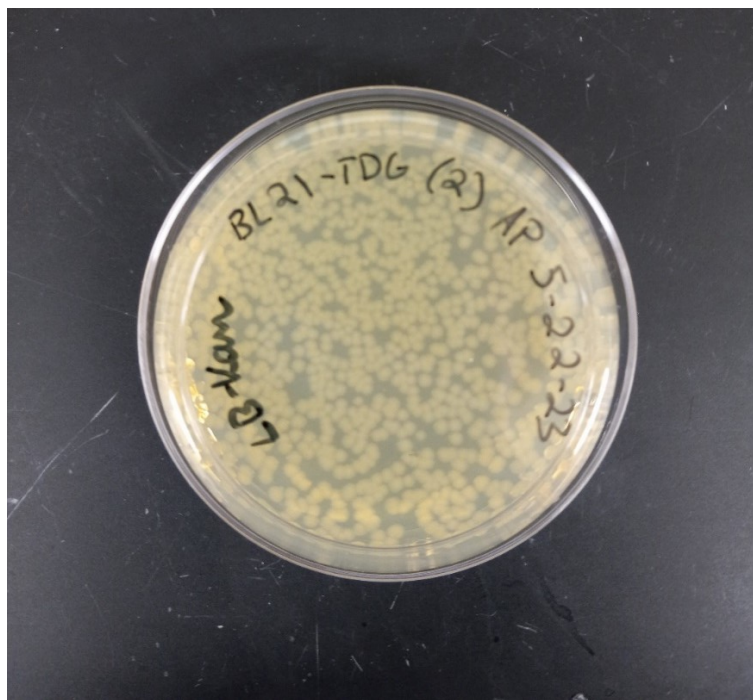
Figure 24. BL21 *E. coli* cells with TDG plasmid streaked on an LB-kanamycin agar plate.

TDG Expression and Lysing of Cells Protocol. LB media was prepared by adding 2.5 g of LB Broth, Miller to a 250 mL erlenmeyer flask with 100 mL DI water. Four 500 mL batches of LB media were also prepared for the scale up step (made by dissolving 12.5 g of LB Broth, Miller in 500 mL DI water in four separate flasks). These LB media preparations were autoclaved and then were cooled to room temperature. 100 microliters of a 50 mg/mL kanamycin solution was added to the 100 mL flask and 500 microliters of the 50 mg/mL kanamycin solution was added to each of the 500 mL LB media flasks. A colony was plucked from the plate using a pipette tip and added to the 100 mL LB media flask (colony and whole pipette tip). The flask was placed in the incubator with shaking at around 190 rpm for overnight at 37°C. Figure 25 shows what this step of adding a colony to the media looks like before incubation and after overnight incubation.
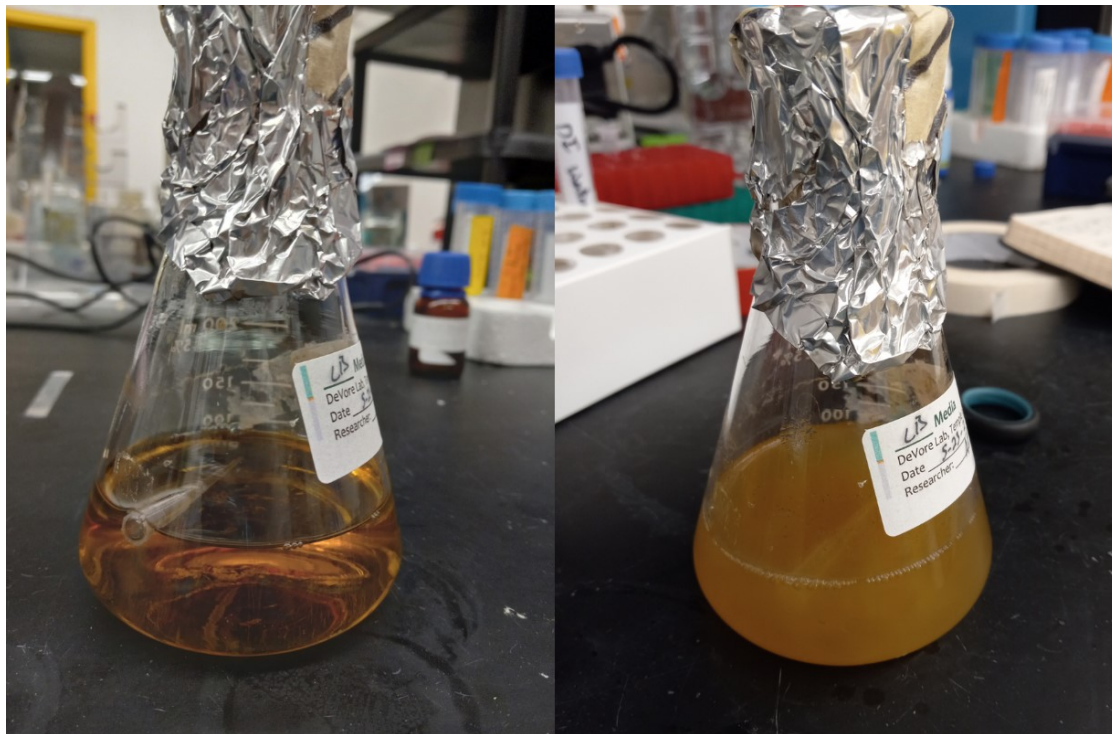
Figure 25. Colony and pipette tip added to LB media before incubation (left) and the result of the overnight culture (right).

The following day approximately 20 mL of overnight culture was added to each of the 500 mL media batch flasks, and the flasks were incubated with shaking at 37°C. The optical density (OD) was checked using a UV-Vis spectrophotometer, with the wavelength set to 600 nm, and the instrument was autozeroed with autoclaved LB media. The approximate OD at induction was 0.8 and 0.5 mM Isopropyl-beta-D-thiogalactoside (IPTG) was used to induce. The samples were placed into the cold room incubator at 15°C for overnight with shaking at ≈160 rpm.

After incubating at 15°C for overnight, the cells were centrifuged at 5,000 rpm for ten minutes. The cell pellets were resuspended in lysis buffer by adding 25 mL lysis to one of the containers with the cell pellet and vortexed to mix. This mixture was then transferred to the next cell pellet container and vortexed. This process was completed until the last container. Once the last container was resuspended, the resulting resuspension was transferred to a sonication cup. This process was completed with the cell pellet containers with an additional 20 mL of lysis buffer, and the resulting mixture from the cell pellet containers was transferred to a second sonication cup. The lysis buffer was 0.3 M NaCl, 50 mM monobasic sodium phosphate anhydrous, 10 mM β-mercaptoethanol (BME), with a pH of 8.0. The samples were then sonicated at 80% amplitude in 30 second intervals for a total of 1.5 minutes for each sonication cup. The samples were then centrifuged at 20,000 rpm for twenty minutes. The subsequent lysate was often thick and syrupy in appearance and texture, so additional lysis buffer was typically added until the thickness had reduced in order to run on the Ni-NTA affinity column. Figure 26 shows a representative image of the isolated cell pellet from the first centrifugation after incubation. Figure 27 shows a representative image of the resulting lysate after the second centrifugation for twenty minutes with the rest of the cell parts pelleted.
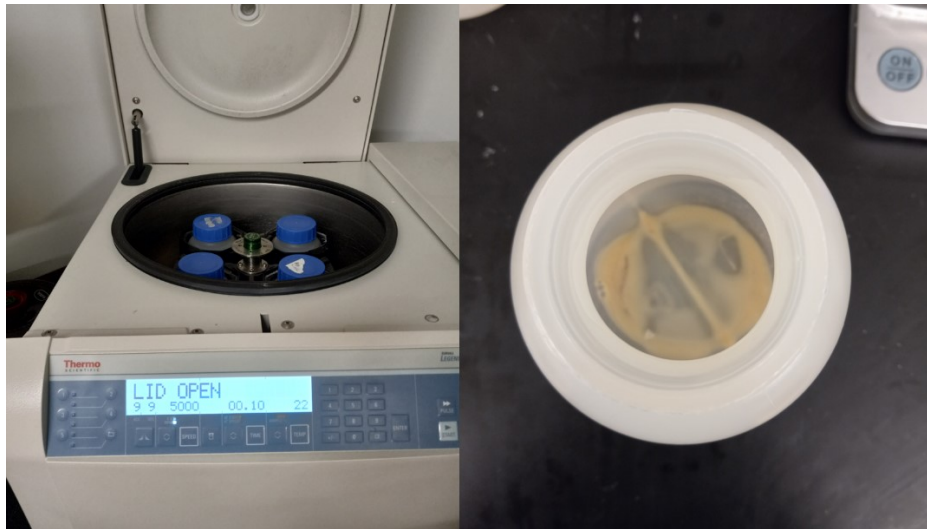
Figure 26. Centrifugation set up and resulting cell pellet after induction incubation.

Figure 27. Resulting lysate and cell parts pellet from the second centrifugation.

TDG Purification Protocol. After the lysate was isolated, a Ni-NTA affinity column was performed in the first step of purification. For the Ni-NTA column, the lysis buffer mentioned in the previous section was used, a wash buffer of 8 mM imidazole, 0.3 M NaCl, 50

mM monobasic sodium phosphate anhydrous, 10 mM β-mercaptoethanol (BME), and a pH of 7.4 was used, as well as an elution buffer of 50 mM monobasic sodium phosphate, 10 mM BME, 0.3 M NaCl, 150 mM Imidazole, and pH 7.4 was used. The Ni-NTA affinity column should be performed in the cold room due to the apparent temperature sensitivities of the enzyme with a flow rate for the lysate of 0.5 mL/min due to the thickness of the sample. After the Ni-NTA column, a SDS-PAGE (Sodium Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis) gel is completed on the elution fractions to determine what fractions contain TDG (the MW of the construct is approximately 26 kDa). After this is determined, a UV-Vis spectrum is taken of the elution fraction to determine the concentration. To obtain the UV-Vis spectrum, the wavelength range was set to 400 nm (start) to 250 nm (end) with a sampling interval of 1.0 nm. The "Go to Wavelength" value was 400 nm. The instrument was autozeroed with elution buffer used for the Ni-NTA column and baselined. The instrument was then autozeroed with the Ni-NTA column elution fraction(s). After autozeroing with the elution fraction(s), data was collected. The absorbance of the peak centered around 250-280 nm on the resulting spectrum was used to calculate the concentration via Beer's Law. The concentration is determined via Beer's Law in Equation 3, which is then used to determine the amount of Tev protease that must be added to digest the sample (to remove the 6 Histidine tag on the protein). Within the Beer's Law equation, A is absorbance, ε is the molar absorptivity of TDG (17420 M^{-1} cm^{-1}), b is the cuvette path length (1 cm), and c is the concentration in M. The Tev protease was then added to the sample to digest overnight in the cold room. Dialysis was also performed simultaneously with the digestion. The dialysis tubes that were used were "Slide-A-Lyzer Mini Dialysis Devices, 20 K MWCO, 2 mL, 25/PK" from ThermoScientific. The dialysis involved the use of the lysis buffer described above as well.

Equation 3: A = ϵbc

After the subsequent digestion and dialysis overnight, a Nickel gravity column was

performed. After the column, the flowthrough typically must be concentrated. At this point, the

SDS-PAGE gel that is performed after this column shows extremely faint bands at around 25,

50, and 75 kDa. Figure 28 shows a gel result after the first Ni-NTA column and of the

concentrated flowthrough from the second Ni column.
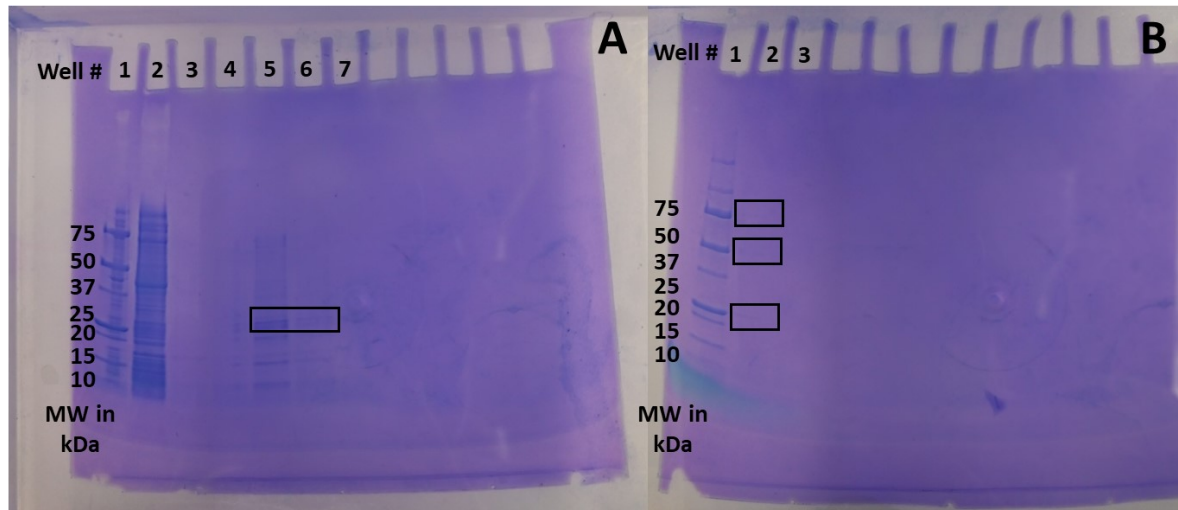
Figure 28. SDS-PAGE gel after the two Nickel columns. (A) SDS-PAGE gel after the first
Nickel column. Well 1 is the ladder and Well 5 and 6 are the elution fractions of interest
containing TDG (denoted within a black box). SDS-PAGE gel after concentrating the flow-
through from the gravity Nickel column (B). Well 1 has the ladder while Well 2 is the
concentrated flowthrough with very faint bands at around 25, 50, and 75 kDa (denoted with the
three black boxes).

RESULTS

The following chapter of results is split into three sections entailing the three projects investigated within my Master's program. The initial project that was continued from undergraduate was work on an 8mer non-palindromic DNA sequence containing a U:G mismatch, which can be found in the first subsection. The second subsection presents the results obtained from work on optimizing the expression and purification of TDG through a collaboration with Dr. Natasha DeVore at Missouri State University, Department of Chemistry and Biochemistry. Finally, the results of the investigation of the influence of sample conditions on palindromic DNA containing a T:G mismatch is presented in the final subsection of the Results.

8mer Non-Palindromic DNA Containing a U:G Mismatch

Prior work concerning phosphodiester backbone equilibrium dynamics of DNA analyzed the impact of lesions on dodecamer palindromic DNA sequences.[22] This research was expanded to non-palindromic DNA with the 8mer sequences: 8mer control, 8mer U4, and 8mer FU4 to analyze the impact of sequence context on the phosphodiester backbone conformational dynamics. The incomplete proton assignments for the 8mer control and 8mer U4 were presented in CHM/UHC 399/499 papers at 283 K and 298 K for the 8mer control and 283 and 285 K for the 8mer U4 within undergraduate. These values and walks are still included in Appendix A and B, respectively, for their relevance to this project. Within the Master's program, the proton assignments were completed at more temperatures for both sequences. In addition, the phosphorus assignments were completed at 278-298 K every 5 K for the 8mer control and 8mer U4.

8mer Control. The phosphorus for the 8mer control were assigned using ^{1}H-^{31}P HSQC and 1D ^{31}P experiments. Figure 29 and Figure 30 show the 293 K HSQC and stacked 1D ^{31}P spectra from 278-298 K for the 8mer control, respectively. Representative NOESY assignments at 283 K and tables of all proton chemical shifts completed for the 8mer control can be found in Appendix A: 8mer control Supplemental Data. The complete table of ^{31}P assignments referenced with a coax insert for the 8mer control can be found in Appendix A as well.
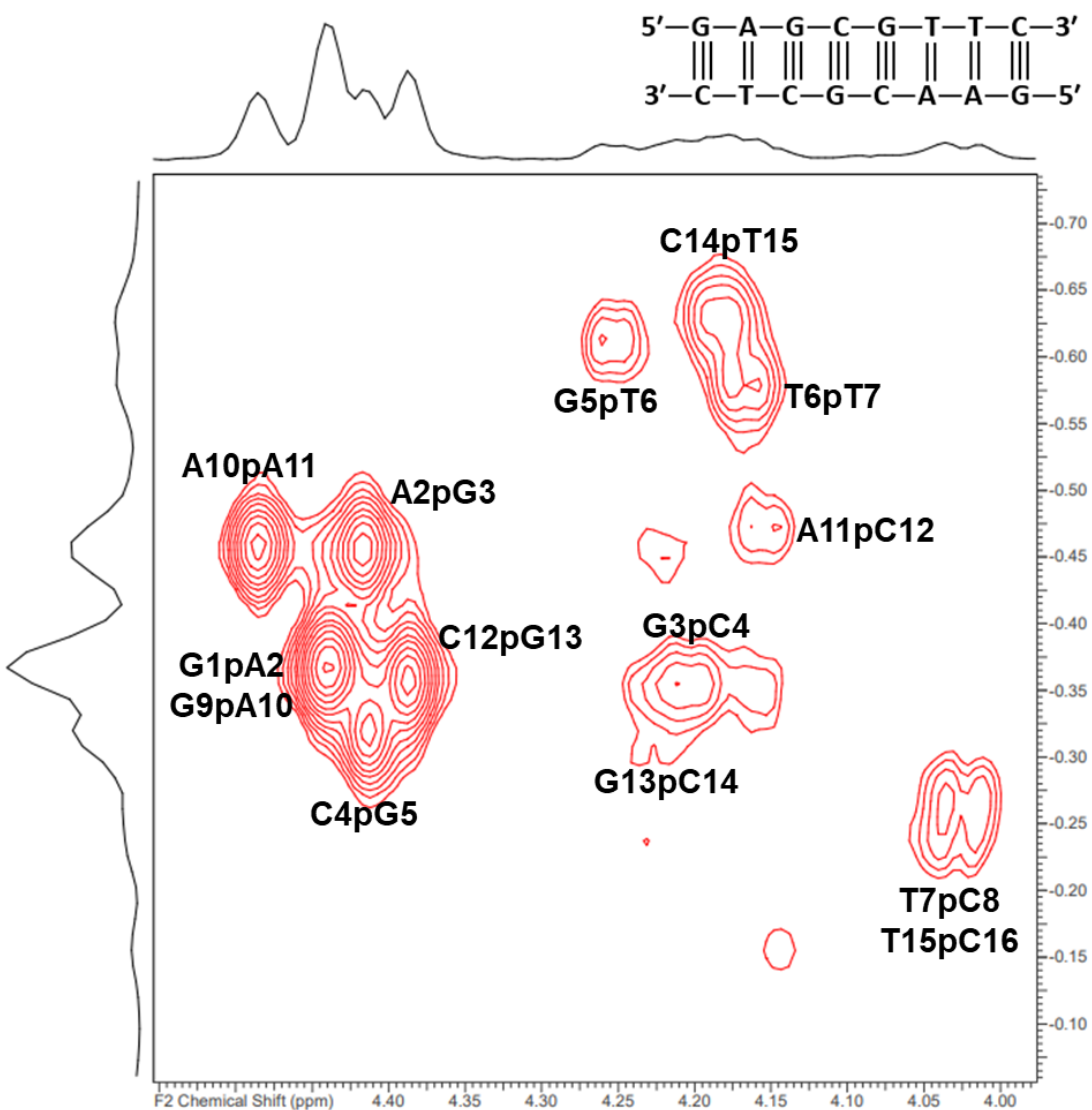
Figure 29. 8mer control HSQC at 293 K with phosphorus assignments.
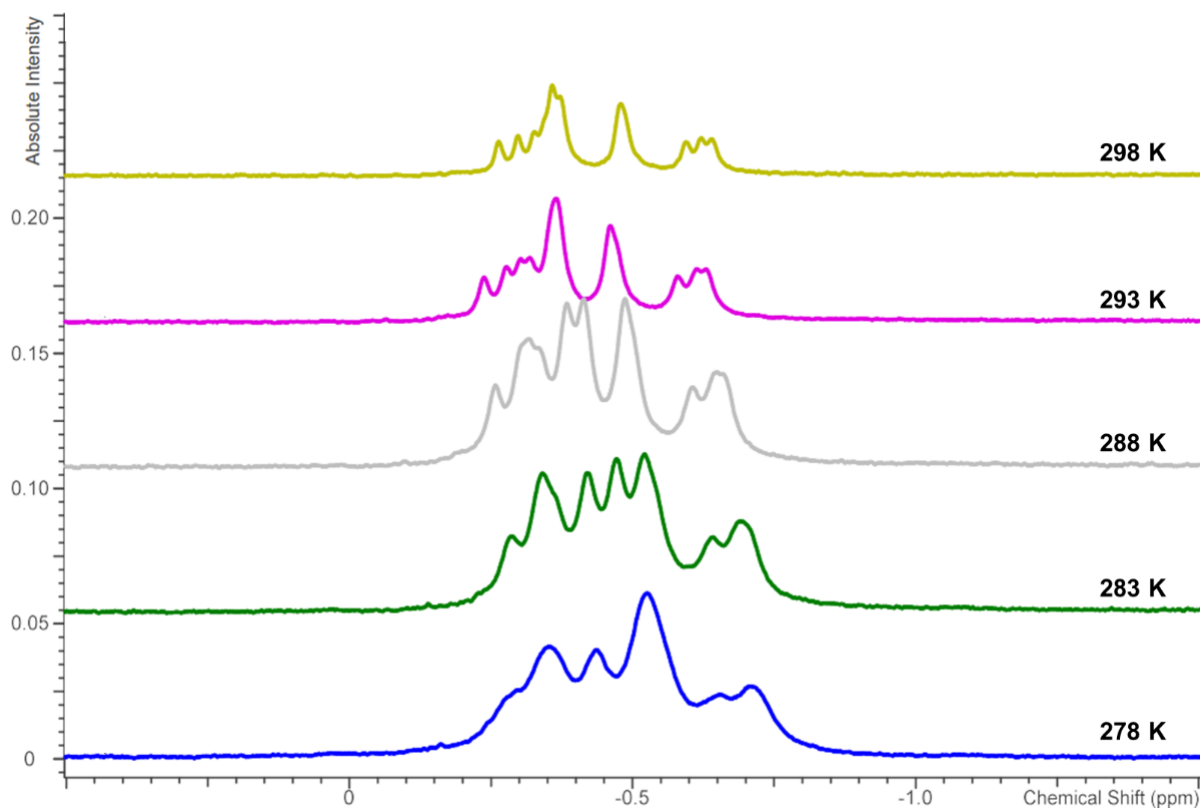
Figure 30. Stacked 1D ^{31}P spectra of the 8mer control from 278-298 K in 5 K increments.

From the ^{31}P chemical shift values, the %BII, equilibrium constant, K, and thus free energy, ΔG, at each phosphate step for the two conformational state equilibrium was calculated based on Equations 1 and 2. Charts of the %BII and ΔG for each phosphate step can be found below in Figure 31 A and B, respectively.

8mer U4. The phosphorus for the 8mer U4 were assigned using ^{1}H-^{31}P HSQC and 1D ^{31}P experiments. Figure 32 and Figure 33 show the 283 K HSQC and stacked 1D ^{31}P spectra from 278-298 K at 5 K for the 8mer U4, respectively. Representative NOESY assignments at 283 K and tables of proton chemical shift values obtained for 8mer U4 can be found in Appendix B: 8mer U4 Supplemental Data. The complete table of ^{31}P assignments referenced with a coax insert for the 8mer U4 can be found in Appendix B as well.
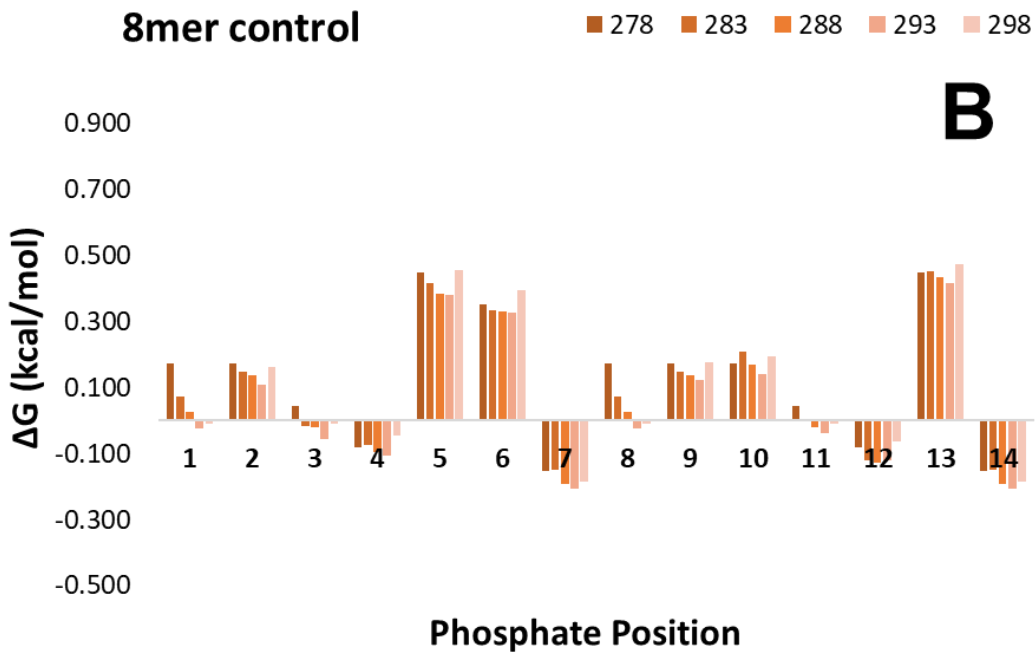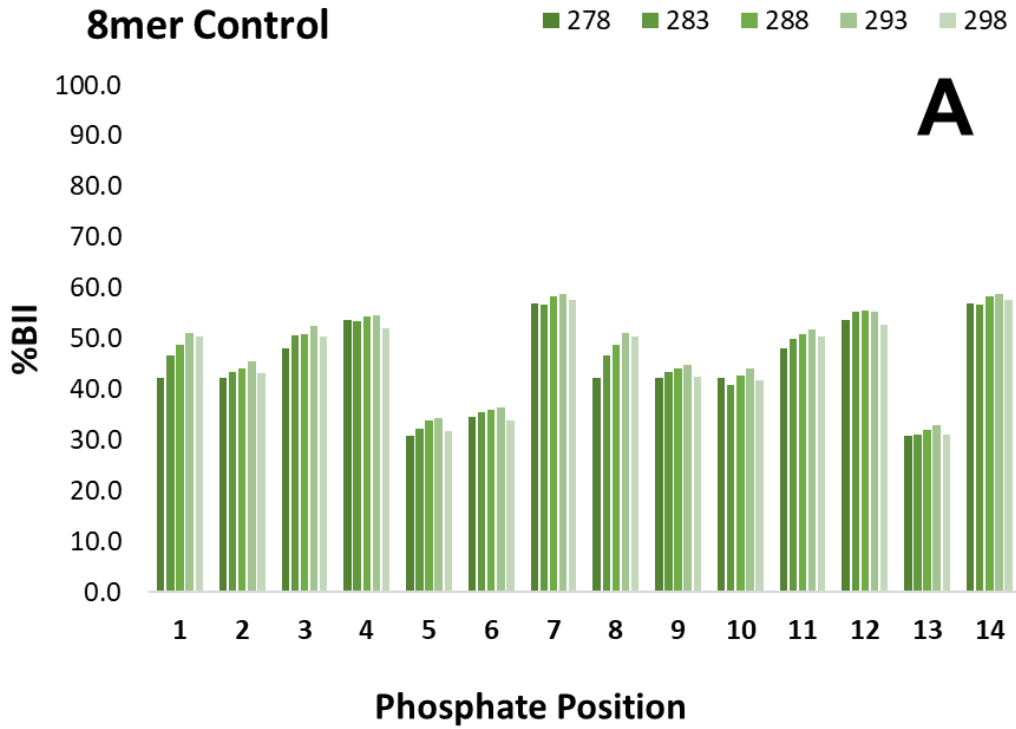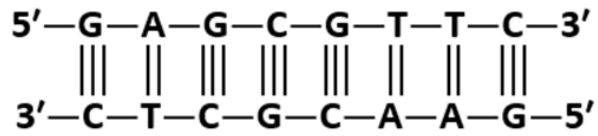
Figure 31. %BII (A) and ΔG (B) for the 8mer control at each phosphate position from 278-298 K.
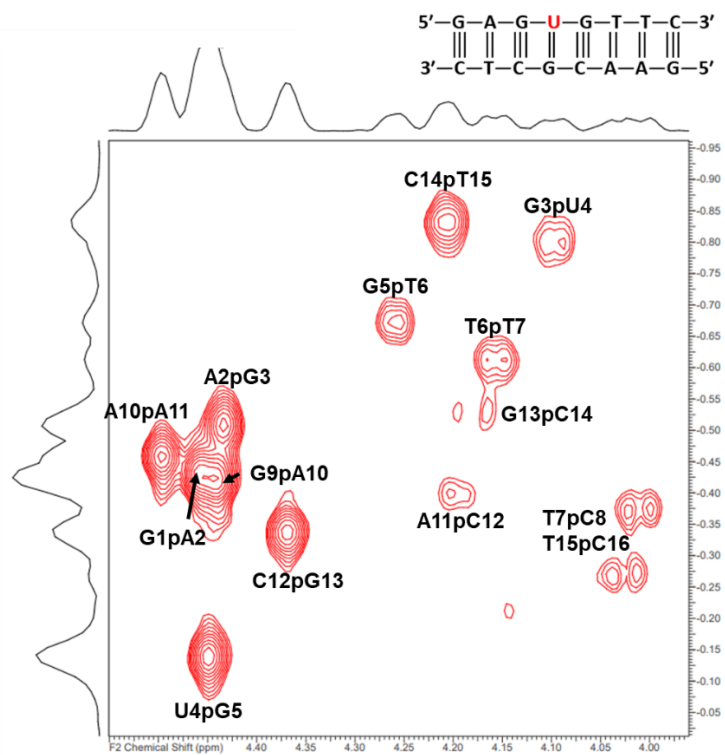
Figure 32. 8mer U4 HSQC at 283 K with phosphorus assignments.
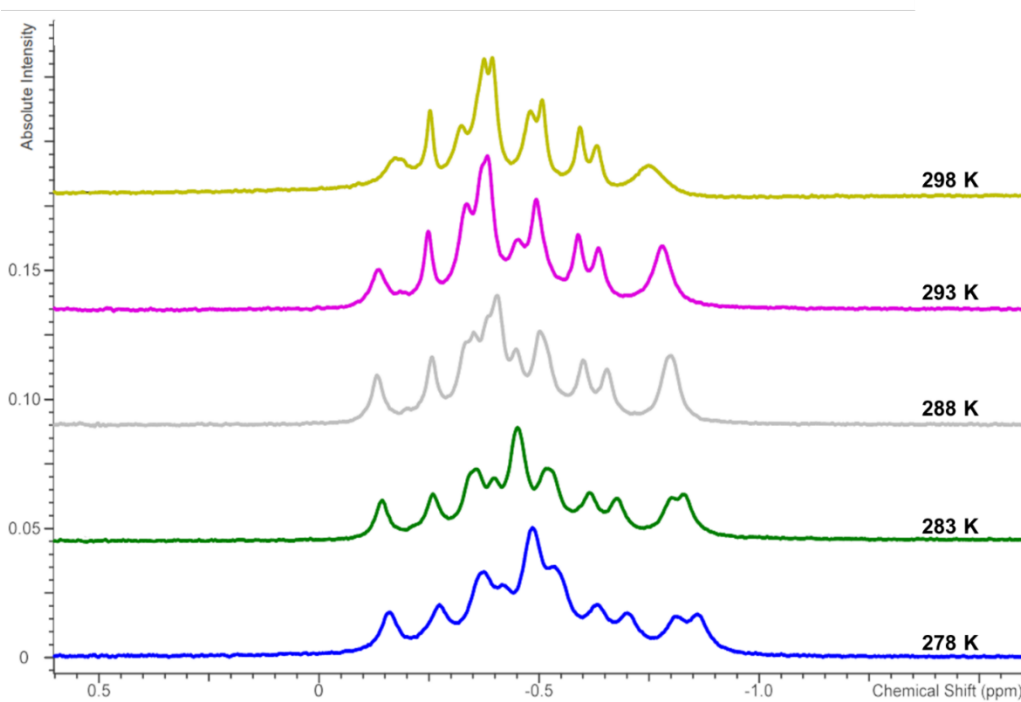
Figure 33. Stacked 1D ^{31}P spectra of the 8mer U4 from 278-298 K in 5 K increments.

From the ^{31}P chemical shift values, the %BII, equilibrium constant, K, and thus free energy, ΔG, at each phosphate step for the two conformational state equilibrium was calculated based on Equations 1 and 2. The charts of %BII and ΔG for the 8mer U4 are below in Figure 34.
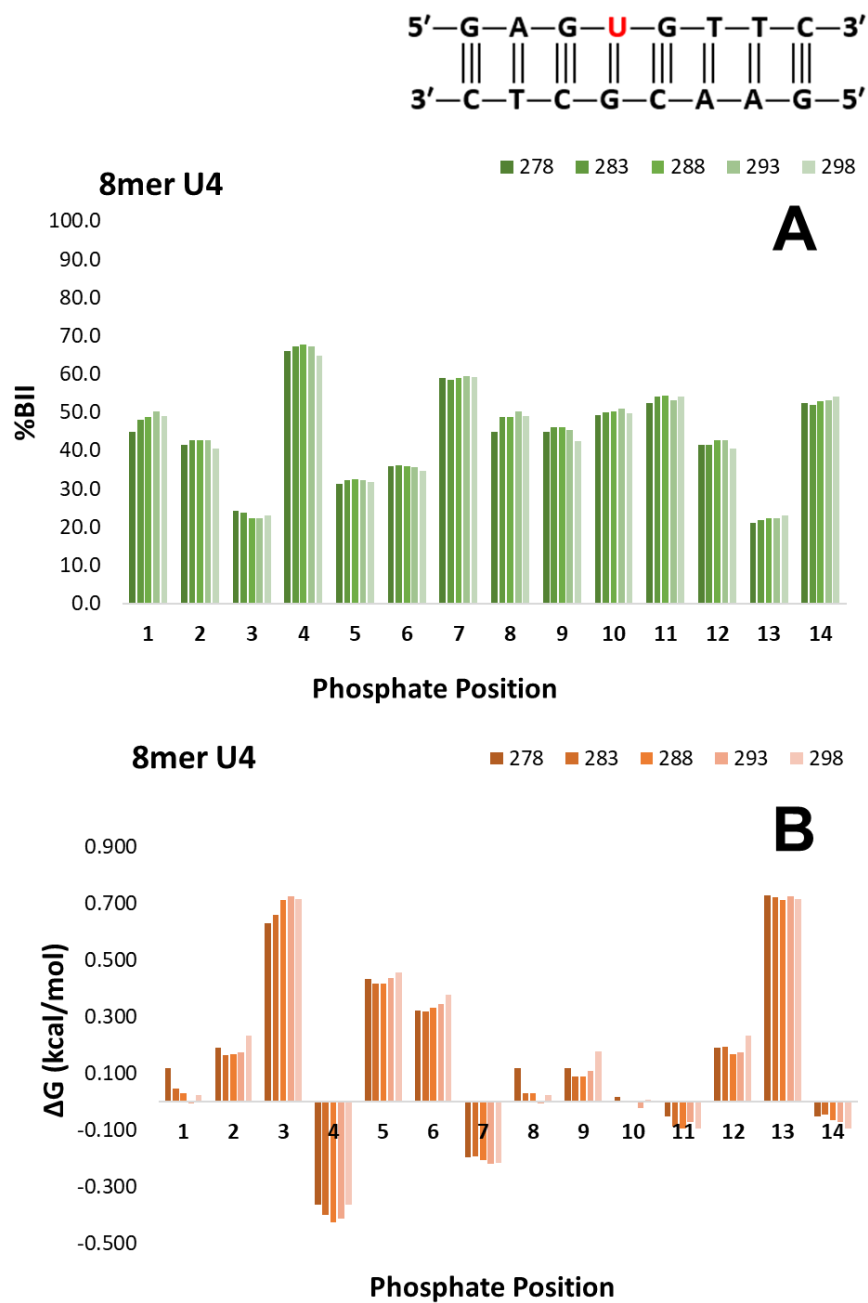
Figure 34. %BII (A) and ΔG (B) for 8mer U4 at each phosphate position from 278-298 K.

The C8 and C16 H4′ were unresolved in the H4′ region of the NOESY, but the phosphates corresponding to these bases with the HSQC are resolved. Due to not being able to distinguish the phosphates corresponding to these bases, T7pC8 and T15pC16, these peaks and thus subsequent chemical shift values were arbitrarily assigned. Thus, the calculations performed for T7pC8 and T15pC16 are questionable.

Comparison of 8mer Control and 8mer U4. Of particular interest is the comparison of the phosphorus chemical shifts between the 8mer control and the 8mer U4 and the subsequent differences in the calculated %BII and ΔG. An overlay of the HSQC spectra of the 8mer control and 8mer U4 at 283 K can be found in Figure 35. As can be seen from the overlay, the phosphates flanking the lesion site shift dramatically. The G3pU4 shifts more upfield while U4pG5 shifts downfield with respect to the control.
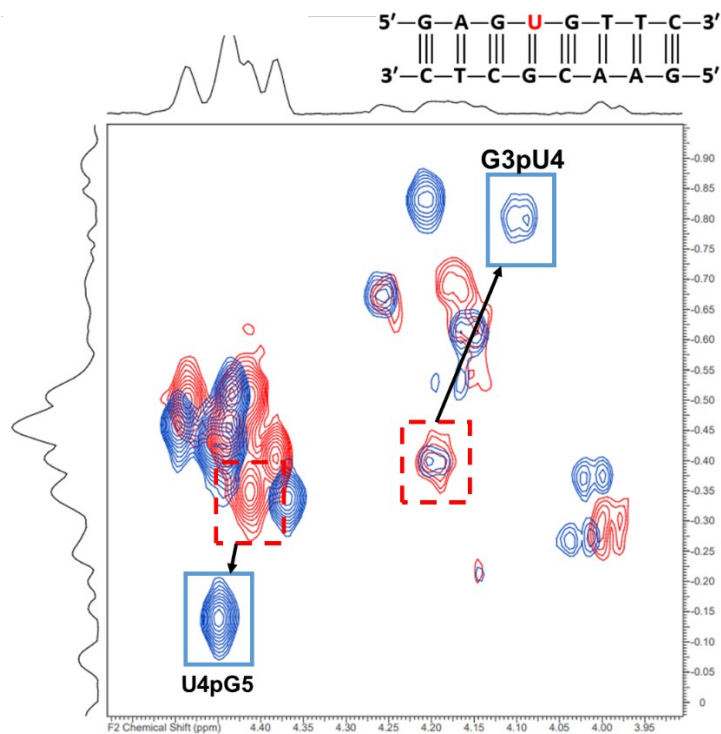
Figure 35. Overlay of the HSQC spectra of the 8mer control and 8mer U4 at 283 K. The 8mer control is the red spectrum and 8mer U4 the blue spectrum.

The change in %BII, Δ%BII, and the change in ΔG, relative ΔG, were calculated by taking the control minus the lesion sequence. These comparisons can be found in Figure 36 and Figure 37, respectively. The data presented here reflects the dramatic shifts observed within the HSQC overlay at the lesion site. There is a stepwise difference in the Δ%BII and relative ΔG around the lesion site (at phosphate position 3 and 4), as well as at the basepairing partner to a lesser extent (at phosphate position 11 and 12).

8mer FU4. The overarching goal of the 8mer project was to see how a uracil lesion, within a non-palindromic sequence, influences the phosphodiester backbone conformational dynamics. To expand this project, the idea was to analyze the backbone of the U4 DNA complexed with TDG. By adding a fluorine at the arabino position of the deoxyribose sugar, the lesioned DNA sequence can be complexed with TDG without being cleaved.[12] The incomplete proton and phosphorus assignments of the 8mer FU4 were made and are presented below. In particular, the H1′ walks of 8mer FU4 can be found in Figure 38 and H4′ assignments in Figure 39, while the rest of the representative NOESY walk at 283 K can be found in Appendix C:8mer FU4 Supplemental Data. Appendix C also contains the incomplete proton chemical shift values at the temperatures assigned. Notice how the FU4 H1′ and FU4 H4′ protons are missing within these NOESY regions. Finally the HSQC and ^{31}P stacked spectra from 278-293 K can be found in Figures 40 and 41.

Because of the missing H4′ proton, the phosphorus of the G3pFU4 step could not be assigned as well, as the phosphorus assignment relies on a confident H4′ proton assignment. This missing peak is reflected in the HSQC found in Figure 40. In addition, the FU4pG5 phosphorus is significantly upfield shifted on the HSQC spectrum and thus the 1D ^{31}P spectra, as evidenced by the peak on the right in Figure 41.
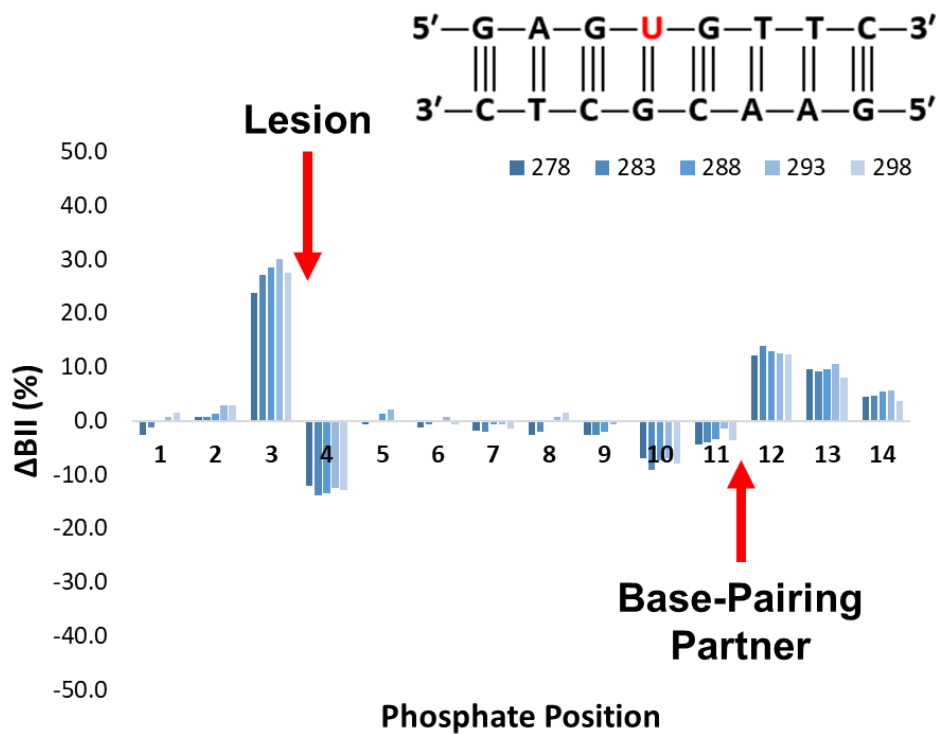
Figure 36. The change in %BII between the 8mer control and 8mer U4 at 278-298 K (calculated by ΔBII = %BII$_{control}$ - %BII$_{U4}$).
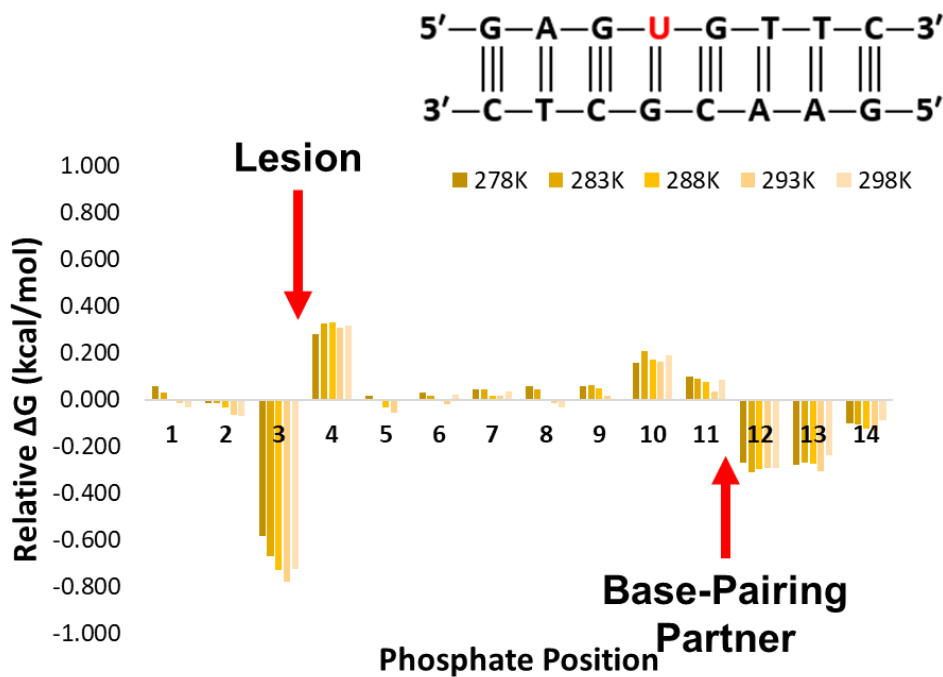
Figure 37. Relative ΔG of each phosphate step between the 8mer control and 8mer U4 at 278-298 K. Calculated by taking the ΔG$_{control}$ – ΔG$_{U4}$.
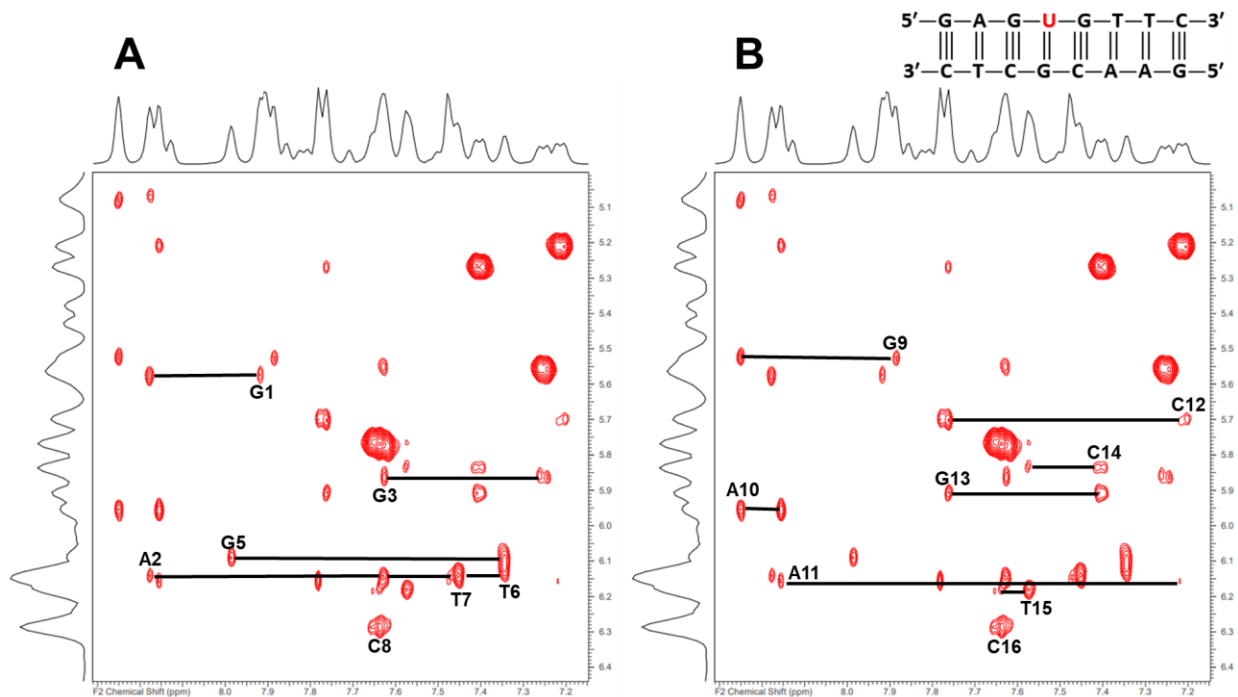
Figure 38. H1′ front-end (A) and back-end (B) walks at 283K for 8mer FU4.
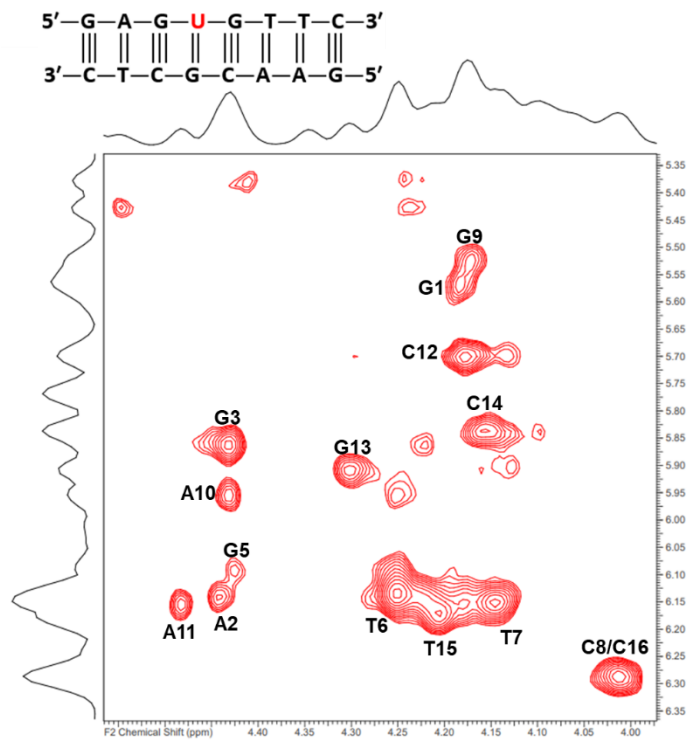
Figure 39. 8mer FU4 H4′ proton assignments at 283 K.
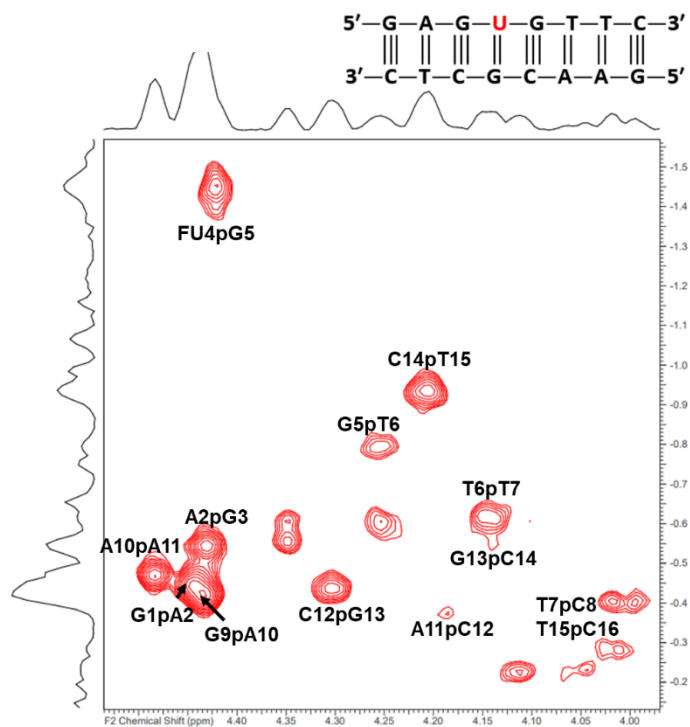
Figure 40. Assigned HSQC of 8mer FU4 at 283 K.
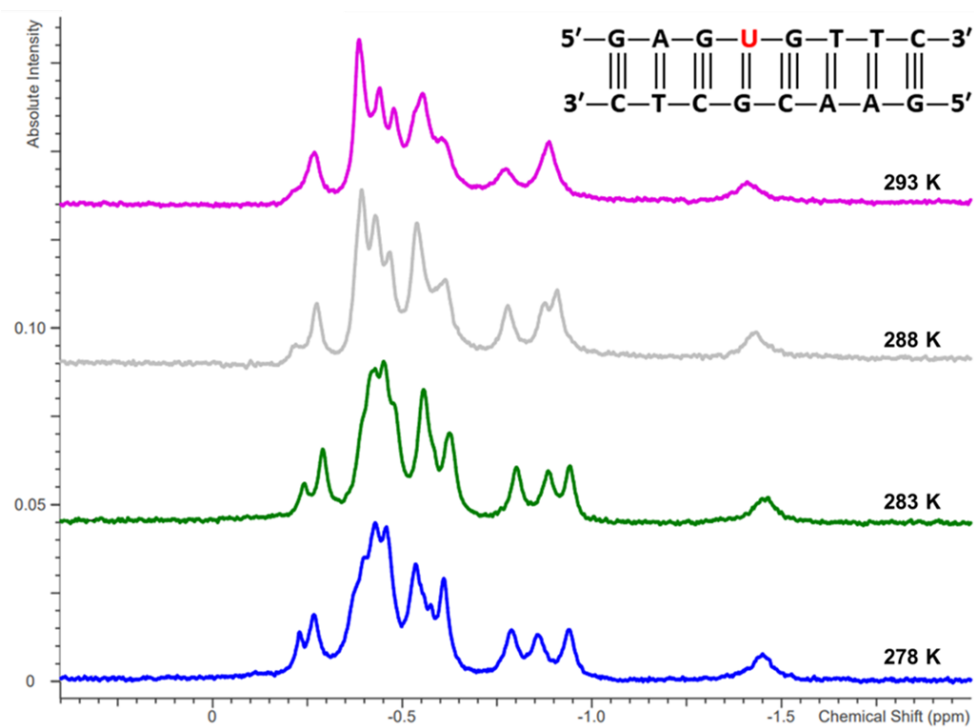
Figure 41. Stacked 1D ^{31}P spectra from 278 to 293 K in 5 K increments of the 8mer FU4.

In addition to allowing us to investigate the phosphodiester backbone conformational equilibrium dynamics, this project would also allow visualization of the flipping step of TDG[12] based on prior research completed by using 1D ^{19}F NMR experiments. Figure 42 shows a preliminary 1D ^{19}F spectrum obtained for the 8mer FU4 at 298 K.
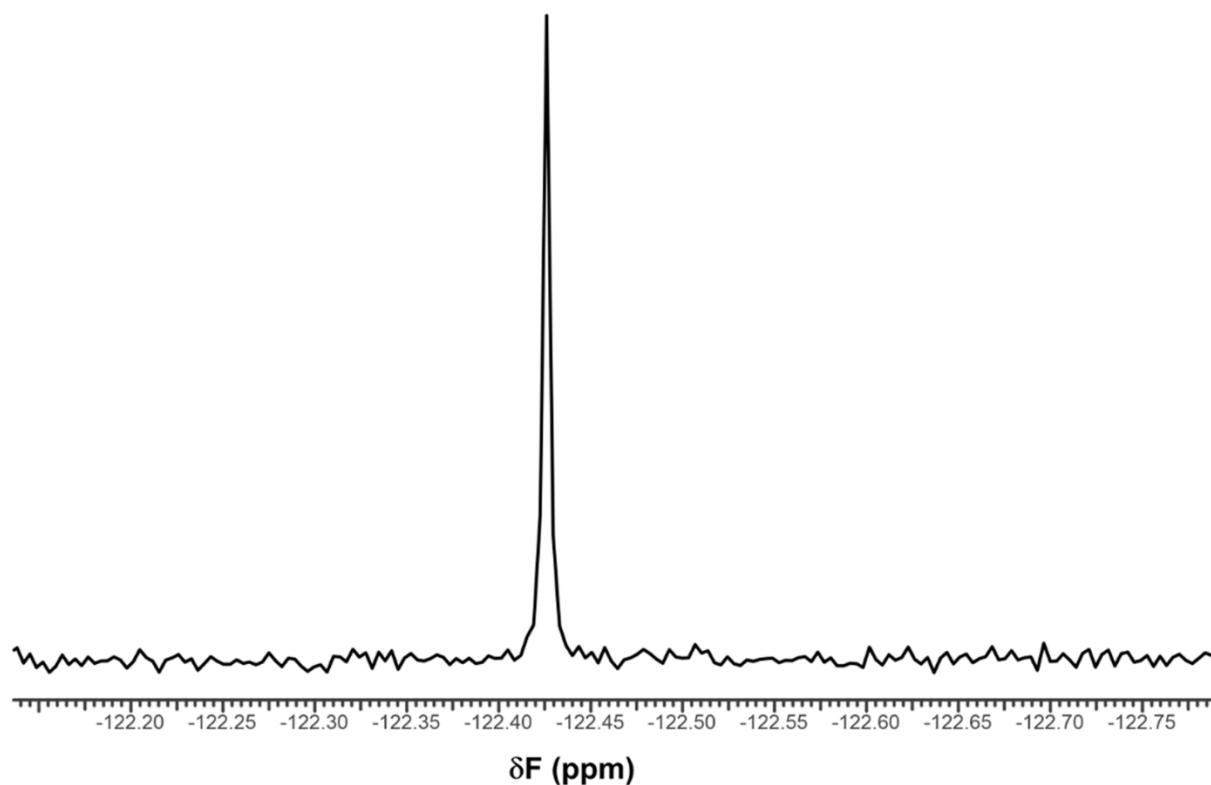
Figure 42. 1D ^{19}F spectrum of the 8mer FU4 at 298 K.

The Evolution of the Expression and Purification of TDG Determined

The initial project was aimed at expressing and purifying hTDG and then complexing the enzyme with a fluorinated version of the 8mer U4 non-palindromic DNA as discussed previously. This would have enabled an investigation of the phosphorus backbone while in the presence of the TDG enzyme. Due to unforseen complexities with the expression and purification of hTDG, there were thirteen attempts to the expression/purification process to yield

better results. The Methods section presented the protocol that was determined to be the most promising, while this section presents a few of the key steps that assisted in determining this protocol.

When first beginning this project, the protocol was based off of methods working with hTDG previously published.[5-6,16] To get to the protocol presented in the Methods section, it took a lot of time and effort with many instances of trial and error. hTDG was found to be a very difficult protein to work with, and many parts of the procedure had to be changed. This section explores a few of the key steps within the procedure that were altered to ultimately lead to the final protocol.

Optimization of Transformation. The first part of any expression and purification is the transformation and plating protocol. A basic protocol for a transformation of proteins was used with a couple important aspects to note. During the transformation, there were a couple attempts investigating what *E. coli* cell line (BL21 or JM109) would be most ideal, as well as brand after the cell line was determined. BL21 was found to produce the best plate of colonies. From this attempt it was also determined that an outgrowth step would be beneficial for the transformation. A later attempt tested Sigma Aldrich derived, Agilent, and Gold Bio derived competent cells, finding that Agilent produced no colonies, Gold Bio resulted in a few colonies, and Sigma Aldrich resulted in many colonies. Figure 43 shows the comparison of the plates testing these brands of BL21 competent cells. After the determination of the competent cell type (BL21), the ideal brand, and the requirement of an outgrowth step, the transformation step was considered optimized.

Expression/Induction Attempts. The second aspect of the procedure that had to be optimized was the expression and induction step. The optimal concentration of IPTG to induce

was determined, as well as the ideal OD at induction and the temperature during expression. Small scale growths were used to determine the ideal IPTG concentration and OD by using 4 mL of overnight culture added to 100 mL of LB media. For the IPTG concentration, the following concentrations were tested: 0.25 mM, 0.5 mM, 1.0 mM, and 1.5 mM. The OD at induction was around 1.2 and the samples after being induced were incubated at around 26-30°C for overnight with shaking. The resulting lysates were isolated as described in the Methods section, and they were tested with a SDS-PAGE gel. This gel result can be found in Figure 44. They generally had similar band thickness, but 0.5 mM IPTG appeared qualitatively to be the thickest band. This concentration was used for subsequent preparations.

Figure 43. The LB-kan agar plates with *E. coli* colonies containing the TDG plasmid vector. The left most plates are the Gold Bio cell results, the middle two plates are Sigma Aldrich, and the right most plates are Agilent cells.
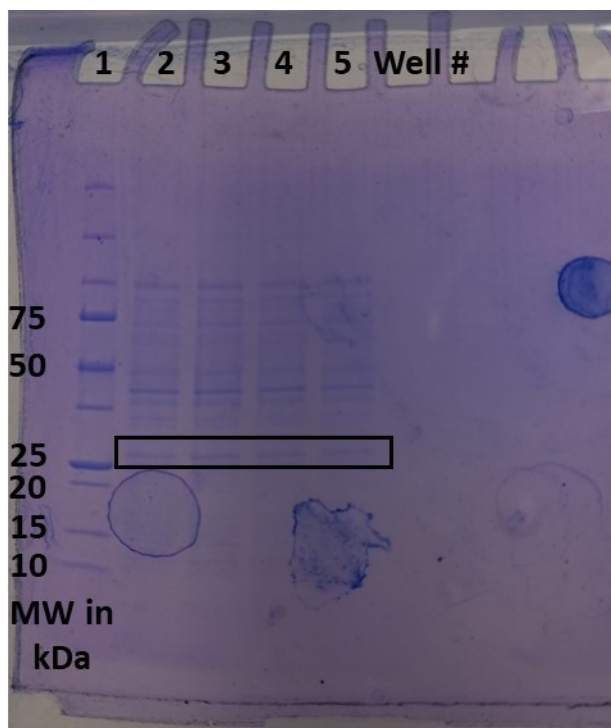
Figure 44. SDS-PAGE gel of the IPTG concentration optimization. Well 1: contains the ladder. Well 2: 0.25 mM IPTG sample. Well 3: 0.5 mM IPTG sample. Well 4: 1.0 mM IPTG sample. Well 5: 1.5 mM IPTG sample. The black box indicates the presumed TDG band.

The optimization of the OD tested 0.8, 1.1, 1.58, and 2.0 OD. At induction, 0.5 mM IPTG was used and the samples were incubated overnight at 25-30°C. The resulting lysates again were isolated and ran on a SDS-PAGE gel. This gel can be found in Figure 45. They all had pretty similar band thickness at around 26 kDa (target protein), but 0.8 OD appeared qualitatively to be the thickest band and an OD of 2 was the lightest. This follows the protocols presented in literature for hTDG expression using an OD of 0.8.[5-6,15]

In addition, in one attempt TB Broth was used instead of LB Broth and a weekend length induction time along with this TB Broth. During this attempt, the quality of the resuspended cell pellet was extremely different, resulting in a very thick, aerated sample. This can be seen in Figure 46. In addition, after the purification was completed, there was a band in the gel around 35 kDa which could point to the wrong protein expressed with the longer induction time. In

addition to this attempt, other expression attempts resulted in a change to the quality of the cultures/cell pellets, to where at times there were black flecks in the sample. Often times these black flecks correlated to a more poor yield.
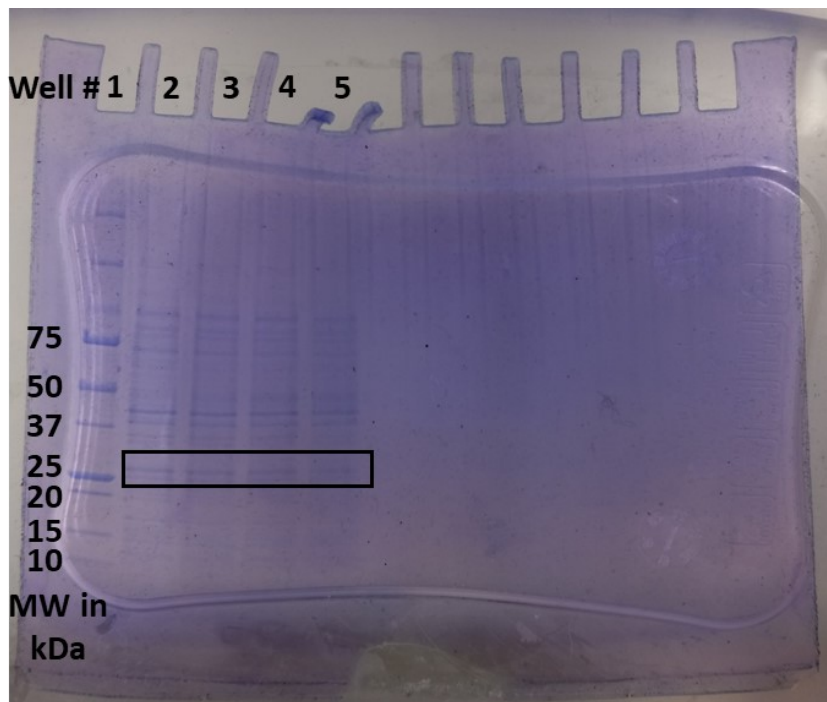
Figure 45. SDS-PAGE gel of the OD optimization results. Well 1: ladder. Well 2: 0.8 OD sample. Well 3: 1.1 OD sample. Well 4: 1.58 OD sample. Well 5: 2.0 OD sample. The black box indicates the TDG band.

Throughout the numerous attempts, it was observed that even after these changes to the protocol, the yield of TDG was quite low. What had been attempted in the first 9 attempts was to get the temperature of the shaking incubator as low as it could go for the overnight induction step. This is because in literature all references induced TDG expression at 15°C.[5-6,16] Due to the lack of a cooling attachment on the incubator, the incubator was unable to get to this lower temperature. To get to this lower temperature, the small incubator was moved into the cold room, which allowed for the incubator to get to 15°C to allow the samples to incubate at this colder

temperature for overnight. In addition, all columns were ran in the cold room. After this was

implemented, the protein stopped precipitating.

Figure 46. Image of the resuspended cell pellet using TB Broth and after a weekend induction.

Optimization of Purification In addition to the expression protocol, the purification

protocol was optimized. The originial protocol involved running a Ni-NTA affinity column, then

a CM ion-exchange column, then a Tev digestion was completed to cut the His tag off of the

protein, then another Ni-NTA affinity column was performed. It was noticed that after the ion-

exchange column there was typically a significant loss in yield. After the second Ni-NTA

column there was typically little-to-no protein yield obtained. The procedure in the last two

attempts removed the ion-exchange column step and instead replaced it with a dialysis step,

followed by a Ni gravity column. Both of these attempts saw some sort of yield of TDG

(presumably. There was a band at around 25, 50, and 75 kDa after the second Ni column) which

was an improvement over the protocol containing ion-exchange. This gel can be found in Figure 28 within the Methods section.

Influence of Sample Conditions on Palindromic DNA Containing a T:G Mismatch

The research that has previously been completed on analyzing the phosphodiester backbone conformational dynamics using solution NMR has been DNA in 25 mM phosphate buffer.[22] Cellular environments are typically very crowded though, with other proteins, small molecules, and other biomolecules present.[30] To accurately mimic cellular conditions and answer the overarching question of how the recognition process of BER occurs, sample conditions should be investigated. The presence of varying ions, ionic strength, pH, and an artificial crowding agent were investigated within the solution conditions of dodecamer palindromic DNA, the Dickerson-Drew Dodecamer (DDD) and the same sequence containing a T:G mismatch at position 9 (T9) as described in the Methods. The specific sample conditions investigated can be found in the Methods section as well. All tables of phosphorus chemical shift values, individual 1D ^{31}P spectra, and individual %BII and ΔG plots can be found within Appendix D: DDD Sample Conditions Supplemental Data for each DDD sample condition and Appendix E: T9 Sample Conditions Supplemental Data for each T9 sample condition. Proton chemical shift values, NOESY, and HSQC assignments can also be found in Appendix D and E for samples that required a NOESY/HSQC to assign the phosphorus.

Dickerson-Drew Dodecamer (DDD) Sample Conditions. The following conditions were used for the DDD sequences: 25 mM phosphate buffer (primary control), 100 mM NaCl, 5 mM MgCl$_2$, 20 mM MgCl$_2$, 300 mM NaCl, 100 mM KCl (Buffers 1-4, 6), pH 5.0, and pH 9.0. A stack of the 1D ^{31}P spectra at 273 and 298 K for DDD in the varying salt conditions can be

found in Figure 47 and 48, respectively. Figure 49 and 50 present the stack of 1D ^{31}P spectra at 273 and 298 K for DDD in the varying pH conditions, respectively. Outside of some resolution and peak quality changes, there are no dramatic differences in the spectra qualitatively, and thus the phosphorus chemical shifts, within each of these conditions outside of the MgCl$_2$ samples.

The %BII and ΔG of each sample at each phosphate step was calculated as described in the Methods. The Δ%BII, calculated by %BII$_{\text{25 mM Phosphate Buffer}}$ - %BII$_{\text{varied sample condition}}$, for each DDD salt and pH condition can be found in Figure 51 and Figure 52, respectively. The relative ΔG, ΔG$_{\text{25 mM Phosphate Buffer}}$ − ΔG$_{\text{varied sample condition}}$, for each DDD salt and pH condition can be found in Figure 53 and 54, respectively.

Figure 47. Stacked 1D ^{31}P spectra for the varying salt conditions of DDD at 283 K.

Figure 48. Stacked 1D ^{31}P spectra for the varying salt conditions of DDD at 298 K.

Figure 49. Stacked 1D ^{31}P spectra for the varying pH conditions of DDD at 283 K.

Figure 50. Stacked 1D ^{31}P spectra for the varying pH conditions of DDD at 298 K.

As can be seen from Figure 53 and 54, and the few changes seen within the 1D ^{31}P spectra in Figure 47-50, there were only minute changes to the conformational equilibrium of each phosphate step when the different conditions were introduced when compared to the control 25 mM phosphate condition. The average changes in %BII for the DDD sequences across all of the sample conditions was less than a ±2% change at 283 and 298 K, except for 20 mM MgCl$_2$ and the pH 9 condition. 20 mM MgCl$_2$ saw a 3.53 and 4.72% change in BII at 283 and 298 K, respectively. The pH 9 condition shows a -2.99 and -3.61% BII change at 283 and 298 K, respectively. The average changes in relative ΔG were less than or equal to a ±0.05 kcal/mol at 283 and 298 K for all sample conditions, except for 20 mM MgCl$_2$ and the pH 9 condition. MgCl$_2$ saw a -0.09 and -0.13 kcal/mol change in ΔG at 283 and 298 K, respectively. The pH 9 condition shows a 0.07 and 0.09 kcal/mol change in ΔG at 283 and 298 K, respectively. The

largest change that was seen in the relative ΔG and quality of the spectra is for 20 mM MgCl$_2$. This sample saw line broadening within the spectra and a slight increase in the BI conformation population compared to the 25 mM phosphate buffer control condition. An important note to make is the 5 mM MgCl$_2$ sample does not follow the trend expected based on the 20 mM MgCl$_2$ data (shifting more to the right), and the referencing coax experiment should be completed again in future experiments.

Figure 51. Δ%BII plots for the DDD salt conditions at 283 (blue) and 298K (orange). The change in %BII is calculated by %BII$_{25 \text{ mM Phosphate Buffer}}$ - %BII$_{\text{varied sample condition}}$.
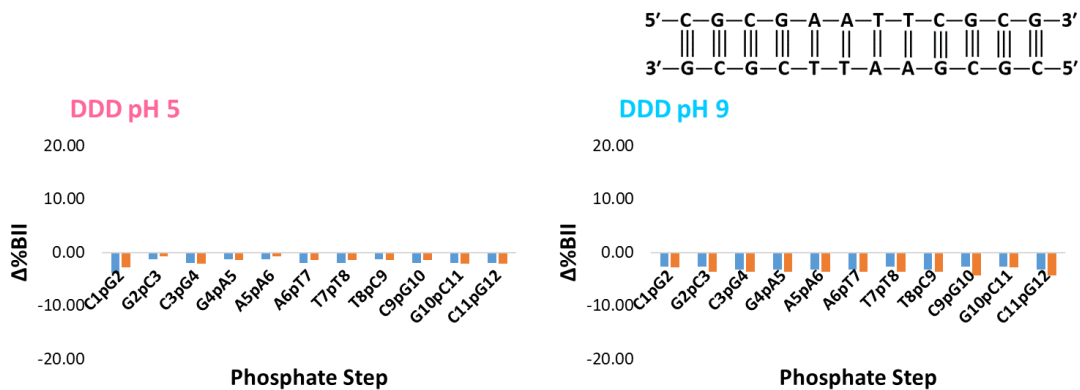
Figure 52. Δ%BII plots for the DDD pH conditions at 283 (blue) and 298K (orange). The change in %BII is calculated by %BII$_{25\text{ mM Phosphate Buffer}}$ - %BII$_{\text{varied sample condition}}$.

Figure 53. Relative ΔG plots for the DDD salt conditions at 283 (blue) and 298K (orange). The relative ΔG is calculated by ΔG$_{25\text{ mM Phosphate Buffer}}$ − ΔG$_{\text{varied sample condition}}$.

5′—C—G—C—G—A—A—T—T—C—G—C—G—3′
3′—G—C—G—C—T—T—A—A—G—C—G—C—5′

DDD pH 5

Rel. ΔG (kcal/mol)

0.25
0.00
-0.25
-0.50
-0.75
-1.00

C1pG2 G2pC3 C3pG4 G4pA5 A5pA6 A6pT7 T7pT8 T8pC9 C9pG10 G10pC11 C11pG12

Phosphate Step

DDD pH 9

Rel. ΔG (kcal/mol)

0.25
0.00
-0.25
-0.50
-0.75
-1.00

C1pG2 G2pC3 C3pG4 G4pA5 A5pA6 A6pT7 T7pT8 T8pC9 C9pG10 G10pC11 C11pG12
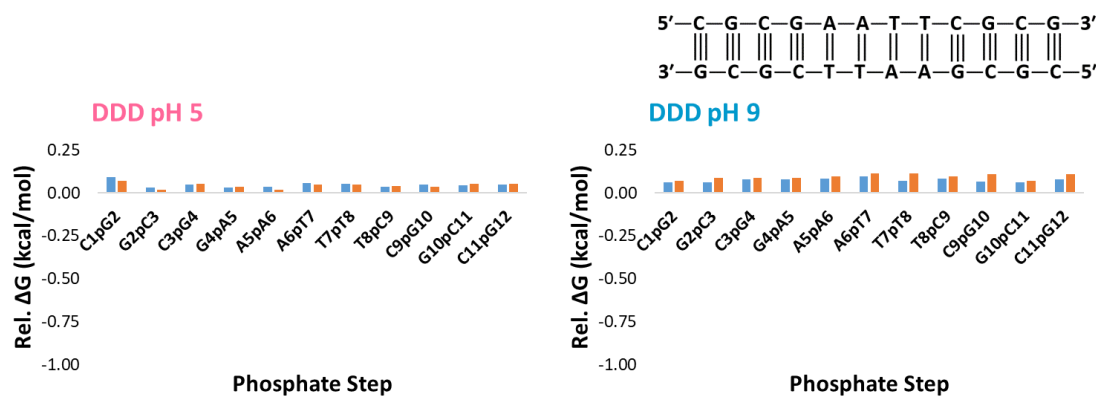
Phosphate Step

Figure 54. Relative ΔG plots for the DDD pH conditions at 283 (blue) and 298 K (orange). The relative ΔG is calculated by $\Delta G_{25 \text{ mM Phosphate Buffer}} - \Delta G_{\text{varied sample condition}}$.

T9 Sample Conditions. The following conditions were used for the T9 sequences: 25 mM phosphate buffer (primary control), 100 mM NaCl, 200 mM NaCl, 300 mM NaCl, 5 mM MgCl$_2$, 20 mM MgCl$_2$, 100 mM KCl (Buffers 1-6), pH 5.0, pH 9.0, 10% w/w PEG-4000, and 20% w/w PEG-4000. The sample preparation is described within the Methods section. A stack of the 1D ^{31}P spectra at 273 and 298 K for T9 in the varying salt conditions can be found in Figure 55 and 56, respectively. Figure 57 and 58 present the stack of 1D ^{31}P spectra at 273 and 298 K for T9 in the varying pH conditions, respectively.

The 1D ^{31}P spectra in Figures 55 and 56 show there are changes that occur in the peaks with a few of the varied salt conditions. The presence of MgCl$_2$ results in line broadening of the phosphorus spectra, similar to what was observed for DDD, but there were also slight shifts in where the peaks are located. In addition, there are instances of improved resolution for some sample conditions. There are also peaks that are perturbed as well, as can be seen in particular for the T9 100 mM KCl sample. For the pH conditions, there are very few changes observed in the 1D ^{31}P spectra. The few changes that were observed were some improved resolution of peaks in the pH 5 condition as can be seen in Figures 57 and 58.
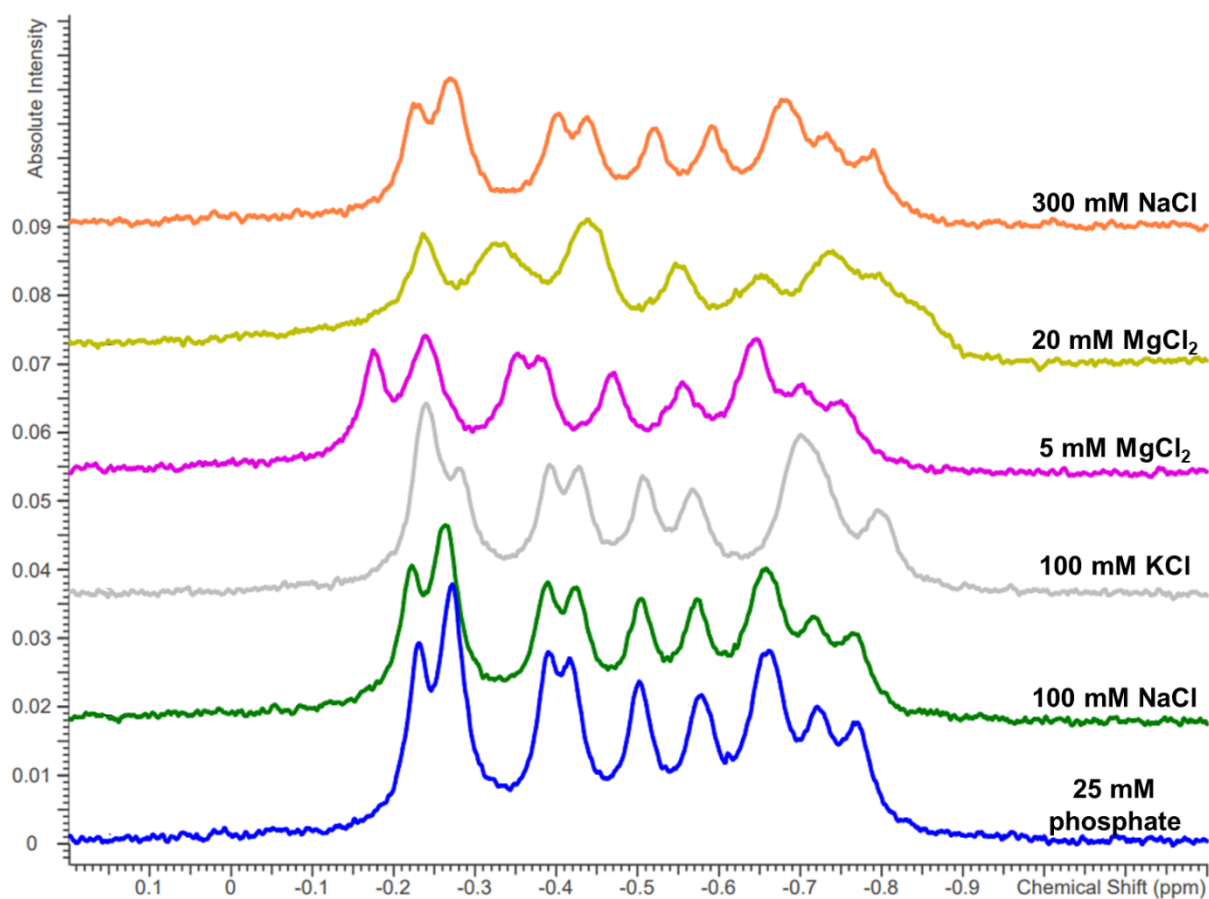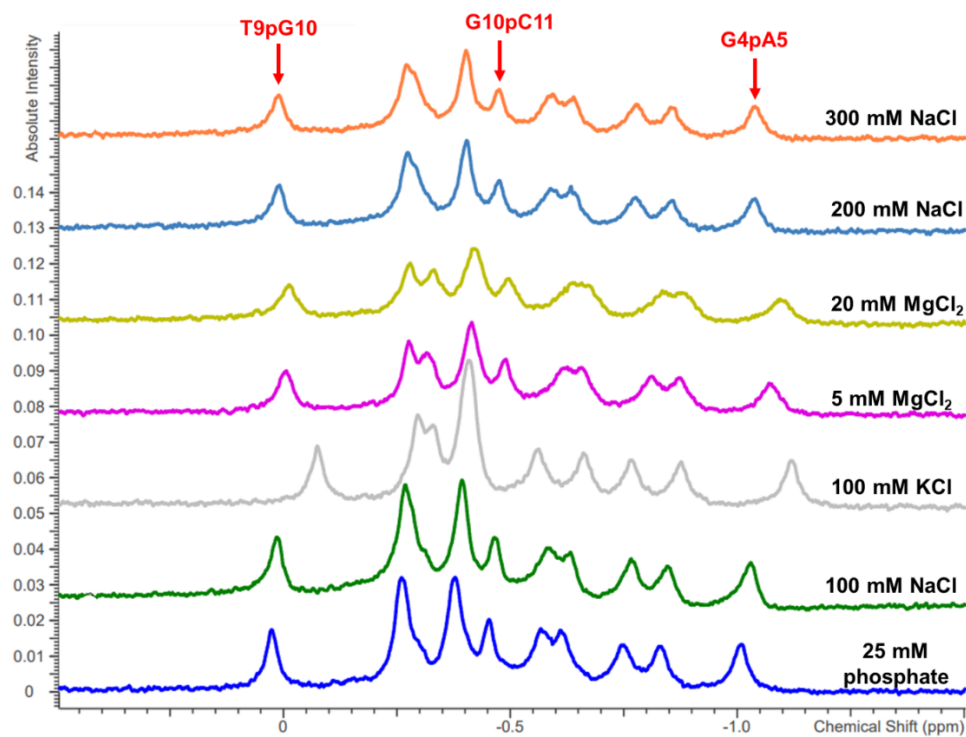
Figure 55. Stacked 1D ^{31}P spectra for the varying salt conditions of T9 at 283 K.
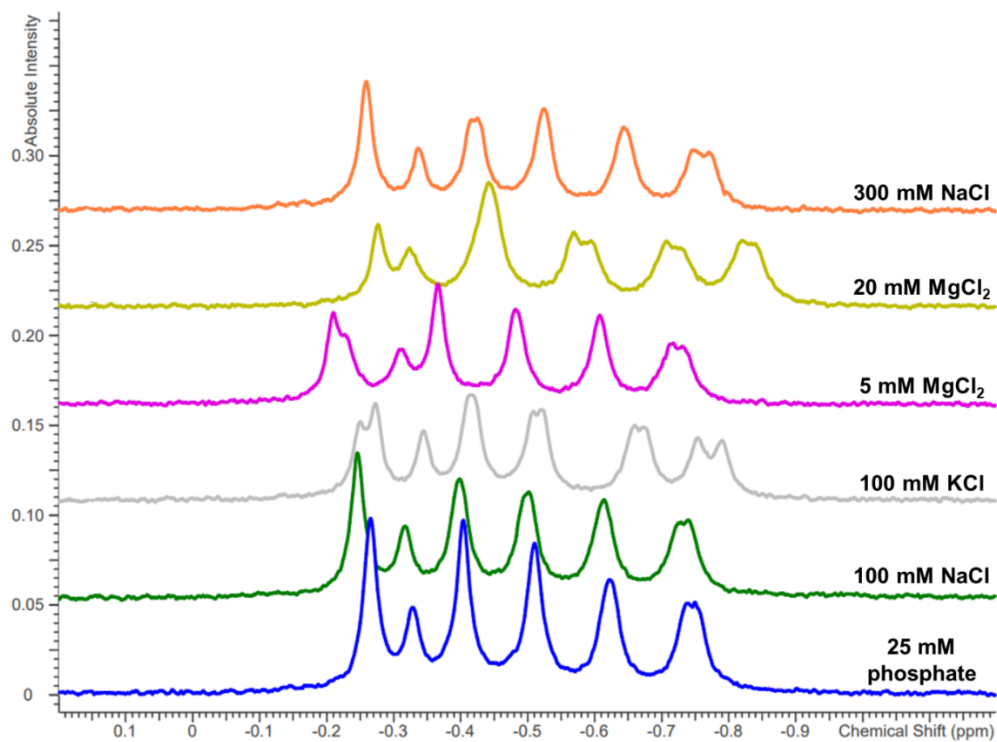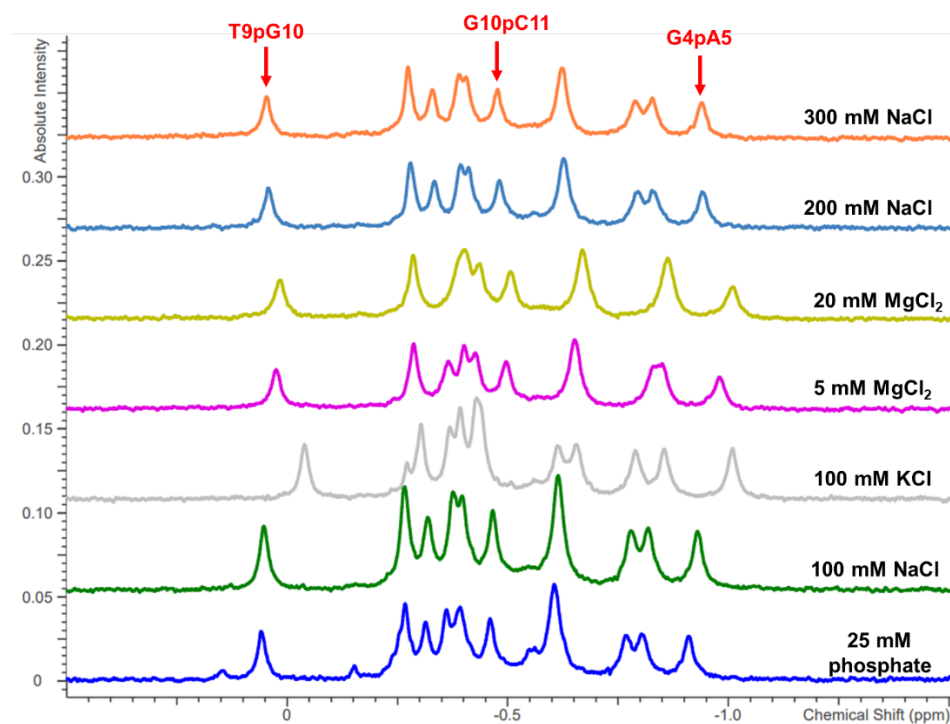

Figure 56. Stacked 1D ^{31}P spectra for the varying salt conditions of T9 at 298 K.
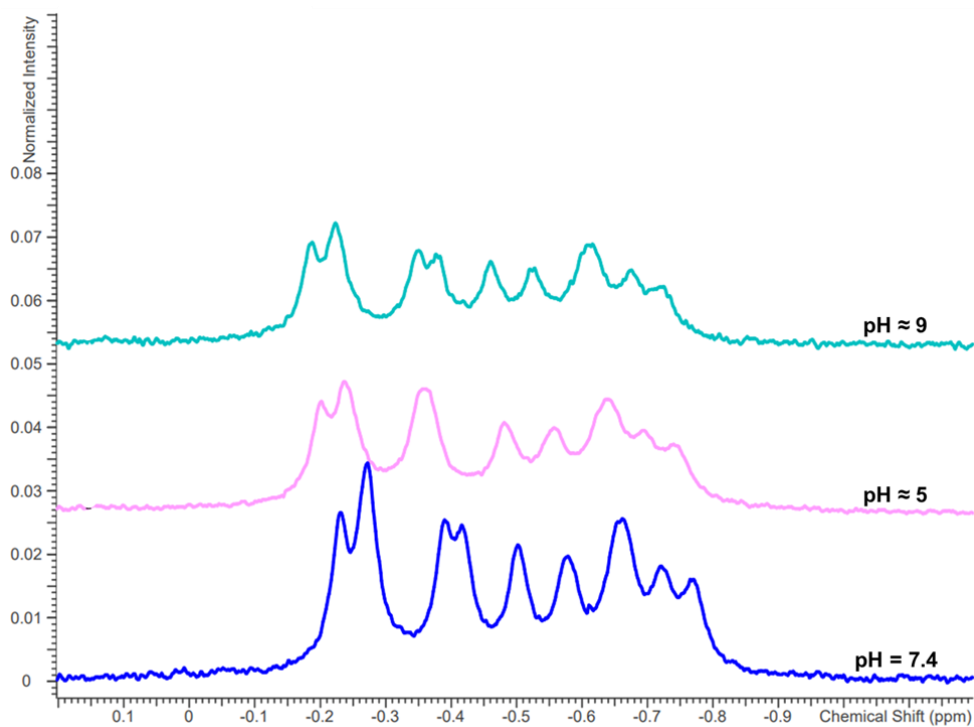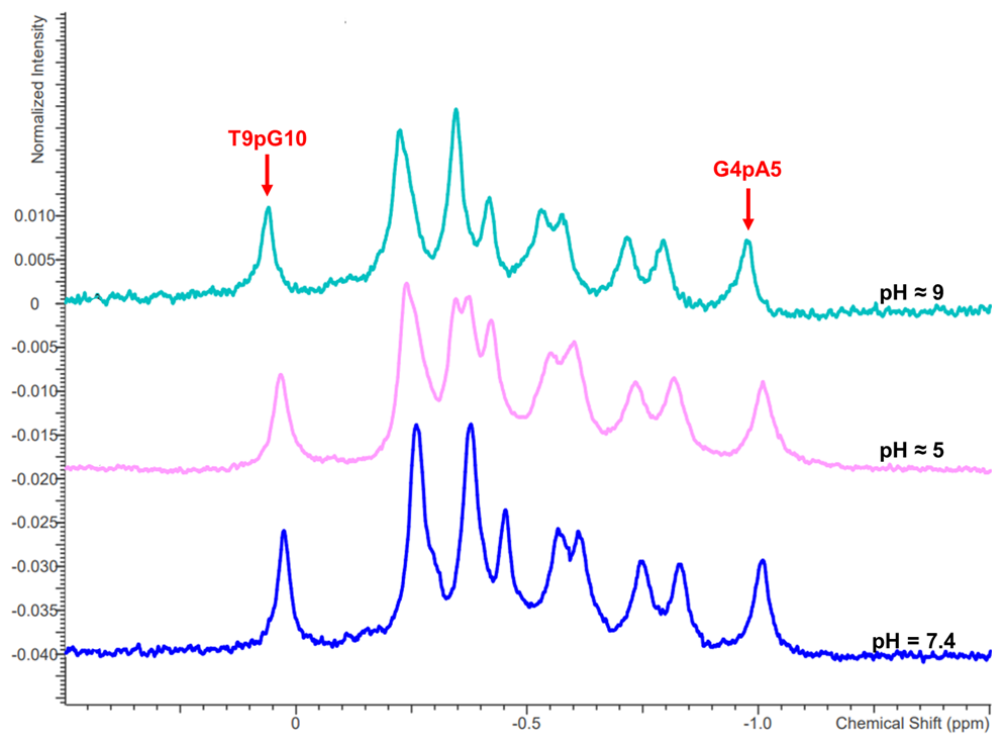
Figure 57. Stacked 1D ^{31}P spectra for the varying pH conditions of T9 at 283 K.
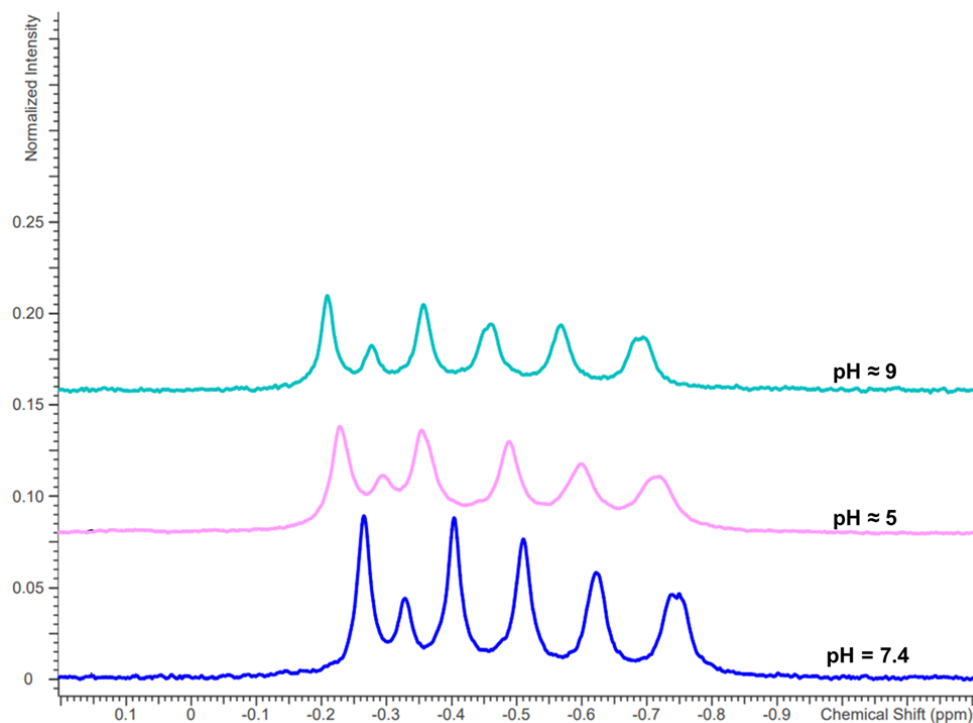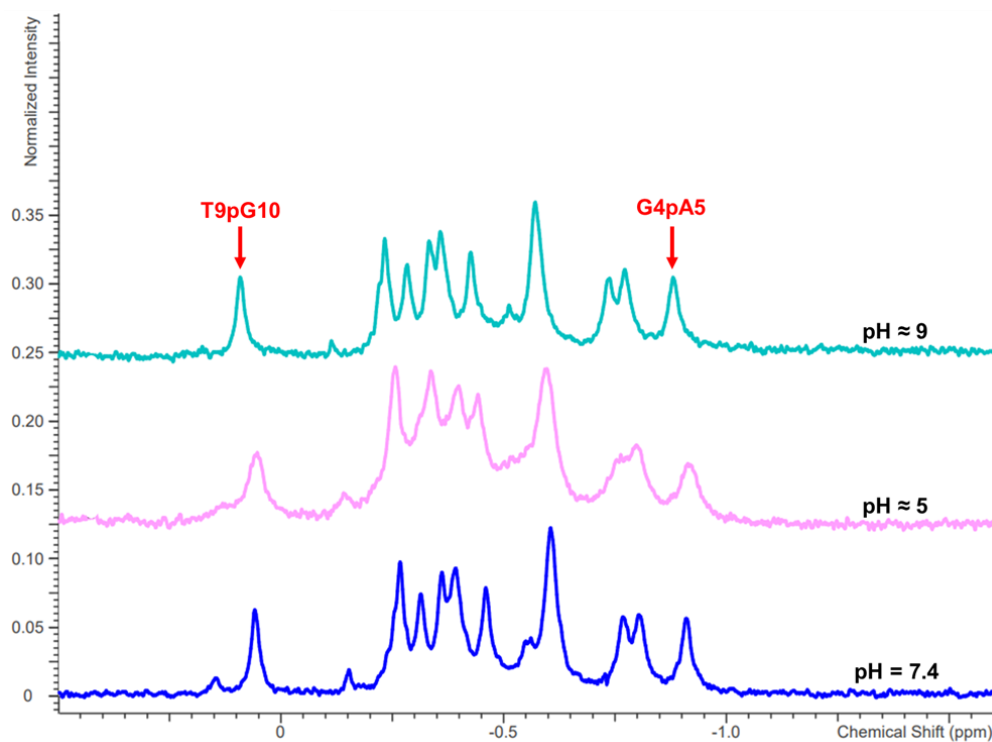

Figure 58. Stacked 1D ^{31}P spectra for the varying pH conditions of T9 at 298 K.

The %BII and ΔG of each sample at each phosphate step was calculated as described in the Methods. The Δ%BII, calculated by %BII$_{25 \text{ mM Phosphate Buffer}}$ - %BII$_{\text{varied sample condition}}$, for each T9 salt and pH condition can be found in Figure 59 and Figure 60, respectively. The relative ΔG, ΔG$_{25 \text{ mM Phosphate Buffer}}$ − ΔG$_{\text{varied sample condition}}$, for each T9 salt and pH condition can be found in Figure 61 and 62, respectively.
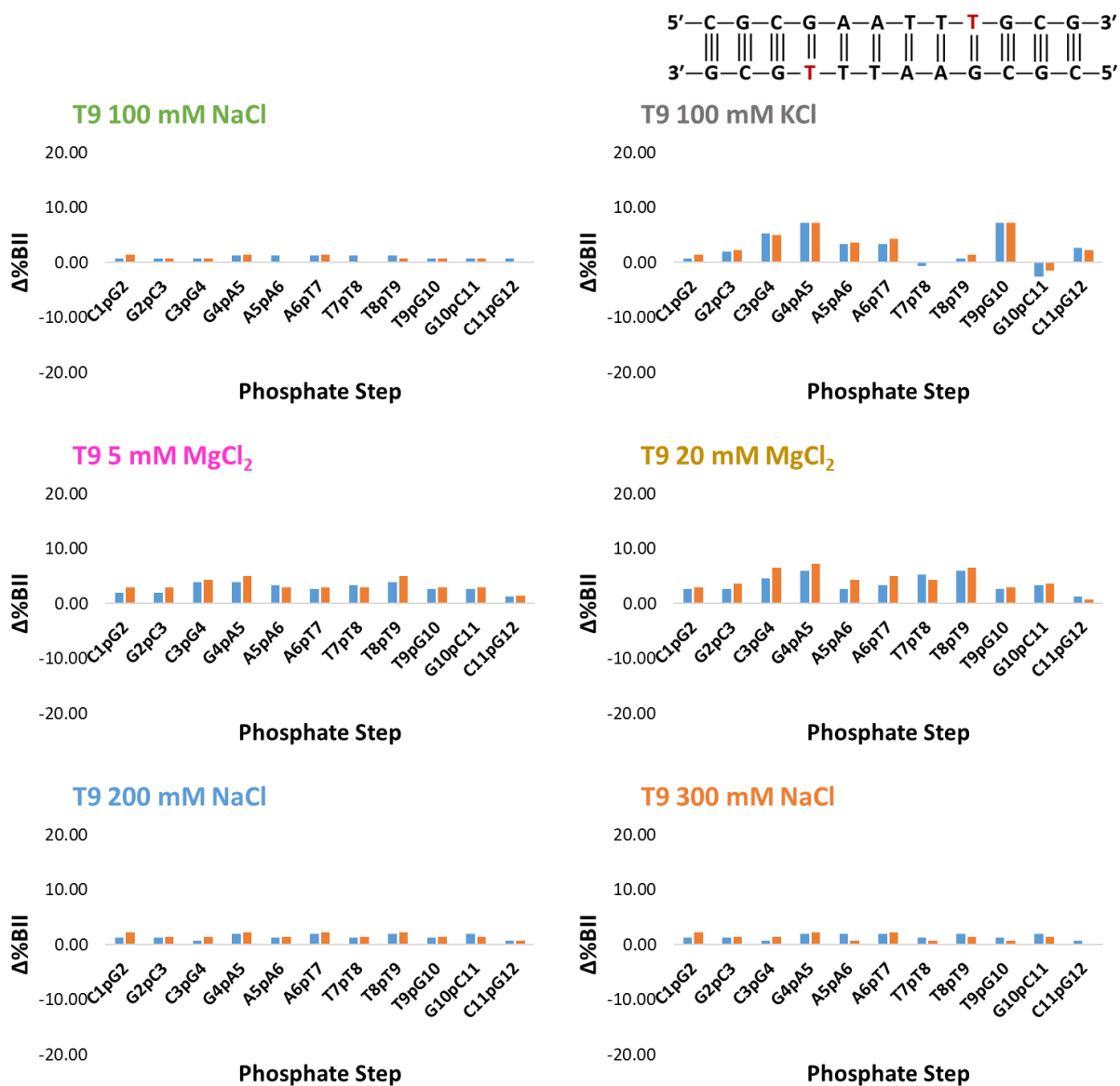
Figure 59. Δ%BII plots for the T9 salt conditions at 283 (blue) and 298 K (orange). The change in %BII is calculated by %BII$_{25 \text{ mM Phosphate Buffer}}$ - %BII$_{\text{varied sample condition}}$.
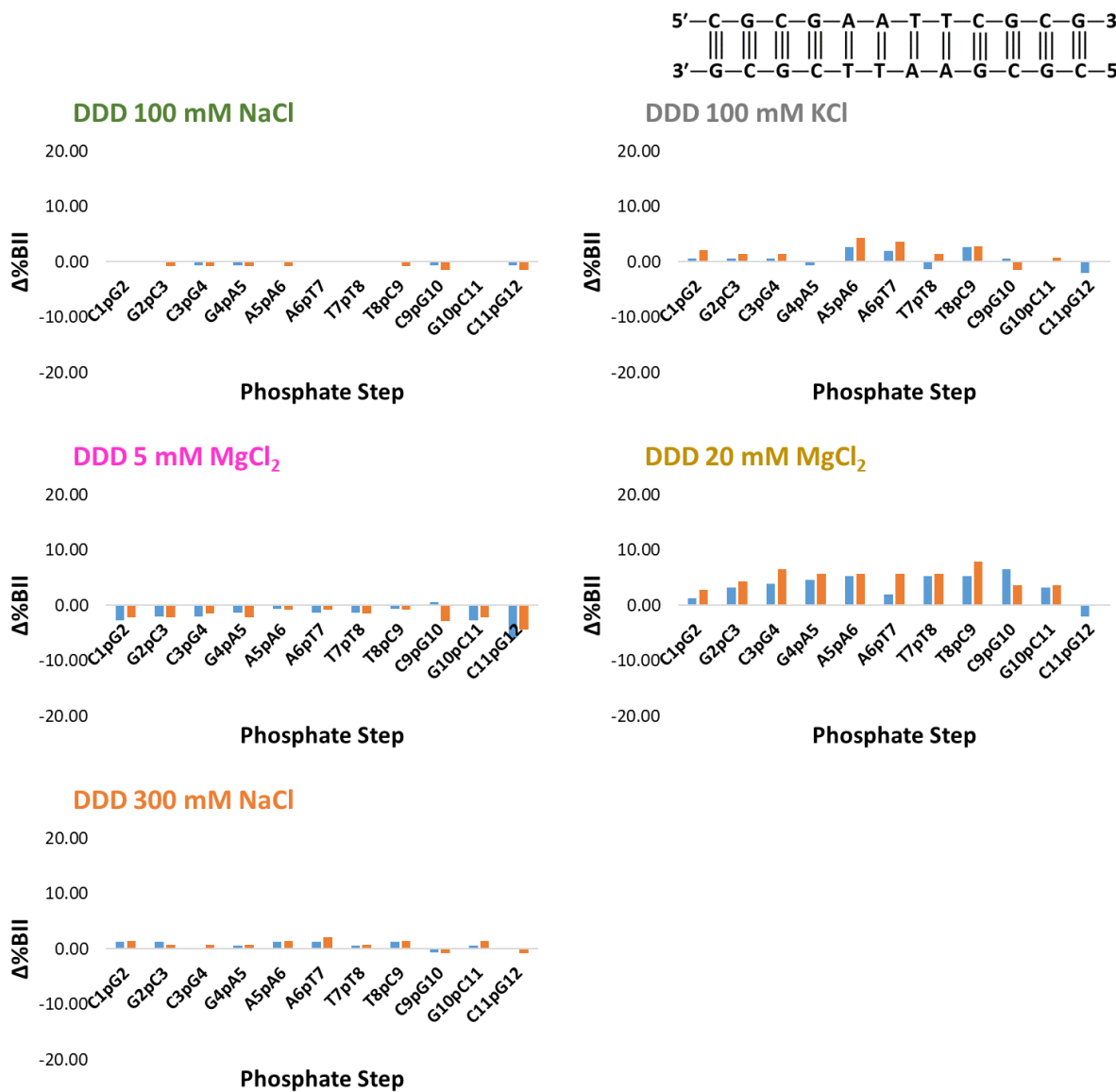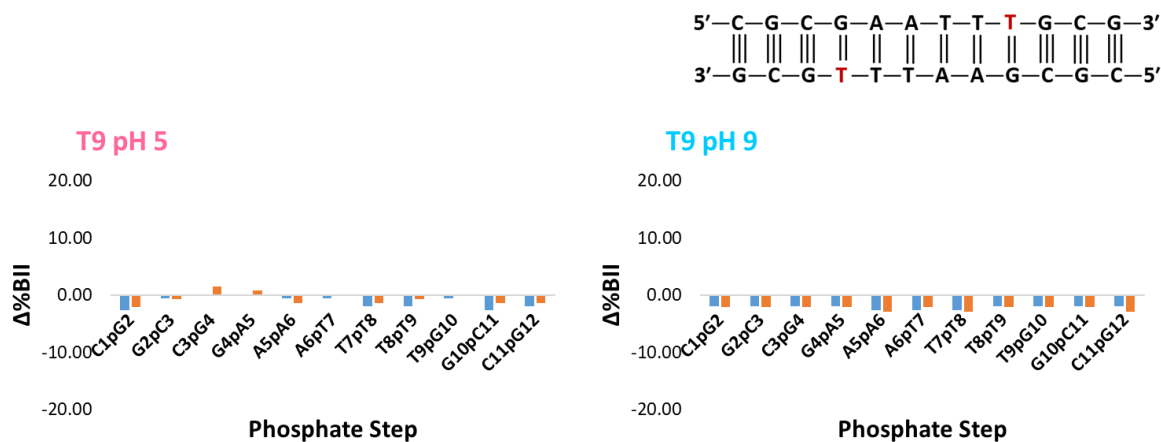
Figure 60. Δ%BII plots for the T9 pH conditions at 283 (blue) and 298 K (orange). The change in %BII is calculated by %BII$_{25 \text{ mM Phosphate Buffer}}$ - %BII$_{\text{varied sample condition}}$.
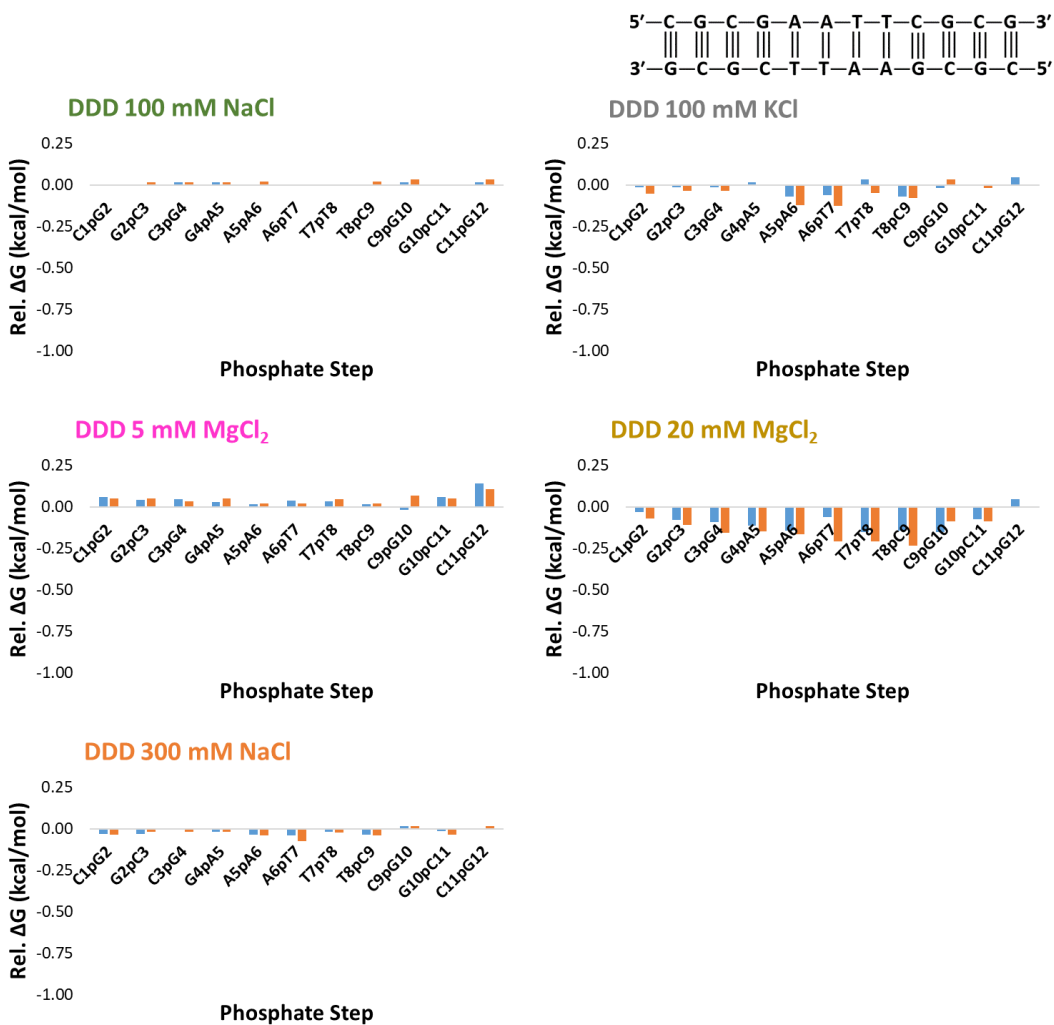
Figures 59 and 61 show significant changes in the %BII and ΔG for the phosphate 3′ to the base-pairing partner (G4pA5) for 100 mM KCl and the MgCl$_2$ samples, and to a lesser extent 3′ to the lesion (T9pG10) for 100 mM KCl. The NaCl samples (100, 200, and 300 mM) and the pH samples saw only slight changes in the %BII and ΔG when compared to the 25 mM phosphate buffer condition. The average changes in %BII for the T9 sequences across all of the sample conditions was less than a ±2.5% change at 283 and 298 K, except for 5 mM MgCl$_2$, 20 mM MgCl$_2$ and the 100 mM KCl. The change in %BII at G4pA5, where the most significant change occurred, for the 5mM MgCl$_2$ sample at 283 and 298 K was 3.95 and 5.05%, respectively. For the 20mM MgCl$_2$ the change in %BII at G4pA5 at 283 and 298 K was 5.92 and 7.21%, respectively. For the 100 mM KCl the change in %BII at G4pA5 at 283 and 298 K was 7.24 and 7.21%, respectively. 100 mM KCl also saw a 7.24 and 7.21% change in %BII at 283 and 298 K, respectively for the T9pG10. The average changes in ΔG for the T9 sequences across all of the sample conditions was less than a ±0.07 kcal/mol change at 283 and 298 K, except for 5 mM MgCl$_2$, 20 mM MgCl$_2$ and the 100 mM KCl. The change in ΔG at G4pA5, where the most significant change occurred, for the 5mM MgCl$_2$ sample at 283 and 298 K was -0.31 and -0.37

kcal/mol, respectively. For the 20mM $MgCl_2$ the change in ΔG at G4pA5 at 283 and 298 K was -0.55 and -0.63 kcal/mol, respectively. For the 100 mM KCl the change in ΔG at G4pA5 at 283 and 298 K was -0.79 and -0.63 kcal/mol, respectively. 100 mM KCl also saw a -0.22 and -0.25 kcal/mol change in ΔG at 283 and 298 K, respectively for the T9pG10. As can be observed, the most significant difference in %BII and ΔG was with 100 mM KCl at the G4pA5 step.
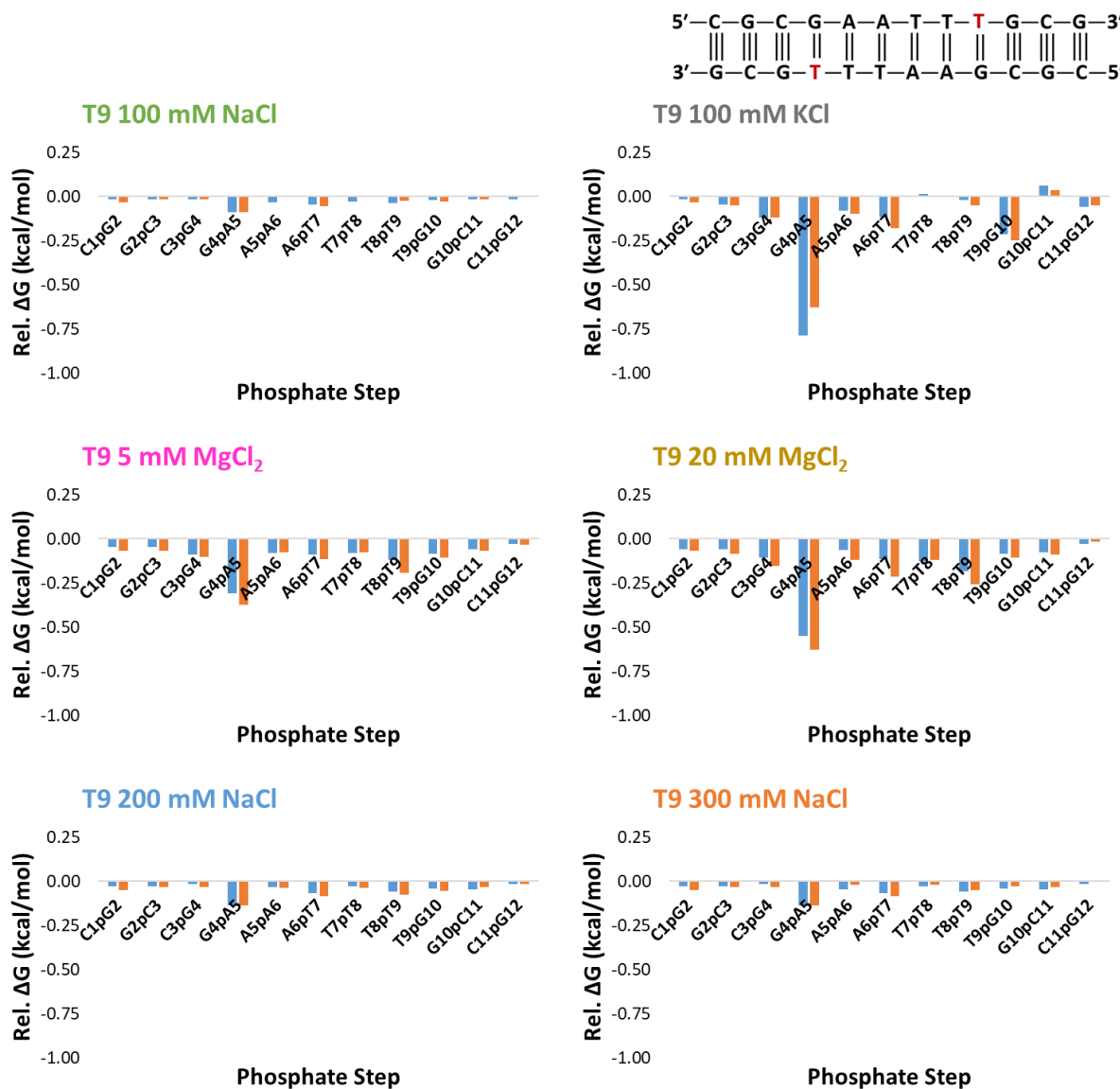
Figure 61. Relative ΔG plots for the T9 salt conditions at 283 (blue) and 298 K (orange). The relative ΔG is calculated by $\Delta G_{\text{25 mM Phosphate Buffer}} - \Delta G_{\text{varied sample condition}}$.
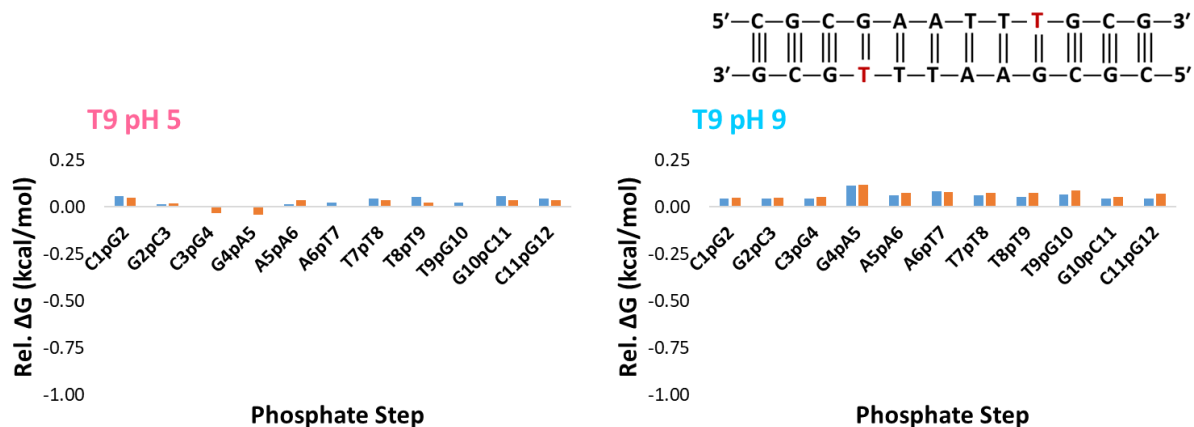
Figure 62. Relative ΔG plots for the T9 pH conditions at 283 (blue) and 298 K (orange). The relative ΔG is calculated by $\Delta G_{25 \text{ mM Phosphate Buffer}} - \Delta G_{\text{varied sample condition}}$.

In addition to the above sample conditions, 10% w/w and 20% w/w PEG-4000 was tested as a sample condition for T9. The unreferenced 1D ^{31}P spectra can be seen in Figures 63 and 64, respectively. As can be seen from these figures, when PEG-4000 was added to the DNA solution, the lines broadened qualitatively quite significantly with increasing concentration.
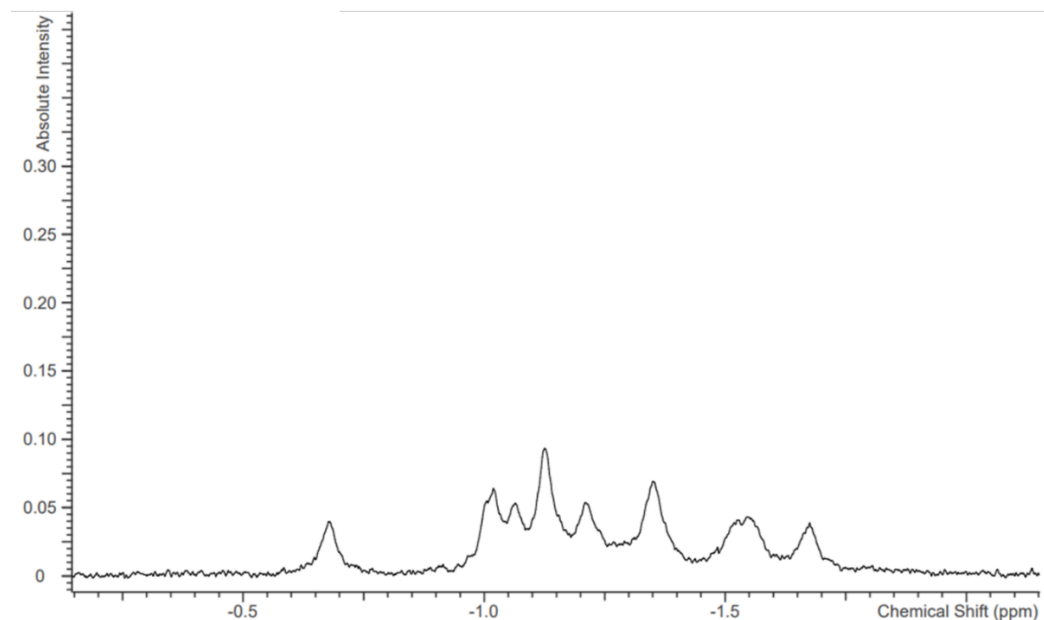
Figure 63. 1D ^{31}P spectrum of PEG-4000 10% w/w at 298 K for T9. This spectrum is unreferenced.
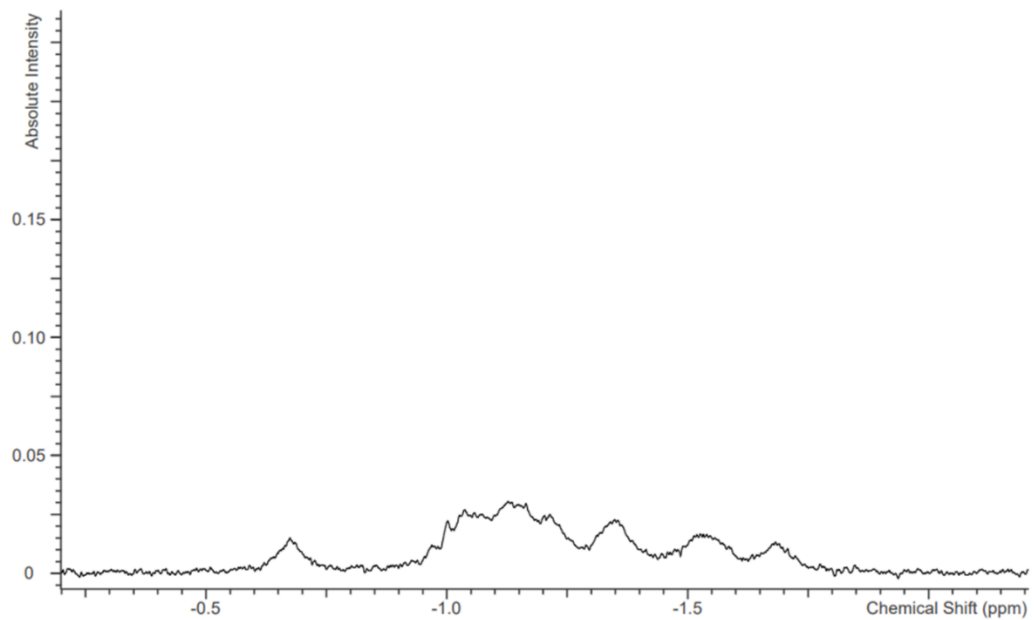
Figure 64. 1D ^{31}P spectrum of PEG-4000 20% w/w at 298 K for T9. This spectrum is unreferenced.

DISCUSSION

8mer Non-Palindromic DNA Sequences

The 8mer control and U4 Δ%BII and Relative ΔG in Figures 36 and 37 show a significant change in the population distribution of the backbone conformations and free energy associated with this exchange process around the lesion site and to a lesser extent around the base-pairing partner. This data is for non-palindromic DNA, which has not been investigated before when analyzing backbone conformation dynamics. Prior research has been published concerning palindromic DNA, which also found a stepwise difference around the lesion site and basepairing partner in the presence of a U:G mismatch.[22] A comparison of the palindromic and non-palindromic HSQC spectra can be seen in Figure 65.

As can be seen from Figure 65 in comparison of the palindromic to non-palindromic HSQC overlay data, the phosphate shifts around the lesion site within the non-palindromic 8mer sequences follows the exact trends of the palindromic data from Westwood et al.[22] The phosphorus 3′ to the lesion (U4pG5) shifts more downfield with respect to the control and the phosphorus 5′ to the lesion (G3pU4) shifts more upfield with respect to the control.

After calculating %BII, these perturbations correlate to a shift to more BII character at the U4pG5 step compared to the control (an average of a -13.0% ΔBII across all temperatures), and a decrease in the BII conformation population at the G3pU4 step compared to the control (an average of a 27.3% ΔBII across all temperatures). The relative ΔG plot showing a stepwise difference in energy around the lesion site and the subsequent base-pairing partner indicates an energetic perturbation to the phosphodiester backbone equilibrium conformational dynamics with a shift in the backbone conformation population distribution.

68
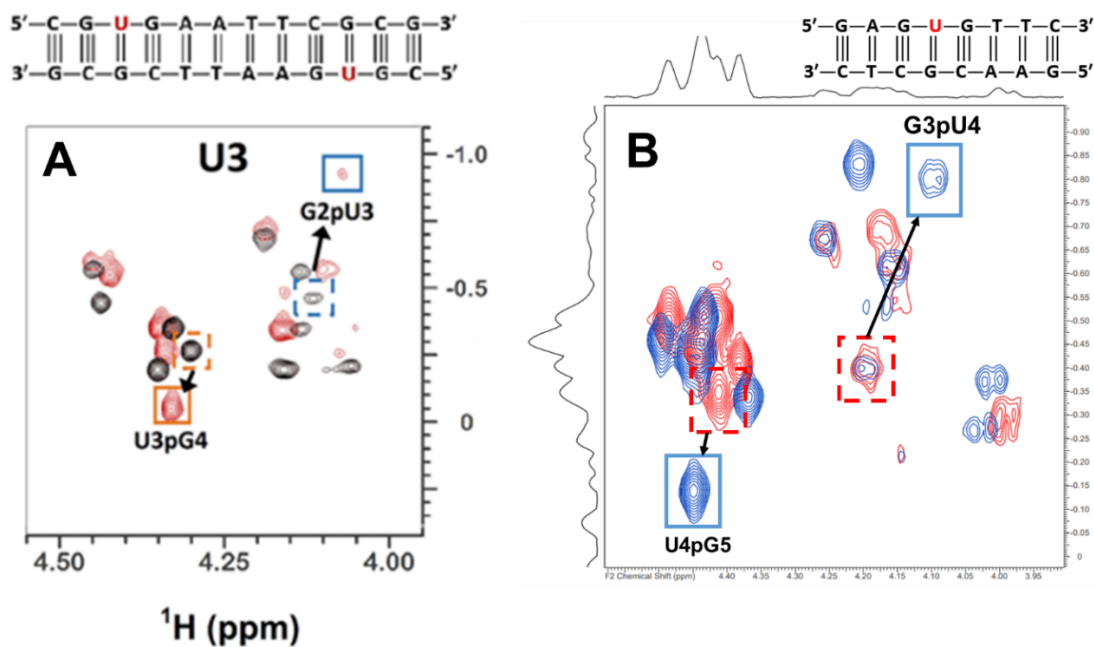
Figure 65. Comparison of the HSQC overlays between (A) palindromic dodecamer DNA containing a U:G mismatch at 298K and (B) HSQC overlay of the 8mer non-palindromic DNA at 283 K. The red spectrum in A is the lesion sequence and the black spectrum is the control sequence. In B the blue spectrum is the lesion sequence, and the red is the control sequence. The HSQC from Figure A was reprinted (adapted) with permission from Westwood, M. N.; Ljunggren, K. D.; Boyd, B.; Becker, J.; Dwyer, T. J.; Meints, G. A. Single-Base Lesions and Mismatches Alter the Backbone Conformational Dynamics in DNA. *Biochemistry*. **2021**, *60*, 873-885. Copyright 2021 American Chemical Society.

The fact that this stepwise difference and the perturbations in chemical shift around the uracil lesion in the non-palindromic data matches the trends observed in the palindromic data supports the idea that sequence context of the DNA farther from the lesion (2 bases away from the lesion) likely does not influence the perturbations observed at the phosphates around the lesion site. In addition, the perturbation to the free energy and the dynamics of the BI/BII interconversion around a lesion site could potentially be a way in which TDG recognizes a lesion to initiate BER. By understanding this process, future research could look to potentially enhance the ability of TDG to recognize single base lesions and mismatches.

In comparison, the 8mer FU4 sequence did not give as clear of results. As stated previously, many FU4 protons were missing within the NOESY spectrum. This resulted in the inability to assign the G3pFU4 phosphate in the HSQC and 1D ^{31}P spectra. In addition, the FU4pG5 step was dramatically shifted upfield with respect to the rest of the phosphates. This dramatic perturbation in the phosphorus chemical shift due to the presence of the fluorine atom would make the %BII calculations unreliable and not an accurate representation of the dynamics of the backbone. This is because rather than a change in chemical shift being due to the BI/BII conformational changes, the change in chemical shift is likely due to being around a highly electronegative atom, thus obscuring the actual dynamics that would be visualized normally, like what was observed with the 8mer U4 sequence. The highly electronegative nature of the fluorine atom could be the potential reason of this lack of reliable data because of how close the fluorine likely is to the other protons and phosphorus, impacting the chemical environment. The average distance based on a DDD sequence between the H2′ position (where the fluorine is located) and H1′, H4′, and the 3′ and 5′ phosphates is under 5 Å (3.01 Å, 3.95 Å, 3.78-3.98 Å, 4.30-4.93 Å, respectively. These values were obtained based off of NMR data published in the RCSB PDB[31] entry #1DUF). Since NMR is dependent on the chemical environment of the nuclei of interest, including its proximity to electronegative atoms, this proximity to the protons and phosphorus of interest may be the reason for this dramatic impact.

Future work would entail attempting to move the fluorine to the 5 position of the uracil base, creating 5-fluorouracil. An issue with this alteration is that 5-FU is an excellent substrate for TDG, with one study by Bennett et al. even finding that it is removed 78-fold faster than uracil by TDG.[14] For the 5-halogenated uracils, it is proposed that the electron withdrawing halogens enhance the leaving ability of uracil.[5] Because of this high activity, the 5-FU base

would no longer prevent cleavage by TDG. Because the ultimate end goal of this project is to complex the fluorinated DNA with TDG, the ability of TDG to excise 5-FU would be a detriment to the complexation portion of the project. It would potentially fix the proton and phosphorus dilemma and would allow for investigation of a 5-FU lesion on the phosphodiester backbone conformational dynamics, but once complexed with TDG it would be excised. If this route were to be taken, studies on mutating TDG in some way would need to be completed to make it to where TDG would not excise 5-FU and would stay as a complex. In addition, as will be discussed in the next section, the TDG optimization of expression and purification would need to be further investigated.

TDG Expression and Purification Optimization

The Methods section presents the start of an optimized protocol for the expression and purification of Thymine DNA Glycosylase and the second subsection within the Results presents some of the factors that had to be changed to get to that point. As evident from the numerous attempts of expressing and purifying this enzyme, TDG is a very particular protein. These attempts showed the temperature sensitivity of the expression conditions (the requirement of a 15°C induction for overnight) and the sensitivity to being stored in the -80°C. Often times when steps of the procedure, such as the resuspended cell pellet or lysate, were stored in the -80°C and then thawed to use on a later date, the quality of the sample worsened. For example, the texture of the cell pellet changed a couple times as well in the attempts. In addition to temperature, the concentration of IPTG was optimized, the optical density at induction was found to be ideal at 0.8, and the length of induction was regulated to only overnight rather than a weekend attempt.

For the purification protocol, the largest change was the removal of a CM ion-exchange column. The reason why after this column the yield was likely so negatively affected is because of the sensitivity of the enzyme to ionic strength. After further reading, an article had been published discussing sample conditions of TDG that mentioned that TDG is prone to aggregation in low ionic strength conditions.[32] The CM column used a wash buffer that contained no salt: 50 mM sodium phosphate, 10 mM β-mercaptoethanol, and a 7.4-7.5 pH. This no salt condition and the sensitivity of TDG to low ionic strength may be the reason why the yield was dramatically impacted after the use of the CM column.

Overall the proposed preliminary protocol has resulted in TDG potentially being present as the band at ≈25 kDa is present. What can be done to test this, if sufficient protein is obtained, is a DNA binding assay and/or obtain a crystal structure. What is interesting is that there are also bands present at 50 and 75 kDa. One potential hypothesis for this result, if it is not an impurity, is that TDG could possibly be dimer- and trimerizing.

Future aspects of the protocol that could be tested to optimize further is watching the length of the overnight culture, as sometimes there were black flecks that appeared in the overnight culture and the time was not followed strictly (it was generally just a guideline of "overnight"). Since the yield is still rather small after the altered procedure, the cell pellet should also be tested to see if there is overexpression within the cells that may not be extracted fully.

Influence of Sample Conditions in the Presence of Canonical and Lesioned DNA

As presented within the Results, the tested ions, ionic strength, ion size, and pH did not significantly change the phosphodiester backbone conformational dynamics at 283 and 298 K for DDD. The average changes in %BII for the DDD sequences across all of the sample conditions

was less than a $\pm2\%$ change at 283 and 298 K, except for 20 mM $MgCl_2$ and the pH 9 condition. In addition, the average changes in relative ΔG were less than a ±0.05 kcal/mol at 283 and 298 K for all sample conditions, except for 20 mM $MgCl_2$ and the pH 9 condition.

The sample conditions that most significantly changed the quality of the 1D ^{31}P spectra and the subsequent $\Delta\%BII$ and relative ΔG calculations were the 5 mM and 20 mM $MgCl_2$ samples for DDD, which appeared to show line broadening, and overall a slight change in the %BII and ΔG. If NMR peaks broaden, this may correlate to a slowing of the dynamic exchange.[23] In addition, a previous MD study found that Mg^{2+} within the major groove influences DNA backbone BI/BII conformational dynamics through intra-strand cross-links.[33] This slight change in the free energy across the board for the phosphates, with an increasing effect with increasing concentration, could correlate to this conclusion as well. Mg^{2+} has been known to neutralize the RNA backbone and thus is important in the folding and function of RNA.[34] Due to the similarities of RNA to DNA, potentially the influence of Mg^{2+} on the structure of DNA is an electrostatic interaction of the Mg^{2+} ion with the negatively charged backbone of the DNA, which could partially neutralize the negative charge and thus reduce the repulsion between the phosphates, leading to a slight shift in the BII conformation population. Other work has investigated Mg^{2+} site-binding to the phosphates of DNA and the stabilization impact it has on DNA.[35]

Besides $MgCl_2$ and slightly the pH 9 condition, and as can be seen from Figures 51-54 and as mentioned above, the %BII and ΔG barely changed across the different conditions for the canonical DDD DNA sequence. These small changes to the %BII and ΔG of the backbone across these conditions imply that DNA BI/BII backbone conformations may be resistant to differing solution conditions. This could be because of the varying solution conditions of cells, allowing

for DNA backbone conformations for canonical DNA to not be disturbed when solution conditions change. Also, if backbone dynamics correlate to enzymatic repair by TDG, this lack of a change across canonical DNA in different conditions may imply that TDG interaction with DNA is more direct and not dependent on solution conditions for canonical DNA.

In comparison, there were a few significant changes in the 1D ^{31}P spectra, %BII, and ΔG of the T9 samples in the presence of a varied sample condition as can be seen in Figures 59-62. The 100, 200, and 300 mM NaCl samples, and the pH 5 and pH 9 samples had only slight changes occur. The average changes in %BII for the T9 sequences across these sample conditions was less than a $\pm$2.5% change at 283 and 298 K. The average changes in ΔG for the T9 sequences across the NaCl and pH samples was less than a $\pm$0.07 kcal/mol change at 283 and 298 K. The only visible difference to the spectra was for pH 5 was two peaks became resolved within the middle of the spectrum at 283 and 298 K.

When MgCl$_2$ was introduced into the samples, there was a change in the %BII and ΔG at the phosphate 3′ to the lesion base-pairing partner (G4pA5). The change in %BII at G4pA5, where the most significant change occurred, for the 5mM MgCl$_2$ sample at 283 and 298 K was 3.95 and 5.05%, respectively. For the 20mM MgCl$_2$ the change in %BII at G4pA5 at 283 and 298 K was 5.92 and 7.21%, respectively. The change in ΔG at G4pA5, where the most significant change occurred, for the 5mM MgCl$_2$ sample at 283 and 298 K was -0.31 and -0.37 kcal/mol, respectively. For the 20mM MgCl$_2$ the change in ΔG at G4pA5 at 283 and 298 K was -0.55 and -0.63 kcal/mol, respectively. As can be seen from the spectra 1D ^{31}P stacks in Figure 55 and 56, these peaks shift significantly upfield.

In addition, 100 mM KCl also sees this trend occur to the most significant extent and a significant change 3′ to the lesion site (T9pG10). For the 100 mM KCl the change in %BII at

G4pA5 at 283 and 298 K was 7.24 and 7.21%, respectively. 100 mM KCl also saw a 7.24 and 7.21% change in %BII at 283 and 298 K, respectively for the T9pG10. For the 100 mM KCl the change in ΔG at G4pA5 at 283 and 298 K was -0.79 and -0.63 kcal/mol, respectively. 100 mM KCl also saw a -0.22 and -0.25 kcal/mol change in ΔG at 283 and 298 K, respectively for the T9pG10. This spectrum also sees an internal peak, G10pC11, shift in the opposite direction (downfield) with respect to the other peaks. This shows for 100 mM KCl there is not a universal shift in the peaks upfield, but there is also one that shifts more downfield. This is in comparison to the $MgCl_2$ samples that saw perturbed peaks shift upfield.

The trends observed with the KCl and $MgCl_2$ samples in the T9 sequence are extremely interesting considering where these changes occur: 3′ to the lesion and 3′ to the base-pairing partner. What makes this fact even more interesting is the lack of a change to the %BII and ΔG within the DDD sequences despite being in the same sample conditions as T9. This data shows that the phosphodiester backbone conformational dynamics is impacted when a lesion is present with these KCl or $MgCl_2$ sample conditions compared to 25 mM phosphate buffer. But this is not necessarily the case for a canonical DNA sequence. The reason for this change when KCl is present but not NaCl is unknown. A plausible reason, and something that can be investigated in the future, is a size influence. The ionic radius of Na^+ is 102 pm, while a potassium ion, K^+, is 138 pm.[36] Due to the larger radius of potassium ion and the more dramatic influence of this ion on the phosphodiester backbone conformational dynamics and the 1D ^{31}P spectra, the size could be allowing for a different interaction of the potassium ion with the phosphodiester backbone, potentially with the minor groove of the DNA. Future work should investigate the impact of size on the conformational dynamics of the same T9 sequence, going down the periodic table testing different alkali metal ions.

In addition to the conformational dynamics being influenced, the spectra quality also did change. The divalent ion Mg^{2+} broadened the spectra, as was observed for DDD. The 100 mM KCl resolved some of the peaks in comparison to the phosphate buffer samples and sodium chloride did in some cases as well (while other peaks became overlapped that were previously well resolved). Due to the improved resolution in places with NaCl and KCl and their relevance to cellular conditions, these salt conditions should potentially be considered to be used in future experiments. The reason for these changes is unknown. In addition, sample conditions with sodium phosphate and magnesium chloride, with no sodium chloride present, should also be investigated.

There were two additional solution properties that were tested: pH and PEG as an artificial crowding agent. TDG activity is dramatically reduced when below a pH of 5.5 and above 9.0 pH.[32] The %BII and ΔG changed only slightly at acidic (an average of a -0.96% change in %BII at 283 and 298 K) and basic (an average of a -2.26% change in %BII at 283 and 298K) pH ranges. When analyzing the results from PEG, it is evident that the presence of a crowding agent broadened the peaks with increasing concentration. Besides the broadening, there does not appear to be much change from a surface level analysis to the phosphate positions. The broadening of the peaks would make it difficult to use this condition in future experiments.

Conclusions

1D and 2D solution NMR Spectroscopy has proven to be an incredibly useful tool to analyze the dynamics of biomolecules, including DNA. By using ^{31}P NMR experiments, the phosphodiester backbone conformational dynamics can be quantified (%BII conformation and thus the free energy associated with the BI/BII backbone interconversion at each phosphate of a

DNA sequence) of short DNA sequences. When analyzing 8mer non-palindromic DNA containing a U:G mismatch, the presence of the uracil lesion resulted in a stepwise difference in the free energy around the lesion site. This follows data previously published analyzing palindromic DNA sequences.[22] The similarities in the trends between the palindromic and non-palindromic data implies that sequence context more than 2 bases away from the lesion site of interest does not influence the dynamics data, and that the phosphates flanking the lesion site and around the base pairing partner appear to have the largest perturbations to free energy when a lesion is present. This perturbation to the free energy of the BI/BII interconversion around the lesion site may be a way in which TDG recognizes a lesion. In addition, it was determined that fluorine at the arabino position of the deoxyribose sugar is not a feasible DNA construct to assign the protons and phosphorus, likely due to the electron dense nature of the fluorine and its proximity to the protons and phosphorus of interest.

For the work on hTDG expression and purification, it was determined that the expression and purification needs to be continued to be optimized in order to possibly be used within NMR and binding experiments. This enzyme was found to be sensitive to temperature, the OD induction, and length of expression. Future work on this project would entail optimizing the steps of expression and purification further.

In addition to the above two projects, the influence of sample conditions on DNA backbone conformational dynamics was investigated on palindromic DNA. For canonical DNA, sample conditions did not significantly influence the conformational dynamics. This could be due to the variety of environments within cells and being resilient to changes would prove to be useful. For DNA containing a T:G mismatch, there were significant changes to the free energy at the phosphate 3′ to the base-pairing partner of the thymine lesion when in the presence of

potassium ion, as well as magnesium ion. The lack of a difference in canonical DNA but the significant difference in lesioned DNA at a particular phosphate proves to be very interesting, potentially implying different interactions of potassium and magnesium ions with the backbone, such as with the minor or major groove, when a lesion is present. Future work will need to be completed to probe what these interactions may be and how the conformational dynamics are being influenced by these ions.

REFERENCES

(1) Matthews, C. K.; Van Holde, K. E.; Appling, D. R. *Biochemistry*, 4th ed. Pearson, 2013.

(2) Schärer, O. D. Chemistry and Biology of DNA Repair. *Angew. Chem. Int. Ed*. **2003**, *42*, 2946-2974.

(3) Friedberg, E. C.; Walker, G. C.; Siede, W.; Wood, R. D.; Schultz, R. A.; Ellenberger, T. *DNA Repair and Mutagenesis*, 2nd ed. ASM Press, 2006.

(4) Shen, J. C.; Rideout III, W. M.; Jones, P. A. The Rate of Hydrolytic Deamination of 5-Methylcytosine in Double-Stranded DNA. *Nucleic Acids Research* **1994**, *22* (6), 972-976.

(5) Morgan, M.T.; Bennett, M. T.; Drohat, A. C. Excision of 5-Halogenated Uracils by Human Thymine DNA Glycosylase: Robust Activity for DNA Contexts Other Than CpG. *Journal of Biological Chemistry* **2007**, *282* (38), 27578-27586.

(6) Maiti, A.; Morgan, M. T.; Pozharski, E.; Drohat, A. C. Crystal Structure of Human Thymine DNA Glyosylase Bound to DNA Elucidates Sequence-Specific Mismatch Recognition. *PNAS*. **2008**, *105* (26), 8890-8895.

(7) Schärer, O. D; Kawate, T.; Gallinari, P.; Jiricny, J.; Verdine, G. L. Investigation of the Mechanisms of DNA Binding of the Human G/T Glycosylase Using Designed Inhibitors. *Proc. Natl. Acad. Sci. USA*. **1997**, *94*, 4878-4883.

(8) Lindahl, T. Instability and Decay of the Primary Structure of DNA. *Nature* **1993**, *362*, 709-715.

(9) Hickson, I. D. *Base Excision Repair of DNA Damage*. Landes Bioscience and Chapman and Hall, 1997.

(10) Drohat, A. C.; Coey, C. T. Role of Base Excision "Repair" Enzymes in Erasing Epigenetic Marks from DNA. *Chem. Rev.* **2016**, *116* (20), 12711-12729.

(11) Stivers, J. T. Extrahelical Damaged Base Recognition by DNA Glycosylase Enzymes. *Chem. Eur. J.* **2008**, *14*, 786-793.

(12) Dow, B. J.; Malik, S. S.; Drohat, A. C. Defining the Role of Nucleotide Flipping in Enzyme Specificity Using ^{19}F NMR. *J. Am. Chem. Soc*. **2019**, *141*, 4952-4962.

(13) Hardeland, U.; Kunz, C.; Focke, F.; Szadkowski, M.; Schär, P. Cell Cycle Regulation as a Mechanism for Functional Separation of the Apparently Redundant Uracil DNA Glycosylases TDG and UNG2. *Nucleic Acids Research* **2007**, *35* (11), 3859-3867.

(14) Bennett, M. T.; Rodgers, M. T.; Herbert, A. S.; Ruslander, L. E.; Eisele, L.; Drohat, A. C. Specificity of Human Thymine DNA Glycosylase Depends on N-Glycosidic Bond Stability. *J. Am. Chem. Soc.* **2006**, *128*, 12510-12519.

(15) RCSB Protein Data Bank. PDB ID #5HF7. https://www.rcsb.org/structure/5HF7 (accessed 2023-07-14).

(16) Coey, C. T.; Malik, S. S.; Pidugu, L. S.; Varney, K. M.; Pozharski, E.; Drohat, A. C. Structural Basis of Damage Recognition by Thymine DNA Glycosylase: Key Roles for N-Terminal Residues. *Nucleic Acids Research* **2016**, *44* (21), 10248-10258.

(17) Teletchéa, S.; Hartmann, B.; Kozelka, J. Discrimination Between BI and BII Conformational Substates of B-DNA Based on Sugar-base Interproton Distances. *Journal of Biomolecular Structure and Dynamics* **2004**, *21* (4), 489-494.

(18) Gorenstein, D. G. Conformation and Dynamics of DNA and Protein-DNA Complexes by ^{31}P NMR. *Chem. Rev.* **1994**, *94*, 1315-1338.

(19) Wecker, K.; Bonnet, M. C.; Meurs, E. F.; Delepierre, M. The Role of the Phosphorus BI-BII Transition in Protein-DNA Recognition: The NF-kB Complex. *Nucleic Acids Research* **2002**, *30* (20), 4452-4459.

(20) Balaceanu, A.; Pasi, M.; Dans, P. D.; Hospital, A.; Lavery, R.; Orozco, M. The Role of Unconventional Hydrogen Bonds in Determining BII Propensities in B-DNA. *J. Phys. Chem. Lett.* **2017**, *8*, 21-28.

(21) Heddi, B.; Foloppe, N.; Bouchemal, N.; Hantz, E.; Hartmann, B. Quantification of DNA BI/BII Backbone States in Solution. Implications for DNA Overall Structure and Recognition. *J. Am. Chem. Soc.* **2006**, *128*, 9170-9177.

(22) Westwood, M. N.; Ljunggren, K. D.; Boyd, B.; Becker, J.; Dwyer, T. J.; Meints, G. A. Single-Base Lesions and Mismatches Alter the Backbone Conformational Dynamics in DNA. *Biochemistry* **2021**, *60*, 873-885.

(23) Cheung, T.; Ramesh, V. Biomolecular NMR Spectroscopy and Structure Determination of DNA. In *Biomolecular and Bioanalytical Techniques: Theory, Methodology, and Applications,* 1st ed.; John Wiley and Sons Ltd., 2019; pp 421-469.

(24) Nobel Prize in Chemistry Press Release 1991. https://www.nobelprize.org/prizes/chemistry/1991/press-release/ (accessed 2023-06-15).

(25) Nobel Prize in Physiology of Medicine Press Release 2003. https://www.nobelprize.org/prizes/medicine/2003/press-release/ (accessed 2023-06-15).

(26) Hare, D. R.; Wemmer, D. E.; Chou, S. H.; Drobny, G.; Reid, B. R. Assignment of the Non-Exchangeable Proton Resonances 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.

(27) Tian, Y.; Kayatta, M.; Shultis, K.; Gonzalez, A.; Mueller, L. J.; Hatcher, M. E. ^{31}P NMR Investigation of Backbone Dynamics in DNA Binding Sites. *J. Phys. Chem. B.* **2009**, *113* (9), 2596-2603.

(28) Barrett, T. E.; Schärer, O. D.; Savva, R.; Brown, T.; Jiricny, J.; Verdine, G. L.; Pearl, L. H. Crystal Structure of a Thwarted Mismatch Glycosylase DNA Repair Complex. *The EMBO Journal* **1999**, *18* (23), 6599-6609.

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

(30) Nakano, S.; Miyoshi, D.; Sugimoto, N. Effects of Molecular Crowding on the Structures, Interactions, and Functions of Nucleic Acids. *Chem. Rev.* **2014,** *114*, 2733-2758.

(31) RCSB Protein Data Bank. PDB ID #1DUF. https://www.rcsb.org/structure/1DUF (accessed 2023-07-14).

(32) Maiti, A.; Drohat, A. C. Dependence of Substrate Binding and Catalysis on pH, Ionic Strength, and Temperature for Thymine DNA Glycosylase: Insights into Recognition and Processing of G·T Mispairs. *DNA Repair (Amst).* **2011**, *10* (5), 545-553.

(33) Guéroult, M.; Boittin, O.; Mauffret, O.; Etchebest, C.; Hartmann, B. Mg^{2+} in the Major Groove Modulates B-DNA Structure and Dynamics. *PLoS ONE.* **2012**, *7* (7).

(34) Yamagami, R.; Sieg, J. P.; Bevilacqua, P. C. Functional Roles of Chelated Magnesium Ions in RNA Folding and Function. *Biochemistry* **2021**, *60* (31), 2374-2386.

(35) Bloomfield, V. A.; Crothers, D. M.; Tinoco, Jr., I. *Nucleic Acids: Structures, Properties and Functions.* University Science Books, 2000.

(36) Gilbert, T. R.; Kirss, R. V.; Foster, N. *Chemistry: An Atoms Focused Approach.* W. W. Norton and Company, Inc., 2014.

Appendix A: 8mer Control Supplemental Data

Appendix A-1: 8mer Control H1′ Front-End and Back-End Walks at 283 K. The

front-end walk is on the left, while the back-end walk is on the right.
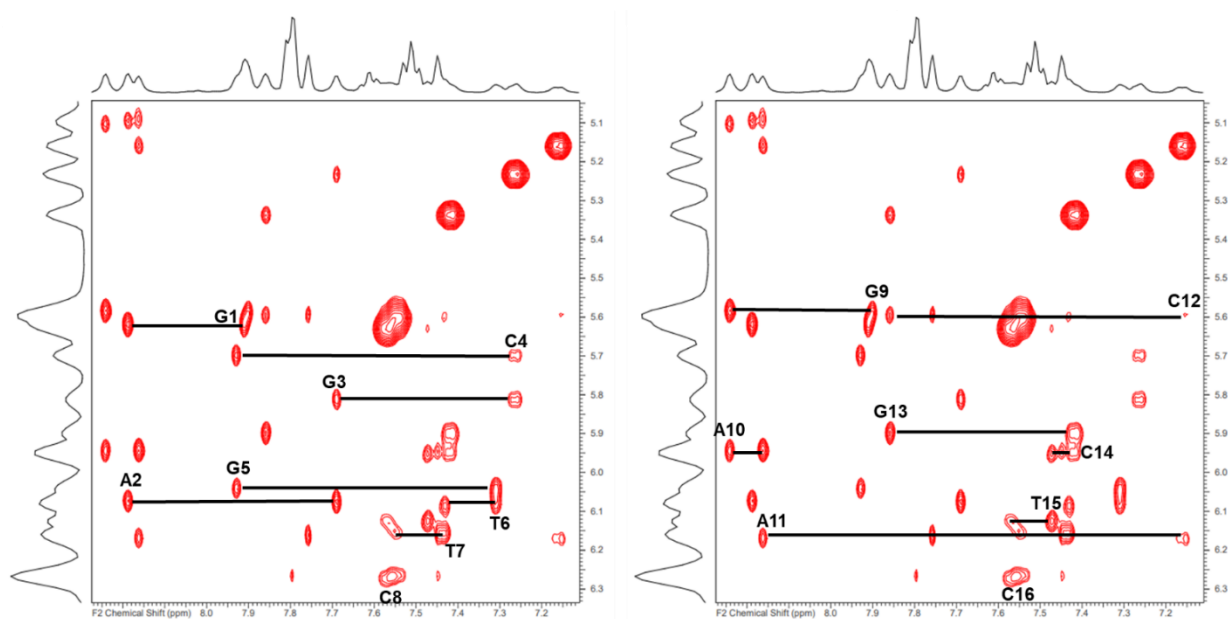
Appendix A-2: 8mer Control H2′ Front-End and Back-End Walks at 283 K. The

front-end walk is on the left, while the back-end walk is on the right.
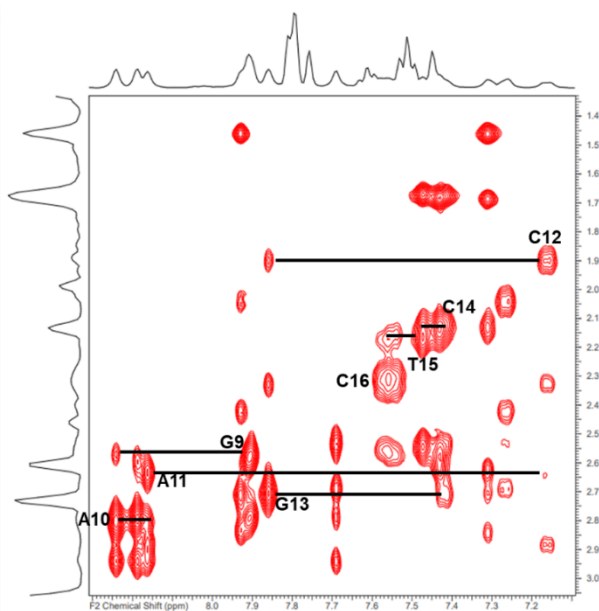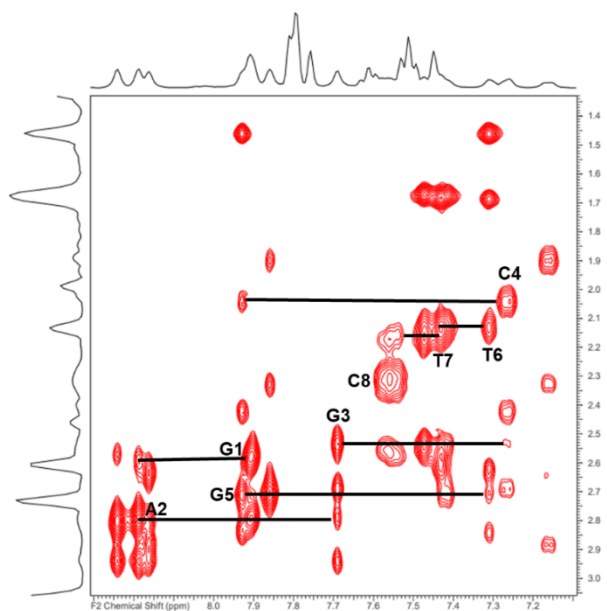
Appendix A-3: 8mer Control H2″ Front-End and Back-End Walks at 283 K. The

front-end walk is on the left, while the back-end walk is on the right.
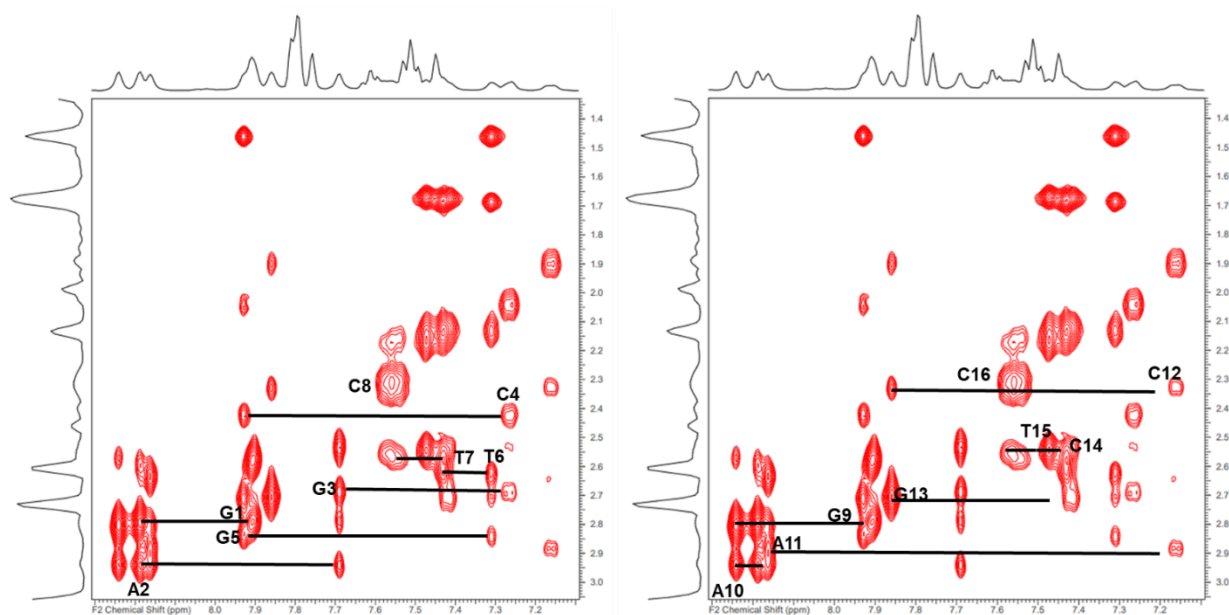
Appendix A-4: 8mer Control H3′ and H4′ Assignments at 283 K. The H3′

assignments are on the left spectrum and the H4′ assignments are on the right spectrum.
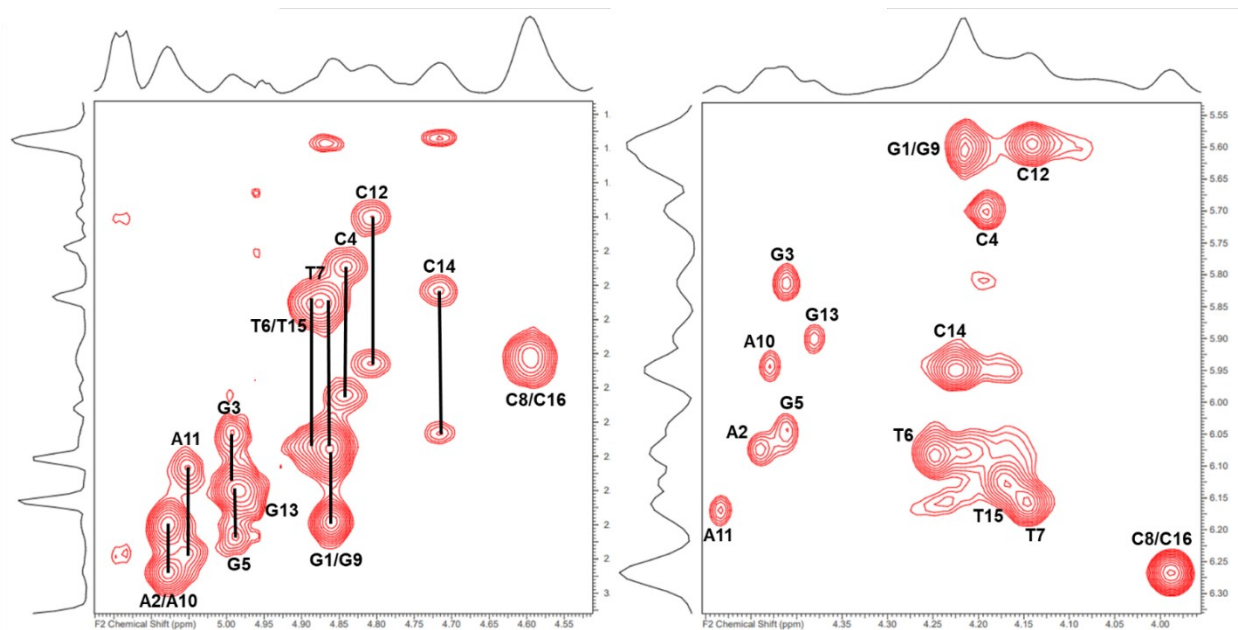
Appendix A-5: Table of Proton Chemical Shift Values for the 8mer Control at 283 K. All values are in parts per million (ppm).

Base	H1'	H2'	H2''	H3'	H4'	H6/H8	H5/methyl
G1	5.62	2.58	2.81	4.86	4.21	7.91	
A2	6.08	2.79	2.94	5.08	4.44	8.19	
G3	5.81	2.53	2.69	4.99	4.41	7.69	
C4	5.70	2.05	2.43	4.84	4.19	7.26	5.23
G5	6.04	2.71	2.84	4.99	4.41	7.93	
T6	6.08	2.13	2.62	n/a	4.25	7.31	1.46
T7	6.16	2.14	2.57	n/a	4.15	7.43	1.68
C8	6.27	2.31	2.31	4.59	3.99	7.54	5.60
G9	5.59	2.57	2.78	4.86	4.21	7.90	
A10	5.94	2.80	2.94	5.08	4.43	8.24	
A11	6.17	2.63	2.89	5.05	4.48	8.16	
C12	5.60	1.90	2.33	4.81	4.14	7.16	5.16
G13	5.90	2.71	2.71	4.98	4.38	7.86	
C14	5.95	2.12	2.54	n/a	4.22	7.42	5.34
T15	6.12	2.16	2.54	4.88	4.17	7.47	1.67
C16	6.27	2.31	2.31	4.59	3.99	7.57	5.63

Appendix A-6: Table of Proton Chemical Shift Values for the 8mer Control at 293 K. All values are in parts per million (ppm).

Base	H1'	H2'	H2''	H3'	H4'	H6/H8	H5/methyl
G1	5.61	2.55	2.75	4.85	4.20	7.90	
A2	6.07	2.80	2.94	5.08	4.44	8.20	
G3	5.82	2.55	2.69	5.00	4.42	7.70	
C4	5.73	2.05	2.43	4.84	4.20	7.27	5.25
G5	6.04	2.71	2.83	4.99	4.41	7.92	
T6	6.09	2.14	2.62	4.87	4.25	7.31	1.48
T7	6.19	2.19	2.58	4.90	4.17	7.45	1.71
C8	6.30	2.32	2.32	4.60	4.02	7.61	5.73
G9	5.58	2.53	2.73	4.85	4.20	7.89	
A10	5.95	2.82	2.94	5.08	4.44	8.25	
A11	6.17	2.64	2.88	5.06	4.48	8.16	
C12	5.61	1.91	2.34	4.81	4.15	7.17	5.18
G13	5.92	2.71	2.71	4.98	4.39	7.86	
C14	5.96	2.12	2.53	4.73	4.23	7.41	5.35
T15	6.16	2.21	2.55	4.90	4.18	7.49	1.69
C16	6.30	2.32	2.32	4.60	4.02	7.64	5.76

Appendix A-7: Table of Proton Chemical Shift Values for the 8mer Control at 298 K. All values in parts per million (ppm).

Base	H1'	H2'	H2''	H3'	H4'	H6/H8	H5/methyl
G1	5.57	2.50	2.68	4.84	4.17	7.88	
A2	6.04	2.78	2.90	5.07	4.43	8.19	
G3	5.80	2.53	2.66	4.99	4.41	7.69	
C4	5.71	2.02	2.41	4.84	4.20	7.27	5.22
G5	6.01	2.68	2.80	4.98	4.40	7.91	
T6	6.06	2.12	2.60	4.86	4.24	7.29	1.46
T7	6.18	2.18	2.55	4.91	4.17	7.46	1.70
C8	6.28	2.29	2.29	4.59	4.03	7.63	5.79
G9	5.53	2.47	2.68	4.84	4.18	7.87	
A10	5.93	2.79	2.91	5.07	4.43	8.24	
A11	6.14	2.62	2.84	5.05	4.47	8.15	
C12	5.59	1.89	2.32	4.80	4.15	7.17	5.16
G13	5.89	2.70	2.70	4.98	4.38	7.85	
C14	5.94	2.10	2.51	4.73	4.22	7.40	5.33
T15	6.14	2.20	2.52	4.89	4.18	7.50	1.68
C16	6.28	2.29	2.29	4.59	4.03	7.67	5.80

Appendix A-8: Table of Phosphorus Chemical Shift Values for 8mer Control. All
values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert,
at 278-298 K.

.Phosphate Step	278K	283K	288K	293K	298K
G1pA2	-0.53	-0.45	-0.41	-0.37	-0.37
A2pG3	-0.53	-0.50	-0.48	-0.45	-0.47
G3pC4	-0.44	-0.39	-0.38	-0.35	-0.37
C4pG5	-0.35	-0.35	-0.33	-0.32	-0.35
G5pT6	-0.71	-0.67	-0.63	-0.61	-0.63
T6pT7	-0.65	-0.62	-0.60	-0.58	-0.60
T7pC8	-0.30	-0.30	-0.27	-0.26	-0.27
G9pA10	-0.53	-0.45	-0.41	-0.37	-0.37
A10pA11	-0.53	-0.50	-0.48	-0.46	-0.48
A11pC12	-0.53	-0.54	-0.50	-0.47	-0.49
C12pG13	-0.44	-0.40	-0.38	-0.36	-0.37
G13pC14	-0.35	-0.32	-0.31	-0.31	-0.34
C14pT15	-0.71	-0.69	-0.66	-0.63	-0.64
T15pC16	-0.30	-0.30	-0.27	-0.26	-0.27

Appendix B: 8mer U4 Supplemental Data

Appendix B-1: 8mer U4 H1′ Front-End and Back-End Walks at 283 K. The front-end walk is on the left and the back-end is on the right.
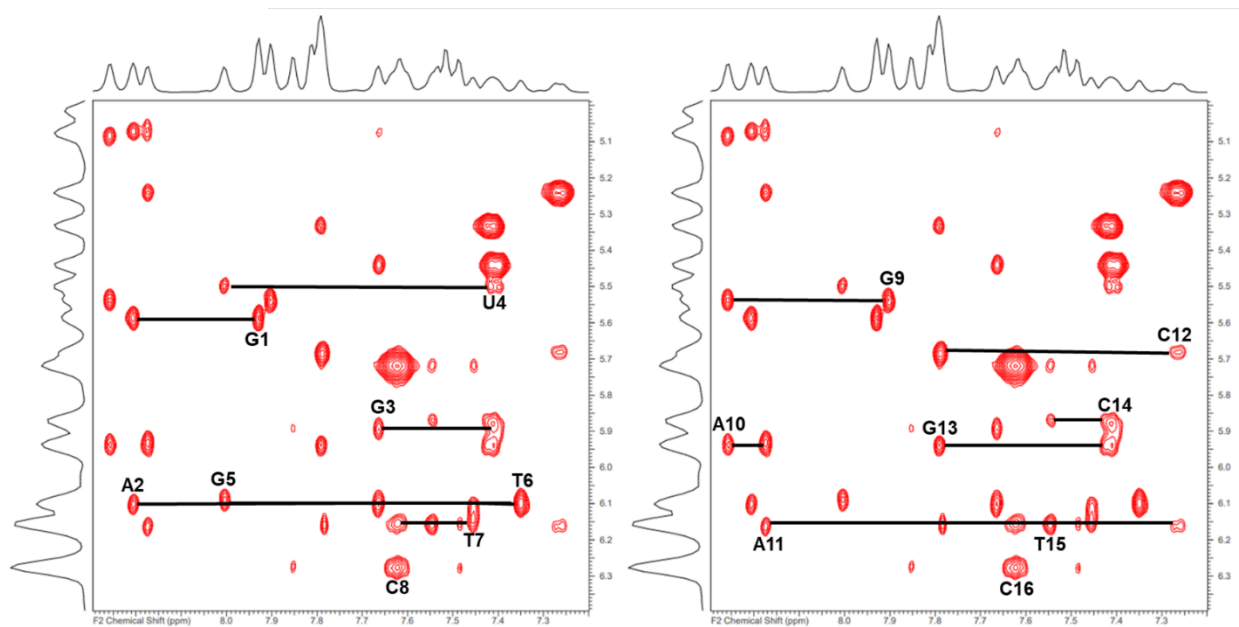
Appendix B-2: 8mer U4 H2′ Front-End and Back-End Walks at 283 K. The front-end walk is on the left and the back-end walk is on the right.
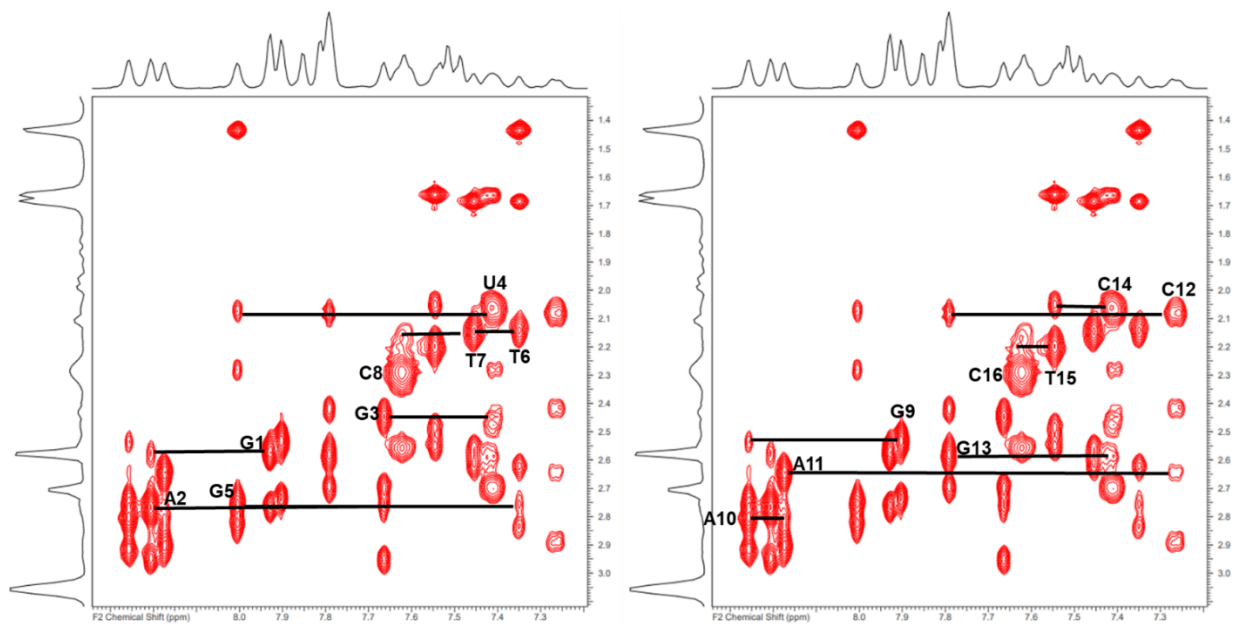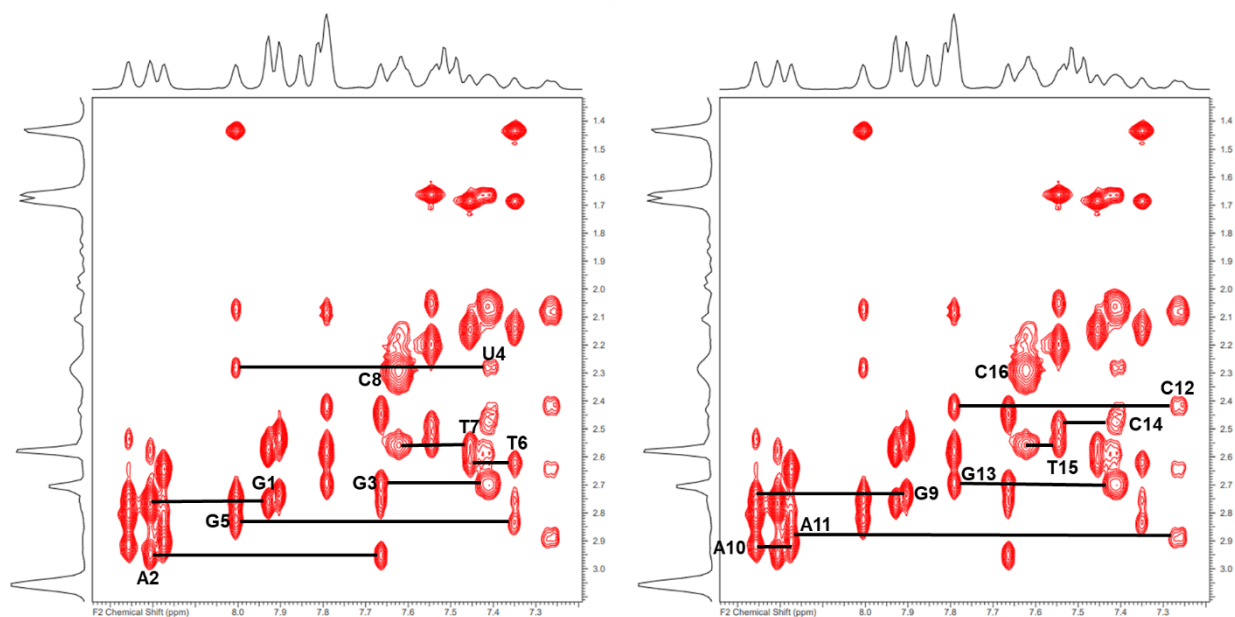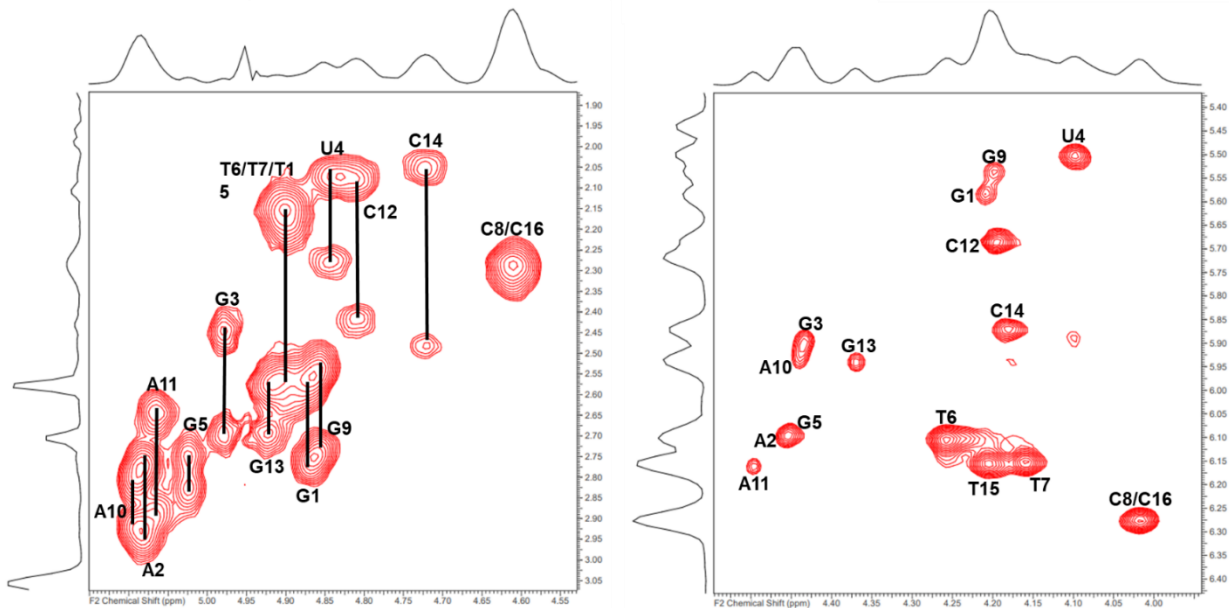
Appendix B-3: 8mer U4 H2″ Front-End and Back-End Walks at 283 K. The front-end walk is on the left and the back-end walk is on the right.

Appendix B-4: 8mer U4 H3′ and H4′ Assignments at 283 K. The H3′ spectrum assignments are on the left and the H4′ assignments are on the right.

Appendix B-5: Table of Proton Chemical Shift Values for the 8mer U4 at 283 K. All

values are in parts per million (ppm).

| Base	H1'	H2'	H2''	H3'	H4'	H6/H8	H5/methyl
G1	5.59	2.57	2.76	4.86	4.21	7.93	
A2	6.10	2.77	2.95	5.08	4.45	8.21	
G3	5.89	2.45	2.70	4.98	4.43	7.66	
U4	5.50	2.07	2.28	4.84	4.10	7.41	5.44
G5	6.09	2.76	2.83	5.02	4.45	8.00	
T6	6.10	2.14	2.62	4.90	4.26	7.35	1.43
T7	6.15	2.15	2.56	4.90	4.16	7.45	1.69
C8	6.28	2.29	2.29	4.61	4.02	7.62	5.72
G9	5.54	2.54	2.73	4.85	4.20	7.90	
A10	5.94	2.80	2.91	5.09	4.44	8.26	
A11	6.17	2.65	2.89	5.06	4.50	8.17	
C12	5.68	2.08	2.42	4.81	4.19	7.26	5.24
G13	5.94	2.59	2.70	4.92	4.37	7.79	
C14	5.88	2.05	2.48	4.72	4.18	7.41	5.33
T15	6.16	2.20	2.55	4.90	4.20	7.55	1.66
C16	6.28	2.29	2.29	4.61	4.02	7.62	5.72

Appendix B-6: Table of Phosphorus Chemical Shift Values for 8mer U4 at 278-298 K. The values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert. The T7pC8 and T15pC16 chemical shifts were arbitrarily assigned based on the two resolved peaks on the spectrum. Due to being on the ends and their distance from the lesion, this would not impact the other results obtained for %BII.

Phosphate Step	278K	283K	288K	293K	298K
G1pA2	-0.49	-0.43	-0.41	-0.38	-0.39
A2pG3	-0.54	-0.51	-0.50	-0.49	-0.51
G3pU4	-0.81	-0.80	-0.80	-0.78	-0.75
U4pG5	-0.16	-0.14	-0.13	-0.14	-0.17
G5pT6	-0.70	-0.67	-0.65	-0.64	-0.63
T6pT7	-0.63	-0.61	-0.60	-0.59	-0.59
T7pC8	-0.27	-0.27	-0.26	-0.25	-0.25
G9pA10	-0.49	-0.42	-0.41	-0.38	-0.39
A10pA11	-0.49	-0.46	-0.45	-0.45	-0.48
A11pC12	-0.42	-0.40	-0.39	-0.37	-0.38
C12pG13	-0.37	-0.34	-0.33	-0.34	-0.32
G13pC14	-0.54	-0.53	-0.50	-0.49	-0.51
C14pT15	-0.86	-0.83	-0.80	-0.78	-0.75
T15pC16	-0.37	-0.37	-0.35	-0.34	-0.32

Appendix C: 8mer FU4 Supplemental Data

Appendix C-1: 8mer FU4 H2′ Front-End and Back-End Walks at 283 K. The front-end walk is on the left and the back-end walk is on the right. Since the fluorine is located at the araribno position, the FU4 H2' proton would not have an assignment.
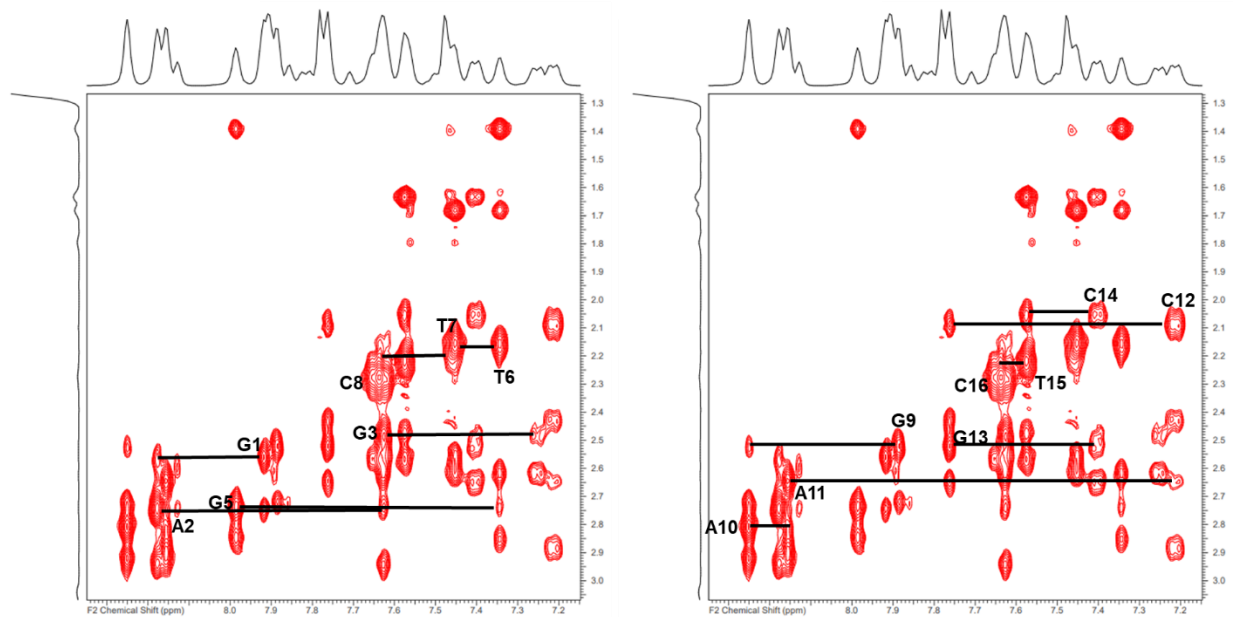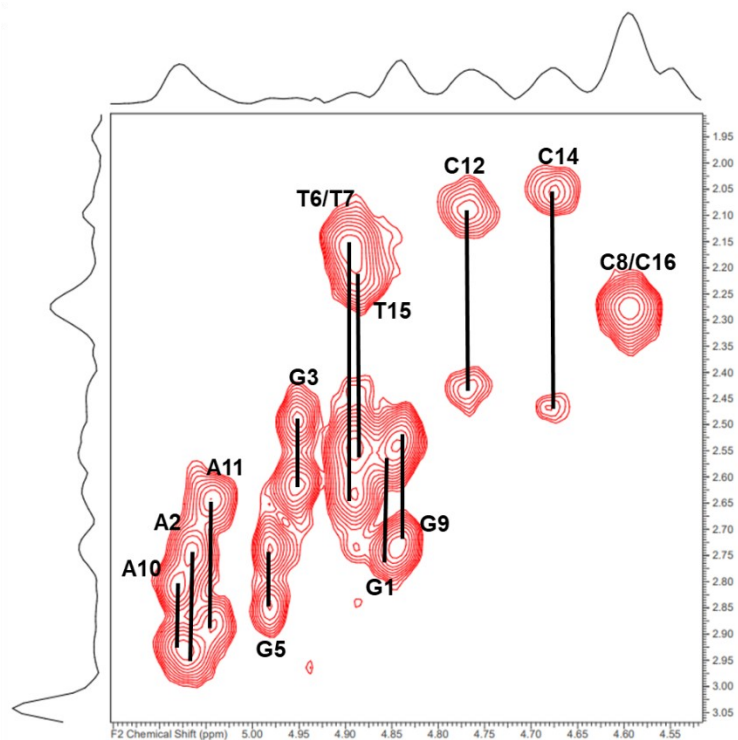
Appendix C-2: 8mer FU4 H2″ Front-End and Back-End Walks at 283 K. The front-end walk is on the left and the back-end walk is on the right. The FU4 H2″ proton could not be assigned due to other missing peaks.
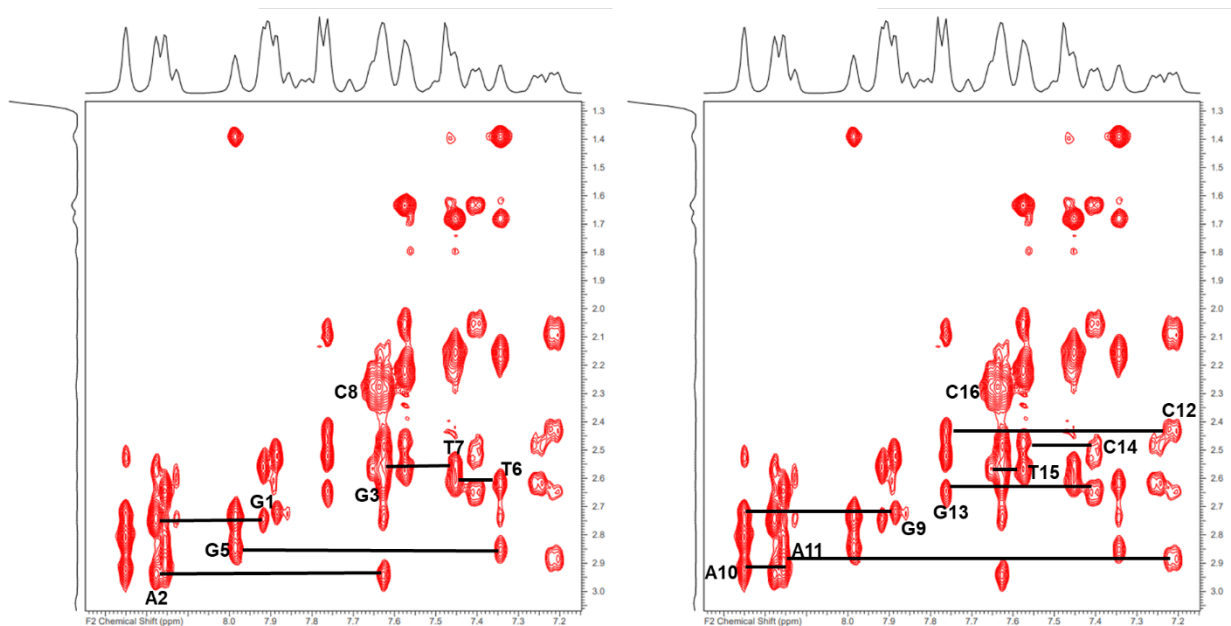
Appendix C-3: 8mer FU4 H3' Assignments at 283 K. Notice the FU4 H3′ proton could not be assigned.

Appendix C-4: Table of Proton Chemical Shift Values for the 8mer FU4 at 278 K.

All values are in parts per million (ppm).

Base	H1'	H2'	H2''	H3'	H4'	H6/H8	H5/methyl
G1	5.58	2.58	2.77	4.86	4.19	7.93	
A2	6.15	2.73	2.95	5.07	4.45	8.18	
G3	5.86	2.49	2.60	4.95	4.43	7.64	
FU4	n/a	n/a	n/a	n/a	n/a	7.25	5.55
G5	6.09	2.74	2.85	4.99	4.43	7.99	
T6	6.13	2.16	2.62	4.89	4.25	7.35	1.39
T7	6.14	2.15	2.57	4.89	4.14	7.46	1.68
C8	6.28	2.28	2.28	4.60	4.00	7.62	5.73
G9	5.54	2.55	2.75	4.84	4.18	7.90	
A10	5.96	2.81	2.93	5.08	4.43	8.26	
A11	6.17	2.65	2.90	5.04	4.49	8.17	
C12	5.69	2.10	2.43	4.77	4.18	7.21	5.21
G13	5.91	2.52	2.65	n/a	4.30	7.77	
C14	5.83	2.06	2.48	4.68	4.15	7.41	5.27
T15	6.18	2.22	2.57	4.88	4.21	7.58	1.63
C16	6.28	2.28	2.28	4.60	4.00	7.62	5.73

Appendix C-5: Table of proton chemical shift values for the 8mer FU4 at 283 K. All values are in parts per million (ppm).

| Base	H1'	H2'	H2''	H3'	H4'	H6/H8	H5/methyl
G1	5.57	2.56	2.74	4.85	4.18	7.92	
A2	6.14	2.74	2.94	5.07	4.44	8.18	
G3	5.87	2.48	2.61	4.95	4.43	7.63	
FU4	n/a	n/a	n/a	n/a	n/a	7.25	5.56
G5	6.09	2.74	2.85	4.98	4.43	7.98	
T6	6.13	2.16	2.62	4.89	4.25	7.34	1.39
T7	6.14	2.16	2.56	4.89	4.15	7.45	1.68
C8	6.29	2.28	2.28	4.59	4.01	7.64	5.77
G9	5.53	2.52	2.72	4.84	4.17	7.89	
A10	5.96	2.81	2.92	5.08	4.43	8.25	
A11	6.16	2.64	2.89	5.04	4.48	8.16	
C12	5.70	2.09	2.43	4.77	4.18	7.21	5.21
G13	5.91	2.53	2.65	n/a	4.30	7.76	
C14	5.83	2.06	2.49	4.68	4.16	7.40	5.27
T15	6.18	2.22	2.56	4.89	4.21	7.57	1.63
C16	6.29	2.28	2.28	4.59	4.01	7.64	5.77

Appendix C-6: Table of Proton Chemical Shift Values for the 8mer FU4 at 288 K.

All values are in parts per million (ppm).

Base	H1'	H2'	H2''	H3'	H4'	H6/H8	H5/methyl
G1	5.57	2.54	2.72	4.84	4.17	7.91	
A2	6.13	2.74	2.94	5.06	4.44	8.17	
G3	5.86	2.49	2.61	4.95	4.43	7.62	
FU4	n/a	n/a	n/a	n/a	n/a	7.26	5.56
G5	6.08	2.73	2.84	4.98	4.41	7.98	
T6	6.13	2.16	2.61	4.89	4.25	7.33	1.39
T7	6.15	2.20	2.55	4.89	4.15	7.45	1.69
C8	6.29	2.28	2.28	4.59	4.02	7.65	5.80
G9	n/a	2.50	2.71	4.83	4.16	7.87	
A10	5.95	2.80	2.92	5.08	4.43	8.24	
A11	6.15	2.64	2.88	5.04	4.48	8.15	
C12	5.70	2.08	2.43	4.76	4.17	7.21	5.21
G13	5.90	2.52	2.65	n/a	4.30	7.76	
C14	5.84	2.06	2.47	4.68	4.15	7.40	5.27
T15	6.18	2.23	2.56	4.88	4.20	7.57	1.64
C16	6.29	2.28	2.28	4.59	4.02	7.65	5.80

Appendix D: DDD Supplemental Data

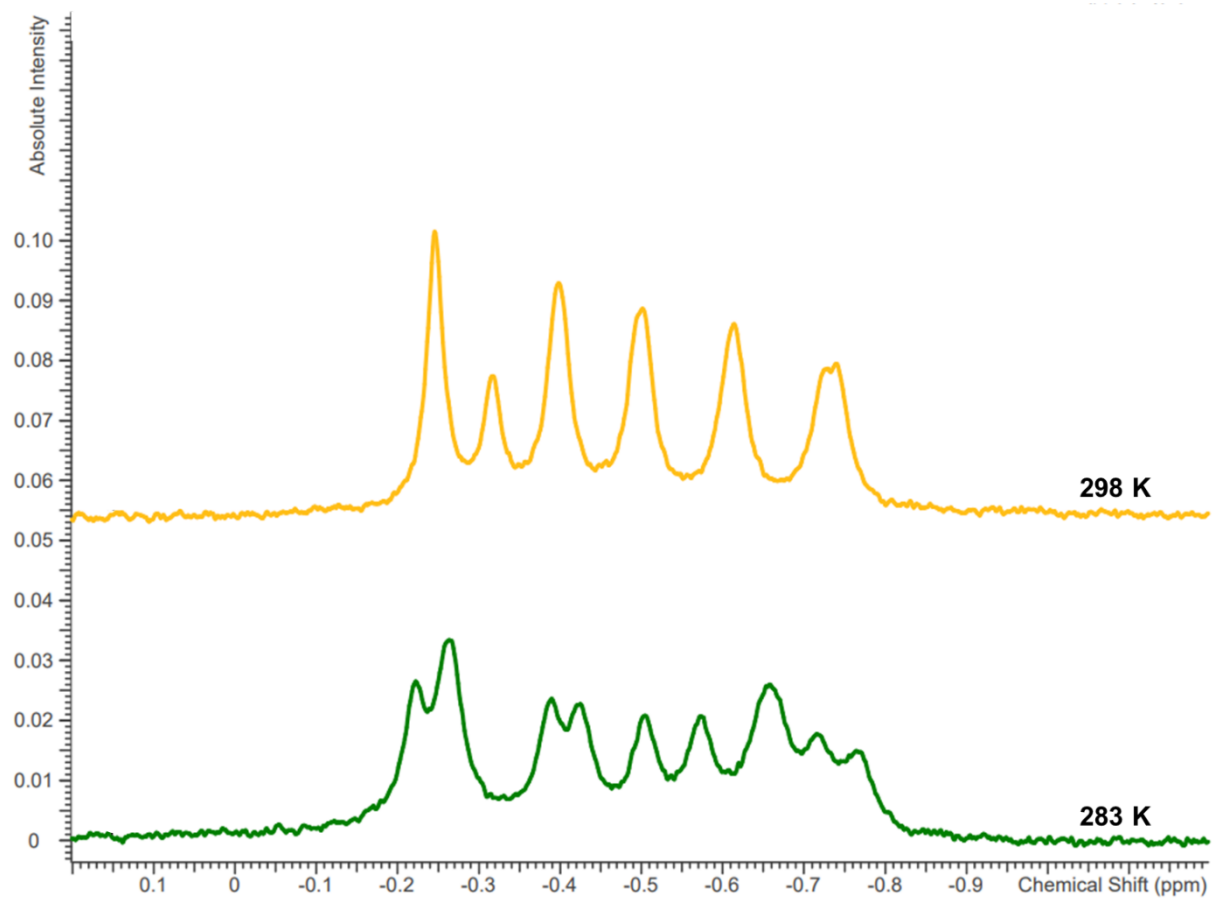**Appendix D-1: Stacked 1D ^{31}P spectra at 283 and 298 K of DDD in 25 mM phosphate buffer.**
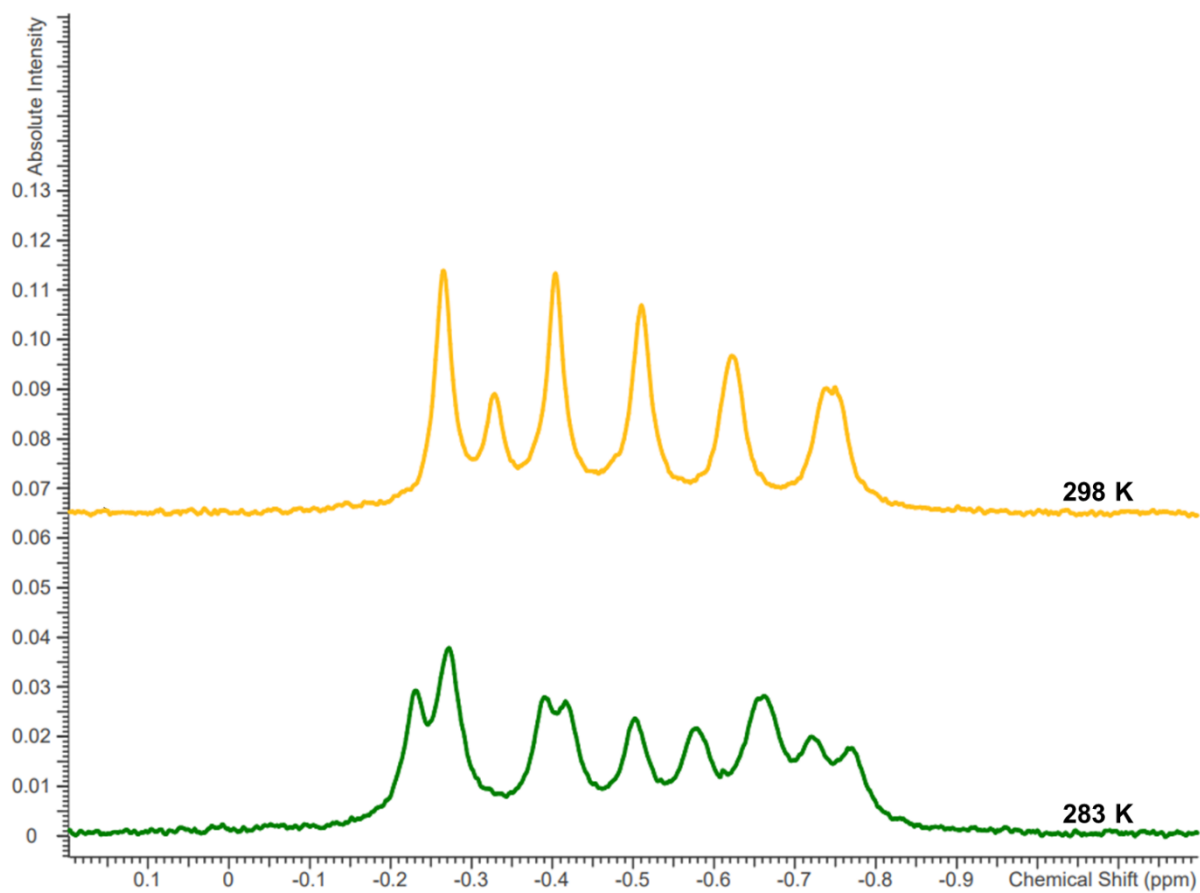
Appendix D-2: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer DDD Sample Condition. The values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.42	-0.40
G2pC3	-0.50	-0.51
C3pG4	-0.27	-0.33
G4pA5	-0.58	-0.51
A5pA6	-0.66	-0.62
A6pT7	-0.77	-0.74
T7pT8	-0.72	-0.74
T8pC9	-0.66	-0.62
C9pG10	-0.23	-0.27
G10pC11	-0.39	-0.40
C11pG12	-0.27	-0.27

Appendix D-3: Stacked 1D ^{31}P Spectra at 283 and 298 K of DDD in 25 mM Phosphate Buffer and 100 mM NaCl.
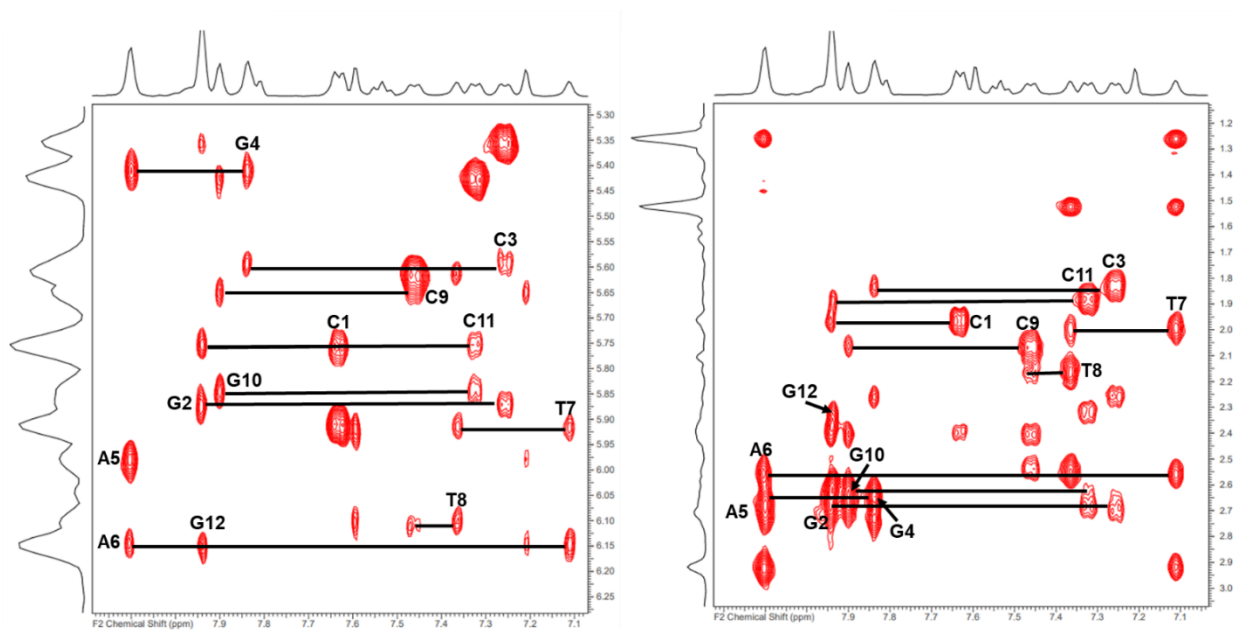
Appendix D-4: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer and 100 mM NaCl DDD Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

| Phosphate Step	283K	298K
C1pG2	-0.42	-0.40
G2pC3	-0.50	-0.50
C3pG4	-0.26	-0.32
G4pA5	-0.57	-0.50
A5pA6	-0.66	-0.61
A6pT7	-0.77	-0.74
T7pT8	-0.72	-0.74
T8pC9	-0.66	-0.61
C9pG10	-0.22	-0.25
G10pC11	-0.39	-0.40
C11pG12	-0.26	-0.25

Appendix D-5: DDD 100 mM KCl Sample Condition H1′ and H2′ Walks at 298 K.

The H1′ walk is on the left and the H2′ walk is on the right.
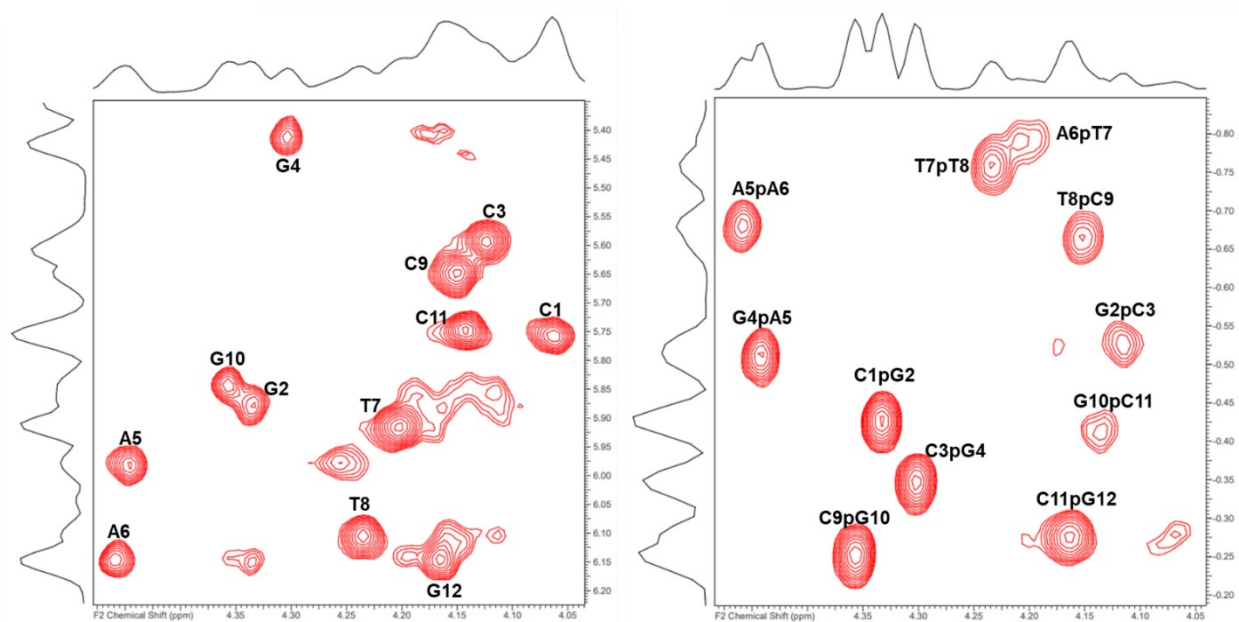
Appendix D-6: DDD 100 mM KCl Sample Condition H2″ Walk and H3′

Assignments at 298 K. The H2″ walk is on the left and the H3′ assignments are on the right.
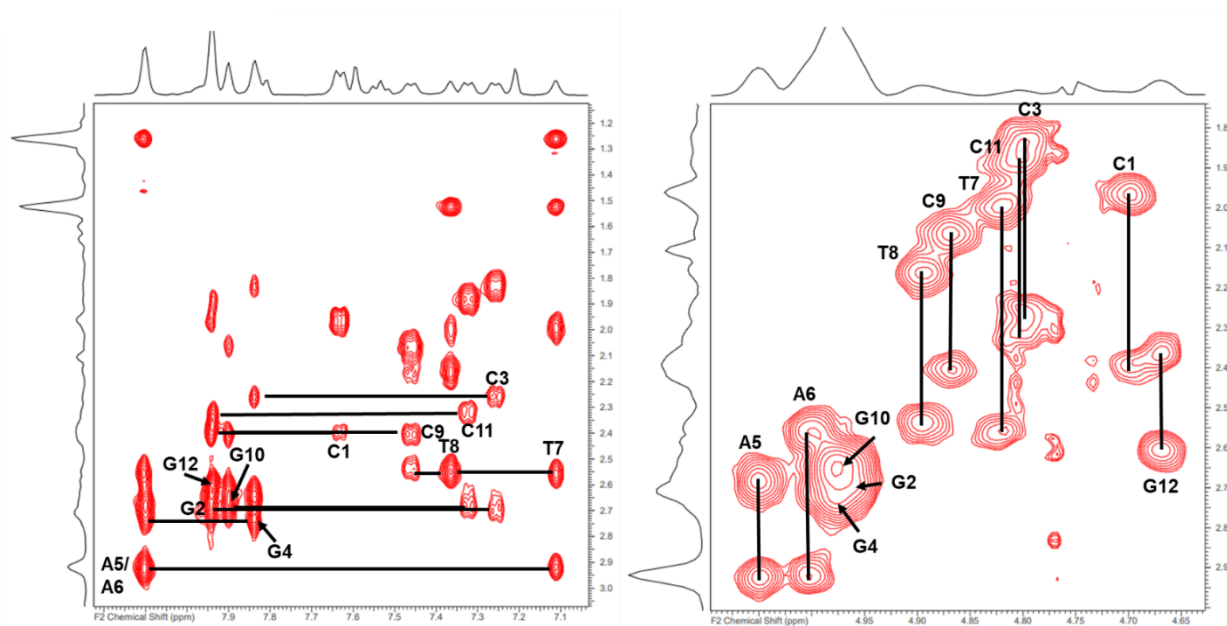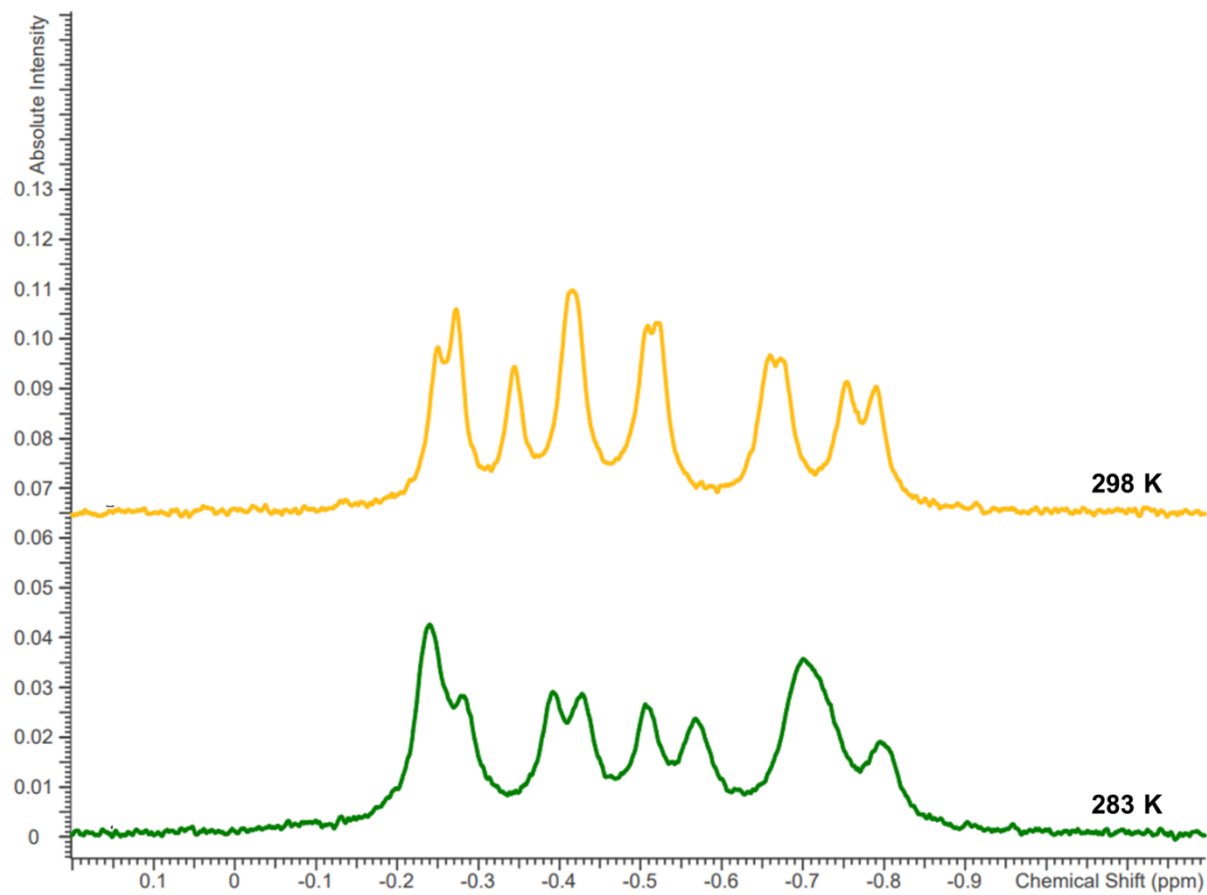
Appendix D-7: DDD 100 mM KCl Sample Condition H4′ and HSQC Assignments at

298 K. The H4′ assignments are on the left and the HSQC assignments are on the right.

Appendix D-8: Stacked 1D ^{31}P Spectra at 283 and 298 K of DDD in 25 mM Phosphate Buffer and 100 mM KCl.

Appendix D-9: Table of Proton Chemical Shift Values for 25 mM Phosphate Buffer and 100 mM KCl DDD Sample Condition at 298 K. All values are in parts per million (ppm).

Base	H1'	H2'	H2''	H3'	H4'	H6/H8	H5/methyl
C1	5.76	1.97	2.40	4.70	4.06	7.63	5.91
G2	5.87	2.64	2.70	4.96	4.33	7.94	
C3	5.59	1.83	2.26	4.80	4.12	7.26	5.35
G4	5.41	2.65	2.74	4.98	4.30	7.84	
A5	5.98	2.69	2.92	5.05	4.45	8.10	
A6	6.14	2.56	2.92	5.00	4.46	8.10	
T7	5.92	1.99	2.56	4.82	4.20	7.11	1.26
T8	6.10	2.16	2.55	4.89	4.24	7.36	1.52
C9	5.65	2.07	2.41	4.87	4.15	7.46	5.62
G10	5.84	2.64	2.68	4.98	4.36	7.90	
C11	5.75	1.88	2.32	4.81	4.14	7.32	5.43
G12	6.15	2.34	2.62	4.67	4.16	7.94	

Appendix D-10: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer and 100 mM KCl DDD Sample Condition. The values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.43	-0.43
G2pC3	-0.51	-0.53
C3pG4	-0.28	-0.35
G4pA5	-0.57	-0.51
A5pA6	-0.70	-0.68
A6pT7	-0.80	-0.79
T7pT8	-0.70	-0.76
T8pC9	-0.70	-0.66
C9pG10	-0.24	-0.25
G10pC11	-0.39	-0.41
C11pG12	-0.24	-0.27

Appendix D-11: Stacked 1D ^{31}P Spectra at 283 and 298 K of DDD in 25 mM Phosphate Buffer, 100 mM NaCl and 5 mM MgCl$_2$.
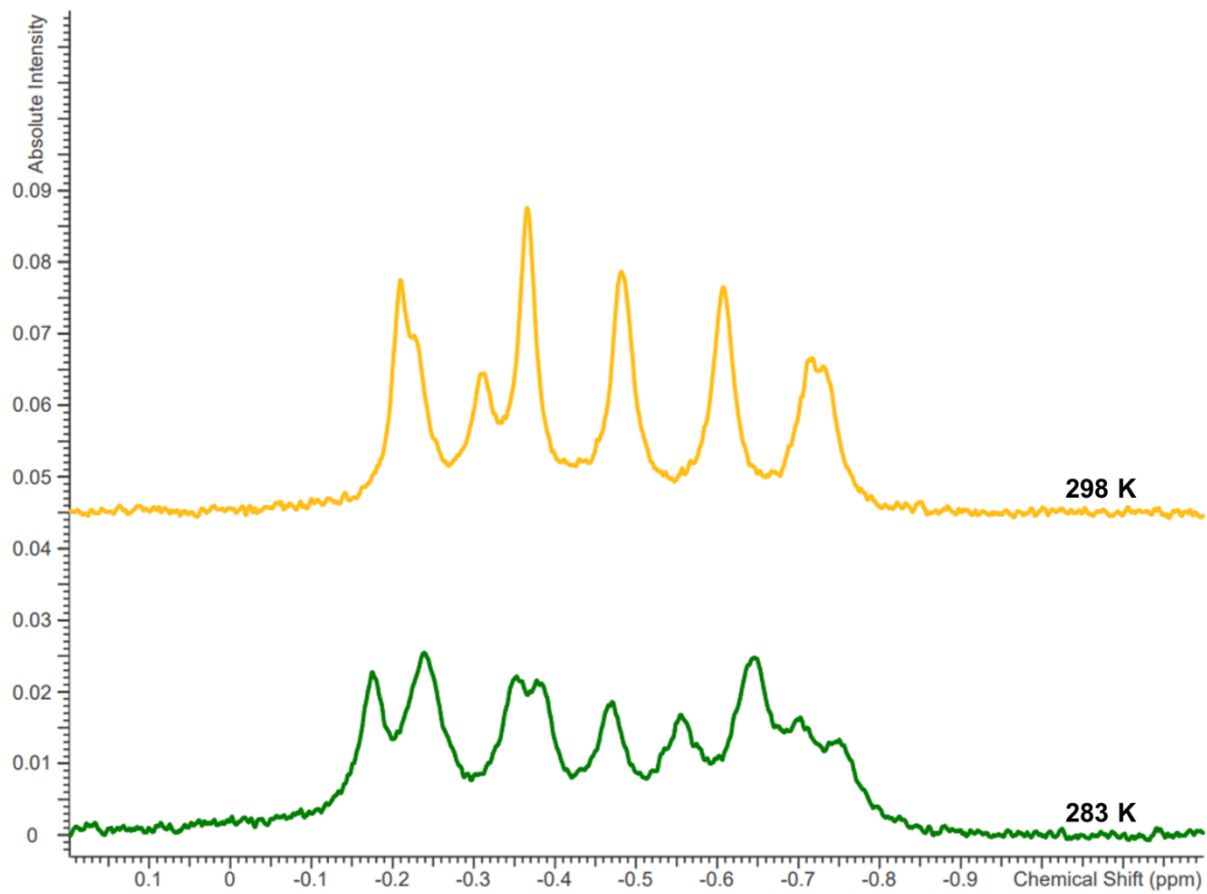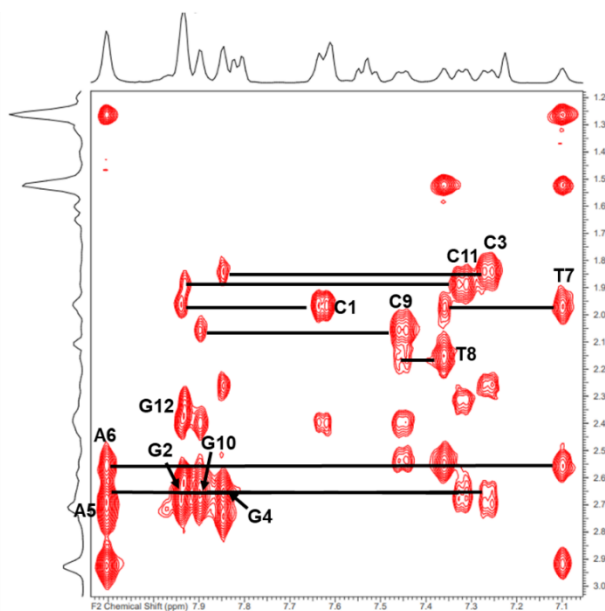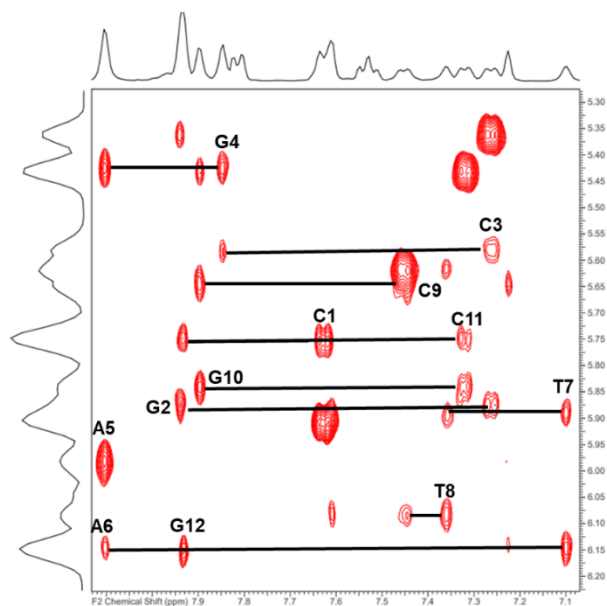
Appendix D-12: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer, 100 mM NaCl, and 5 mM MgCl₂ DDD Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K. These values should be double checked with a new set of coax experiments.

| Phosphate Step	283K	298K
C1pG2	-0.38	-0.37
G2pC3	-0.47	-0.48
C3pG4	-0.24	-0.31
G4pA5	-0.56	-0.48
A5pA6	-0.65	-0.61
A6pT7	-0.75	-0.73
T7pT8	-0.70	-0.72
T8pC9	-0.65	-0.61
C9pG10	-0.24	-0.23
G10pC11	-0.35	-0.37
C11pG12	-0.18	-0.21

Appendix D-13: DDD 20 mM MgCl₂ Sample Condition H1′ and H2′ Walks at 298 K.

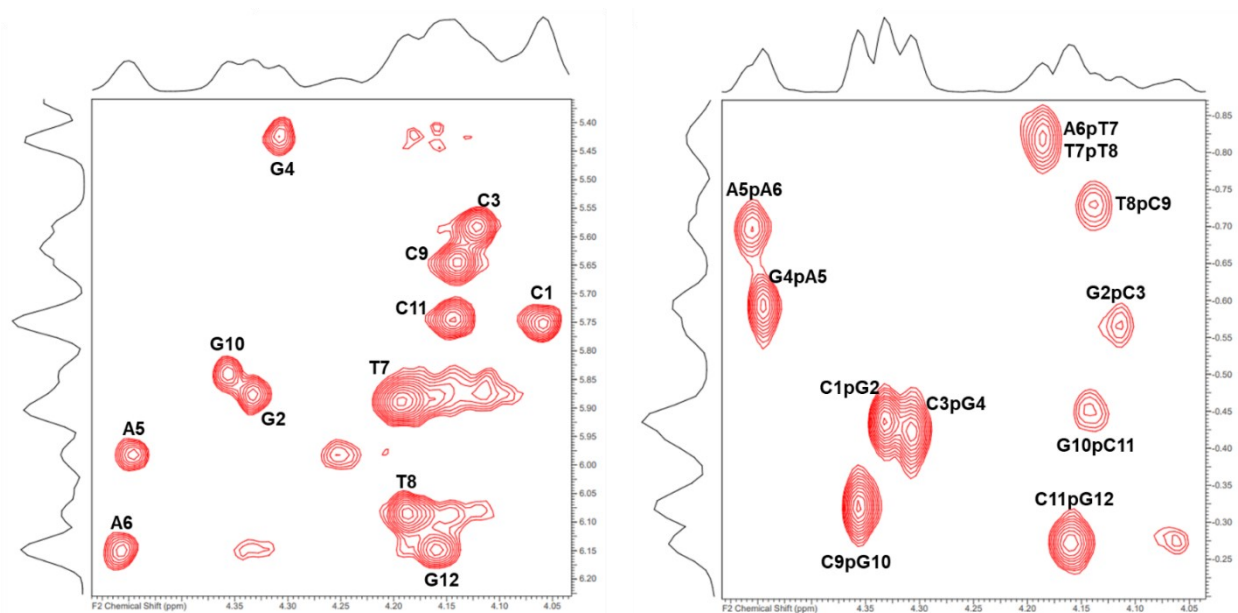The H1′ walk is on the left and the H2′ walk is on the right.

Appendix D-14: DDD 20 mM MgCl₂ Sample Condition H2″ Walk and H3′

Assignments at 298 K. The H2″ walk is on the left and the H3′ assignments are on the right.
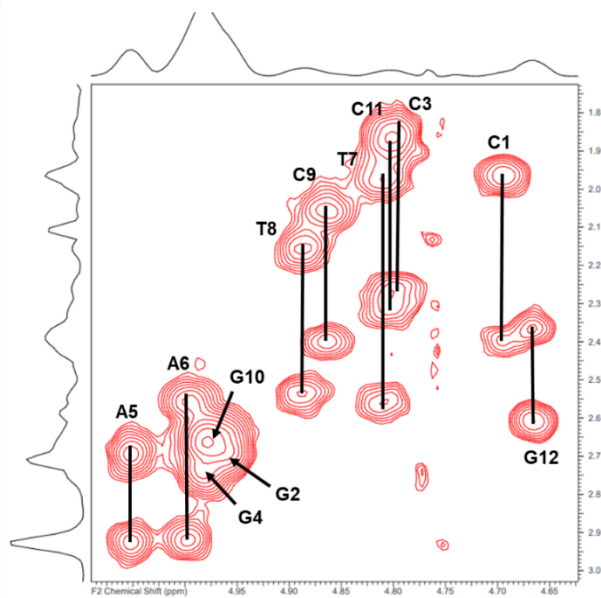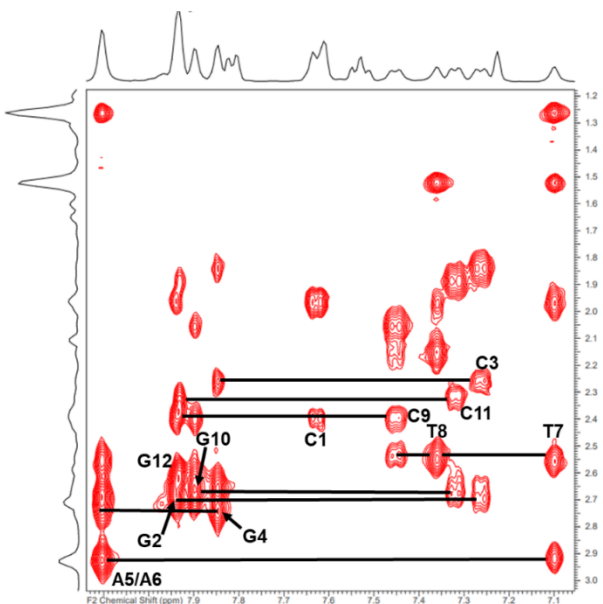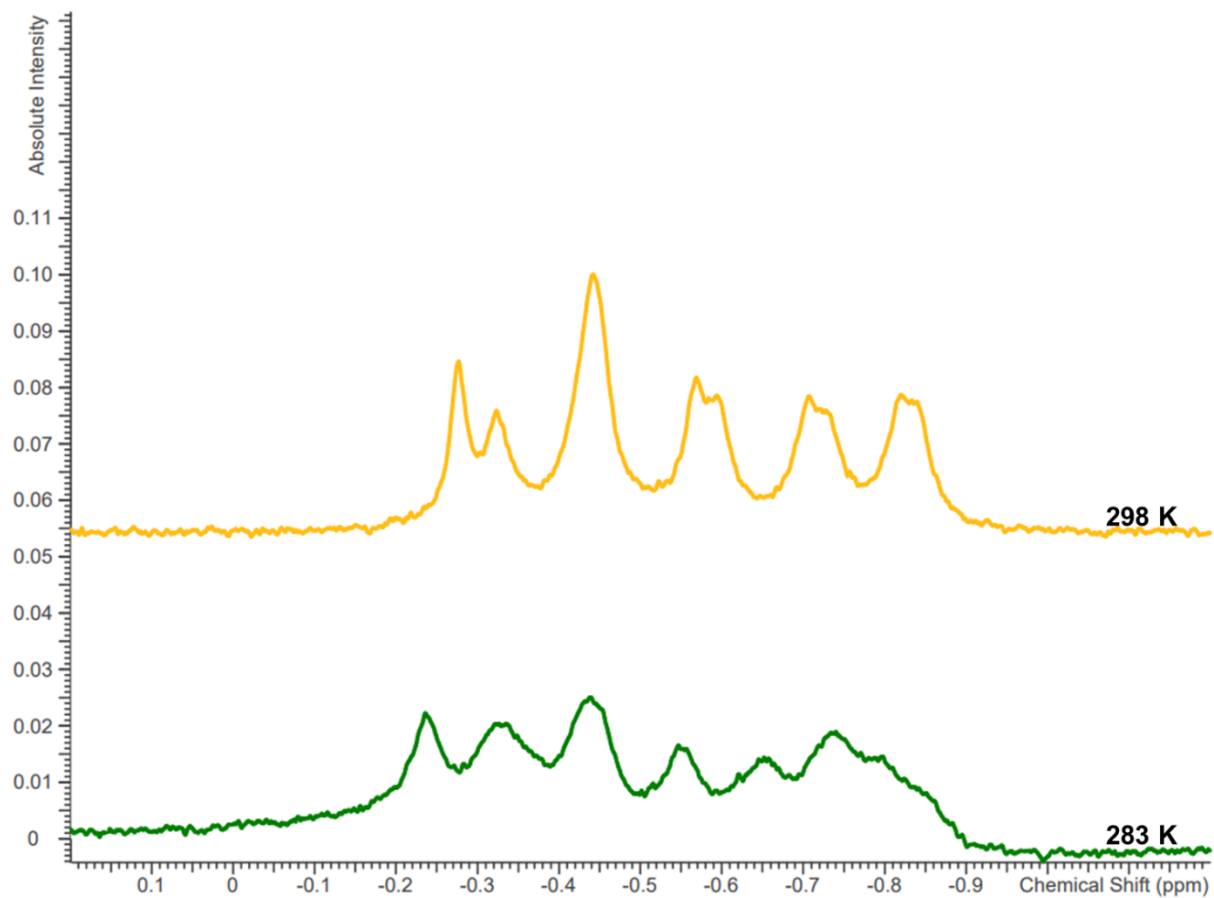
Appendix D-15: DDD 20 mM MgCl₂ Sample Condition H4′ and HSQC Assignments

at 298 K. The H4′ assignments are on the left and the HSQC assignments are on the right.

Appendix D-16: Stacked 1D ^{31}P Spectra at 283 and 298 K of DDD in 25 mM Phosphate Buffer, 100 mM NaCl and 20 mM MgCl₂.

Appendix D-17: Table of Proton Chemical Shift Values for 25 mM Phosphate Buffer, 100 mM NaCl, and 20 mM MgCl₂ DDD Sample Condition at 298 K. All values are in parts per million (ppm).

Base	H1'	H2'	H2''	H3'	H4'	H6/H8	H5/methyl
C1	5.76	1.97	2.40	4.70	4.06	7.63	5.91
G2	5.87	2.63	2.70	4.96	4.33	7.94	
C3	5.58	1.84	2.26	4.80	4.12	7.26	5.36
G4	5.43	2.66	2.74	4.98	4.31	7.85	
A5	5.98	2.70	2.93	5.05	4.45	8.10	
A6	6.15	2.55	2.93	5.00	4.46	8.10	
T7	5.89	1.97	2.56	4.81	4.19	7.10	1.26
T8	6.08	2.15	2.55	4.89	4.19	7.36	1.52
C9	5.65	2.06	2.40	4.86	4.14	7.45	5.62
G10	5.84	2.64	2.68	4.98	4.36	7.90	
C11	5.75	1.89	2.32	4.81	4.14	7.32	5.43
G12	6.15	2.34	2.62	4.67	4.16	7.93	

Appendix D-18: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer, 100 mM NaCl, and 20 mM MgCl₂ DDD Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.44	-0.44
G2pC3	-0.55	-0.57
C3pG4	-0.33	-0.42
G4pA5	-0.65	-0.59
A5pA6	-0.74	-0.70
A6pT7	-0.80	-0.82
T7pT8	-0.80	-0.82
T8pC9	-0.74	-0.73
C9pG10	-0.33	-0.32
G10pC11	-0.44	-0.45
C11pG12	-0.24	-0.27

Appendix D-19: Stacked 1D ^{31}P Spectra at 283 and 298 K of DDD in 25 mM Phosphate Buffer and 300 mM NaCl.
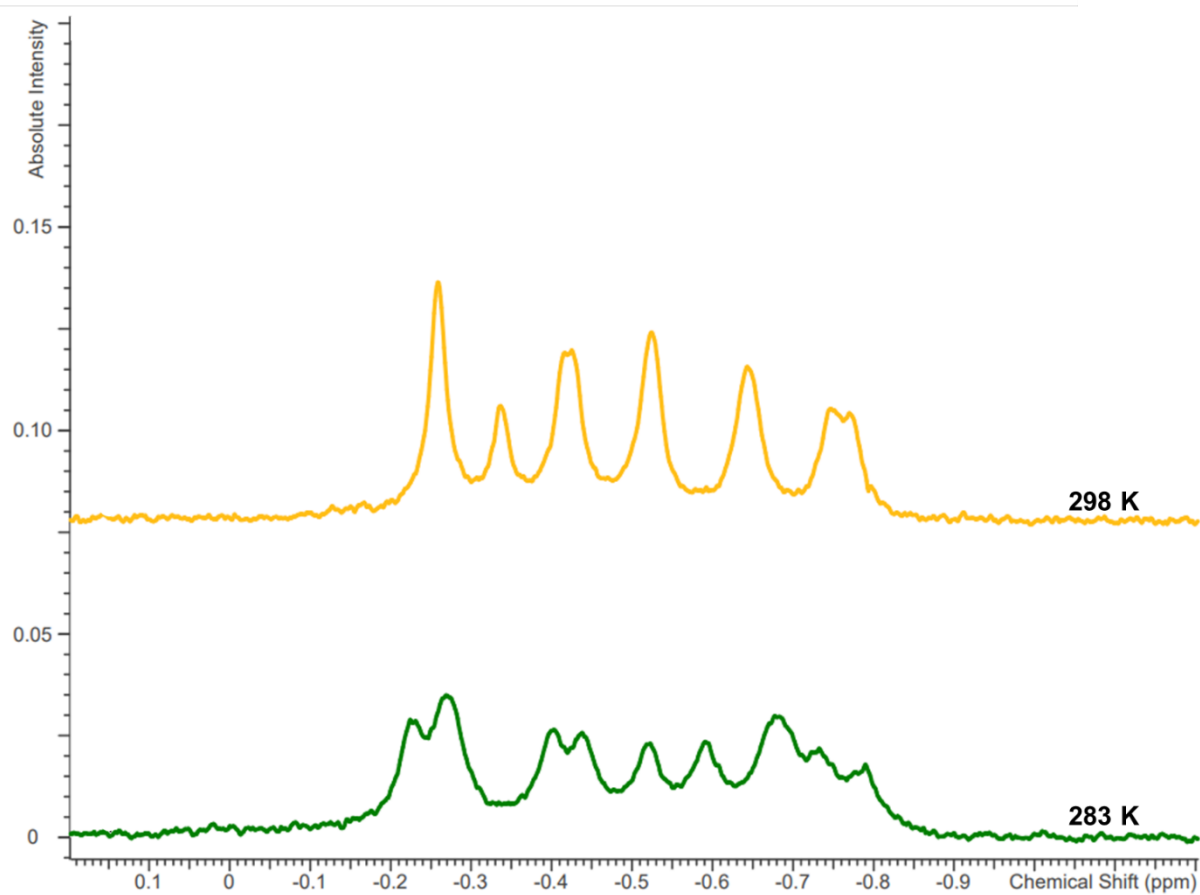
Appendix D-20: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer and 300 mM NaCl DDD Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

| Phosphate Step	283K	298K
C1pG2	-0.44	-0.42
G2pC3	-0.52	-0.52
C3pG4	-0.27	-0.34
G4pA5	-0.59	-0.52
A5pA6	-0.68	-0.64
A6pT7	-0.79	-0.77
T7pT8	-0.73	-0.75
T8pC9	-0.68	-0.64
C9pG10	-0.22	-0.26
G10pC11	-0.40	-0.42
C11pG12	-0.27	-0.26

Appendix D-21: DDD pH 5 Sample Condition H1′ and H2′ Walks at 298 K. The H1′

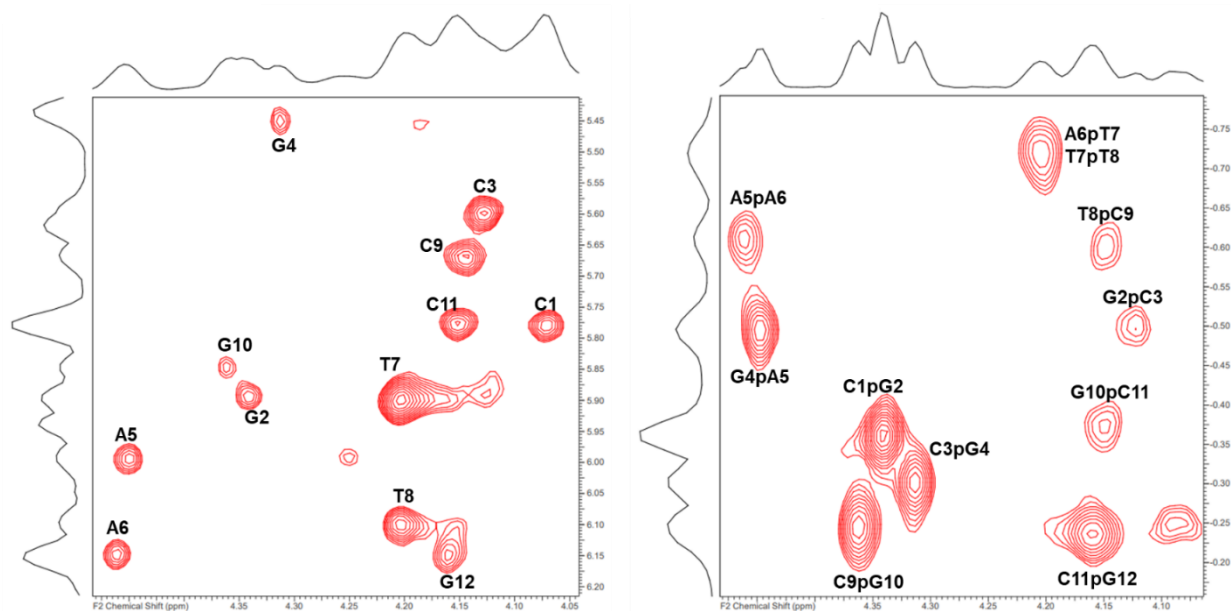walk is on the left and the H2′ walk is on the right.
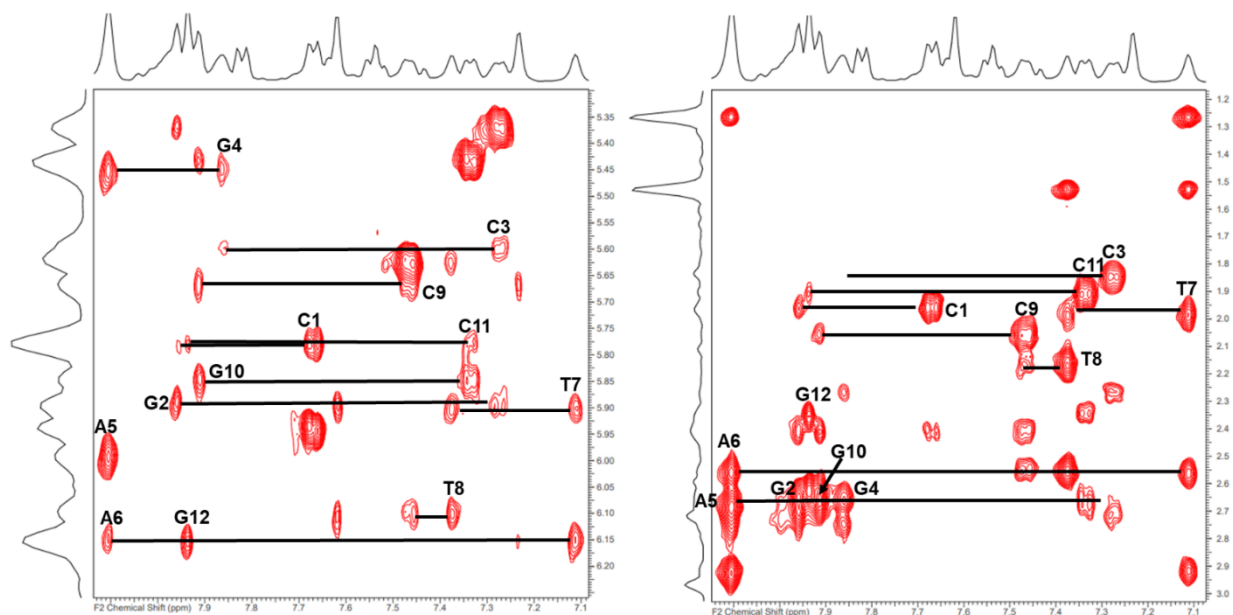
Appendix D-22: DDD pH 5 Sample Condition H2″ Walk and H3′ Assignments at

298 K. The H2″ walk is on the left and the H3′ assignments are on the right.
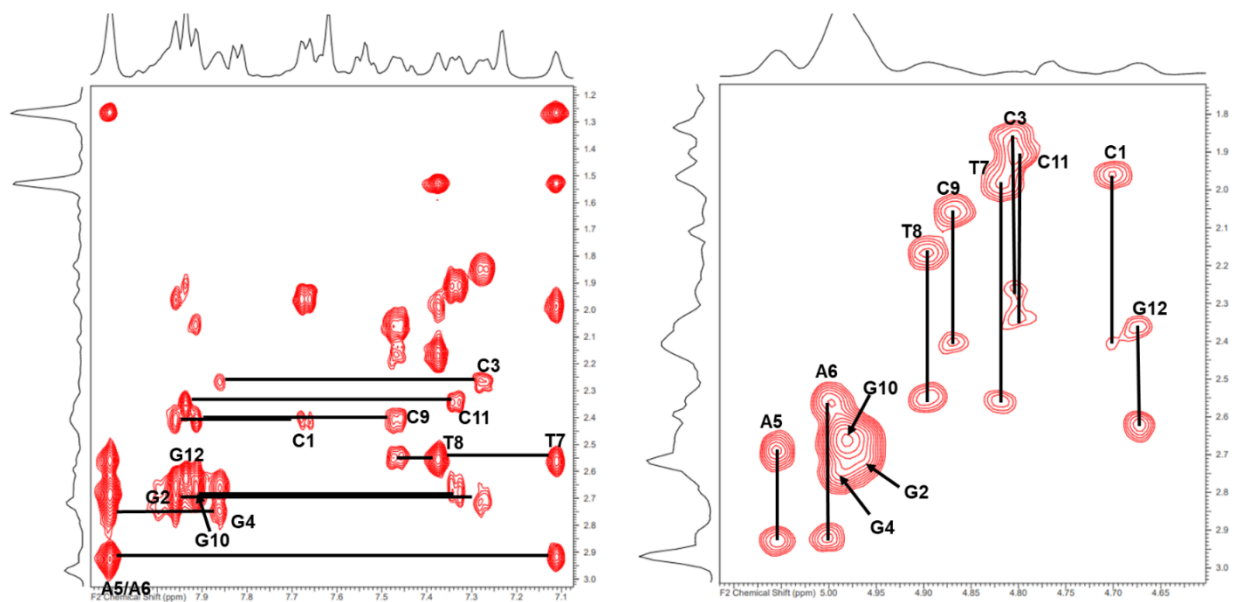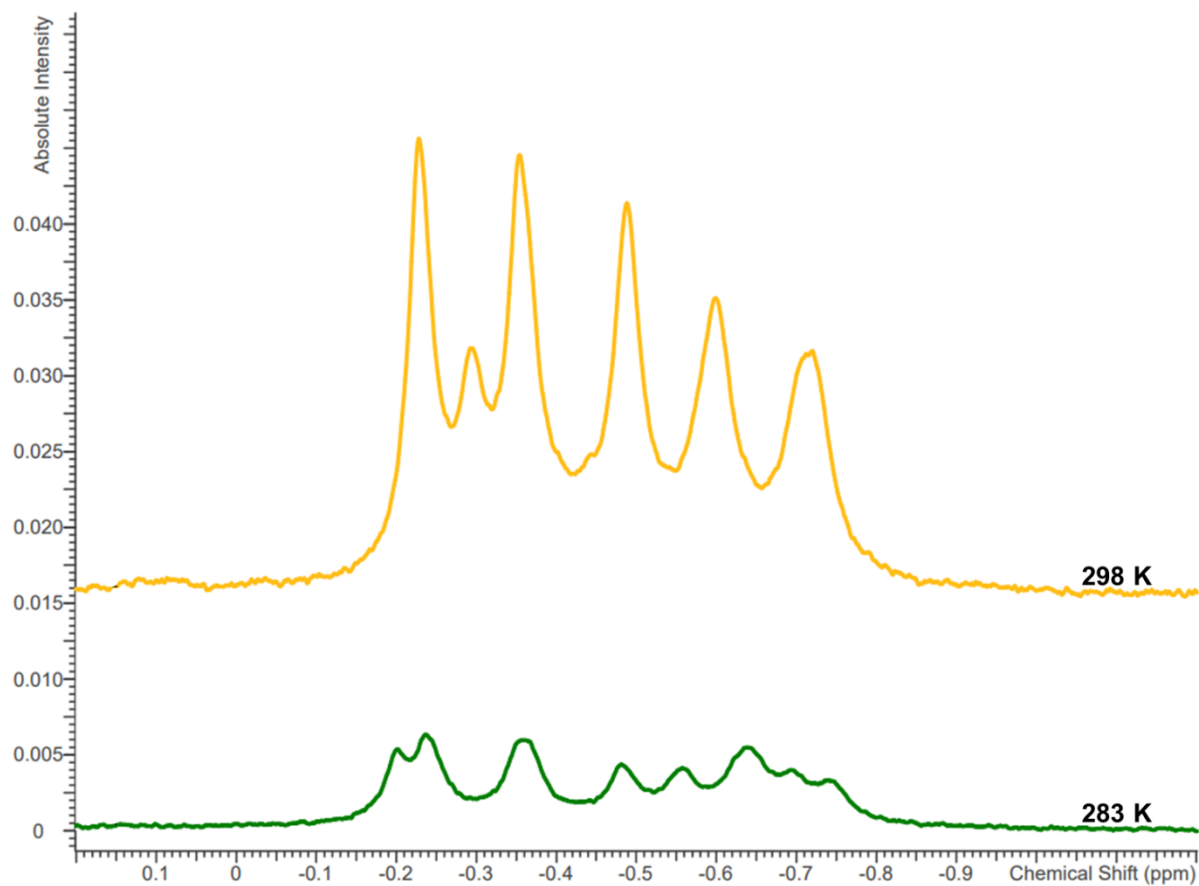
Appendix D-23: DDD pH 5 Sample Condition H4′ and HSQC Assignments at 298 K.

The H4′ assignments are on the left and the HSQC assignments are on the right.

Appendix D-24: Stacked 1D ^{31}P Spectra at 283 and 298 K of DDD in 25 mM Phosphate Buffer with a pH of 5.

Appendix D-25: Table of Proton Chemical Shift Values for 25 mM Phosphate Buffer with a pH of 5 DDD Sample Condition at 298 K. All values are in parts per million (ppm).

Base	H1'	H2'	H2"	H3'	H4'	H6/H8	H5/methyl
C1	5.78	1.96	2.41	4.70	4.07	7.67	5.94
G2	5.89	2.65	2.71	4.96	4.34	7.96	
C3	5.60	1.85	2.26	4.80	4.13	7.27	5.37
G4	5.45	2.66	2.75	4.99	4.31	7.86	
A5	5.99	2.69	2.93	5.06	4.45	8.10	
A6	6.15	2.56	2.93	5.00	4.46	8.10	
T7	5.90	1.98	2.56	4.82	4.20	7.11	1.26
T8	6.10	2.17	2.56	4.90	4.20	7.37	1.53
C9	5.67	2.06	2.41	4.87	4.14	7.47	5.63
G10	5.85	2.64	2.68	4.98	4.36	7.91	
C11	5.78	1.91	2.34	4.80	4.15	7.34	5.43
G12	6.15	2.35	2.63	4.67	4.16	7.94	

Appendix D-26: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer with a pH of 5 DDD Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.36	-0.36
G2pC3	-0.48	-0.50
C3pG4	-0.24	-0.30
G4pA5	-0.56	-0.49
A5pA6	-0.64	-0.61
A6pT7	-0.74	-0.72
T7pT8	-0.69	-0.72
T8pC9	-0.64	-0.60
C9pG10	-0.20	-0.25
G10pC11	-0.36	-0.37
C11pG12	-0.24	-0.24

**Appendix D-27: Stacked 1D ^{31}P Spectra at 283 and 298 K of DDD in 25 mM Phosphate Buffer with a pH of 9.**
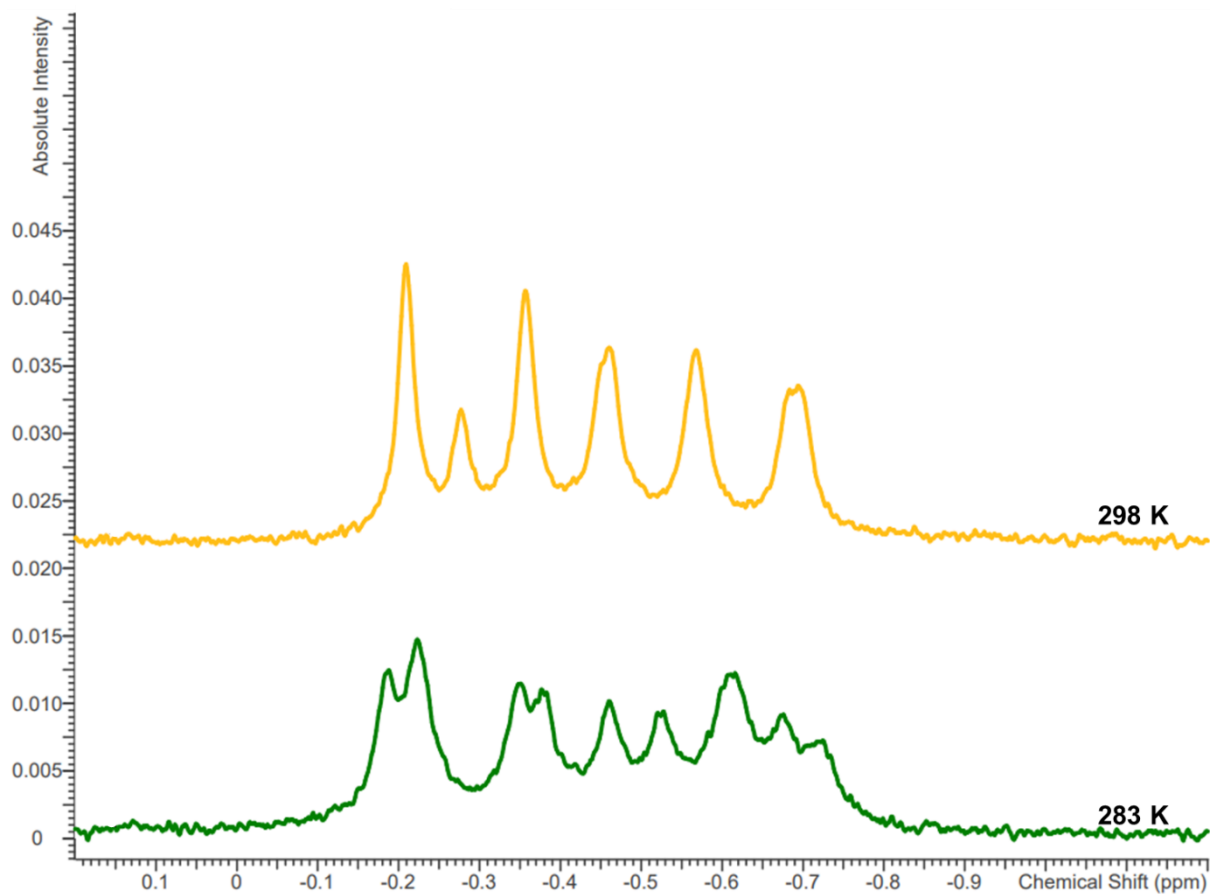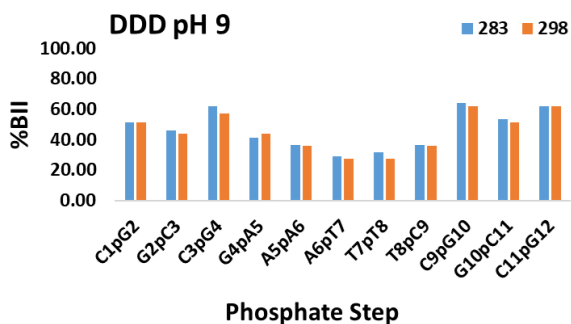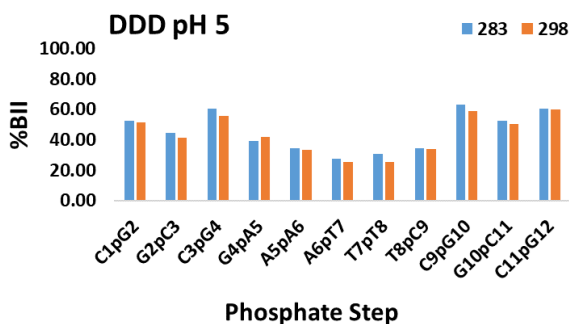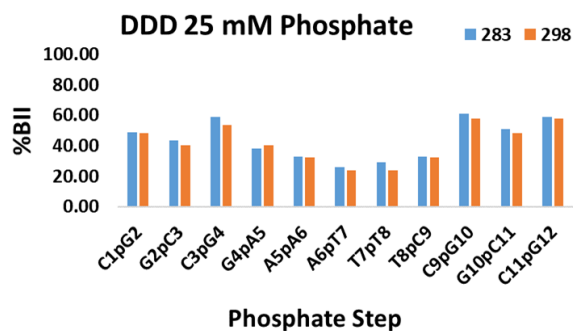
Appendix D-28: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer with a pH of 9 DDD Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.38	-0.36
G2pC3	-0.46	-0.46
C3pG4	-0.22	-0.28
G4pA5	-0.53	-0.46
A5pA6	-0.61	-0.57
A6pT7	-0.72	-0.69
T7pT8	-0.68	-0.69
T8pC9	-0.61	-0.57
C9pG10	-0.19	-0.21
G10pC11	-0.35	-0.36
C11pG12	-0.22	-0.21

Appendix D-29: %BII Plots for all DDD Salt Conditions at 283 and 298 K. These values were calculated via Equation 1. The blue bars correspond to 283 K and the orange correspond to 298 K.
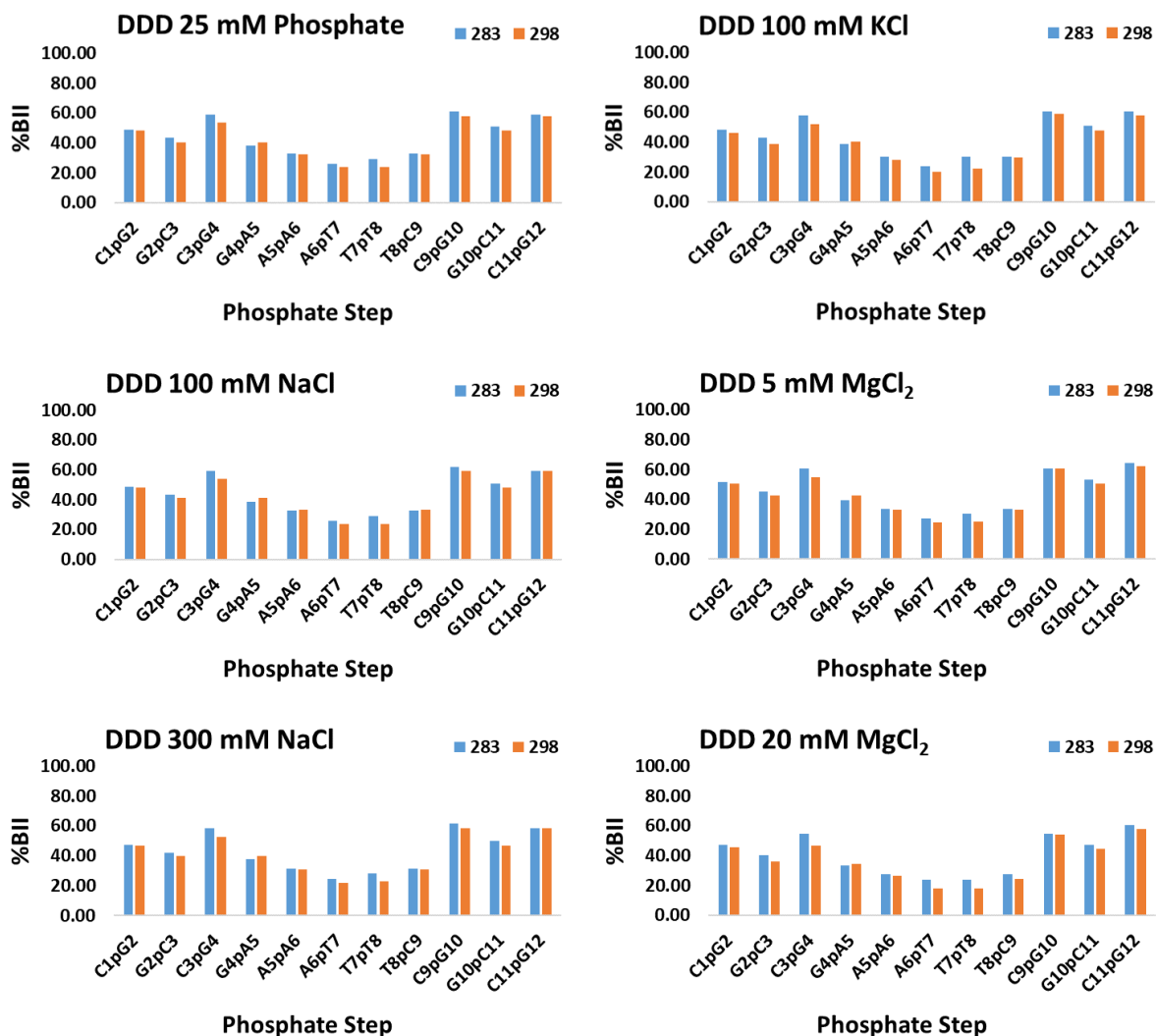
Appendix D-30: %BII Plots for all DDD pH Conditions at 283 and 298 K. These values were calculated via Equation 1. The blue bars correspond to 283 K and the orange correspond to 298 K.

Appendix D-31: ΔG Plots for all DDD Salt Conditions at 283 and 298 K. These

values were calculated via Equation 2. The blue bars correspond to 283 K and the orange

correspond to 298 K.

Appendix D-32: ΔG Plots for all DDD pH Conditions at 283 and 298 K. These values were calculated via Equation 2. The blue bars correspond to 283 K and the orange correspond to 298 K.
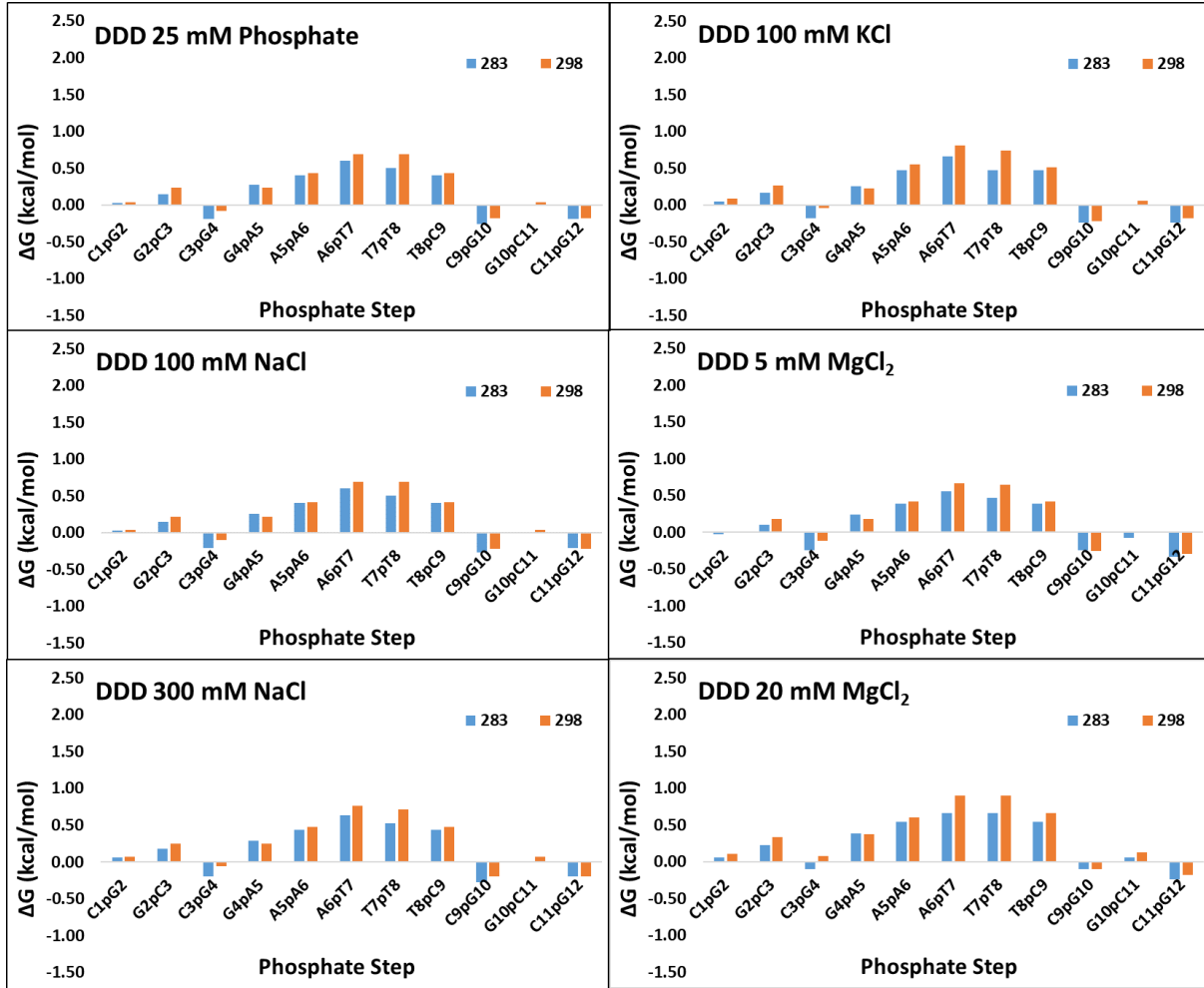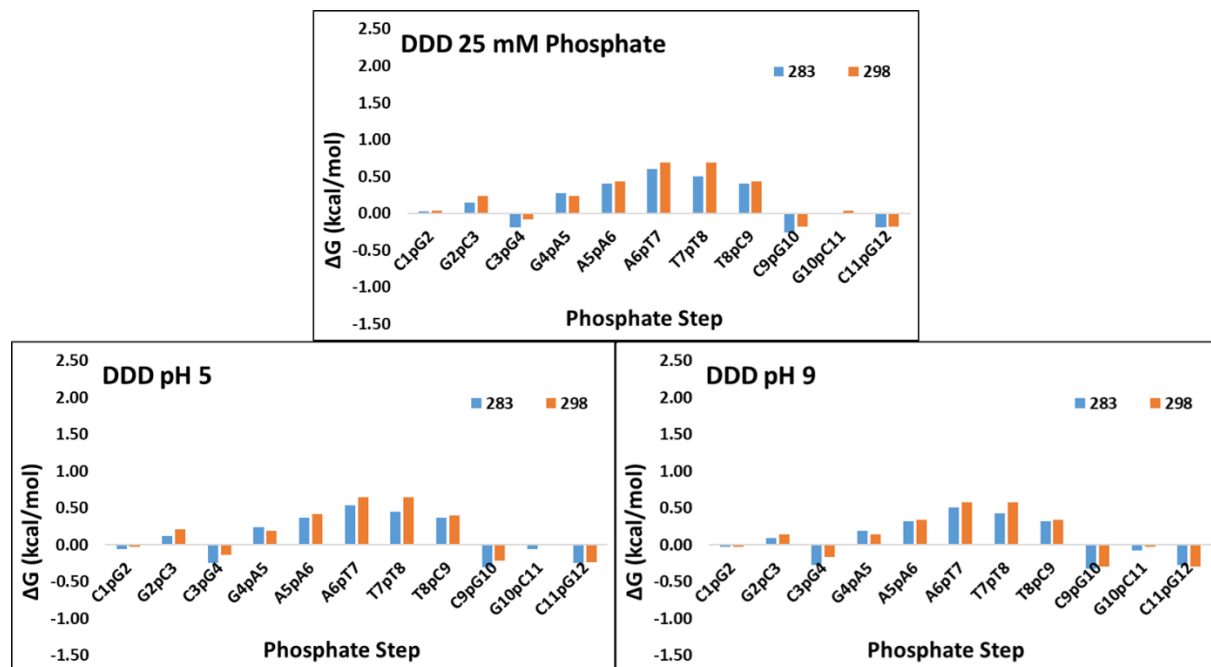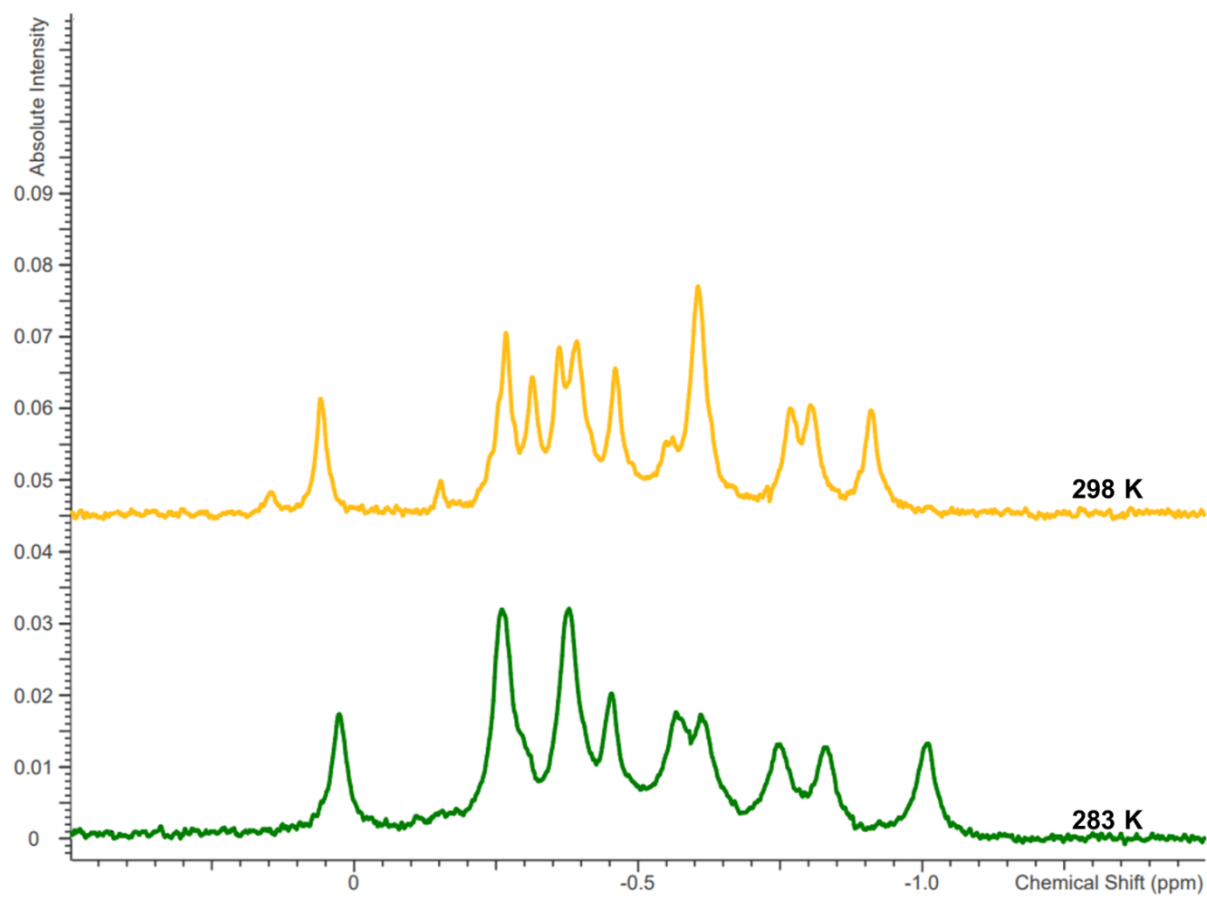
Appendix E: T9 Supplemental Data

        Appendix E-1: Stacked 1D ^{31}P Spectra at 283 and 298 K of T9 in 25 mM Phosphate Buffer.

Appendix E-2: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer T9 Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.38	-0.36
G2pC3	-0.38	-0.39
C3pG4	-0.26	-0.31
G4pA5	-1.01	-0.91
A5pA6	-0.61	-0.61
A6pT7	-0.83	-0.80
T7pT8	-0.57	-0.61
T8pT9	-0.75	-0.77
T9pG10	0.03	0.06
G10pC11	-0.45	-0.46
C11pG12	-0.26	-0.27

Appendix E-3: Stacked 1D ^{31}P Spectra at 283 and 298 K of T9 in 25 mM Phosphate Buffer and 100 mM NaCl.
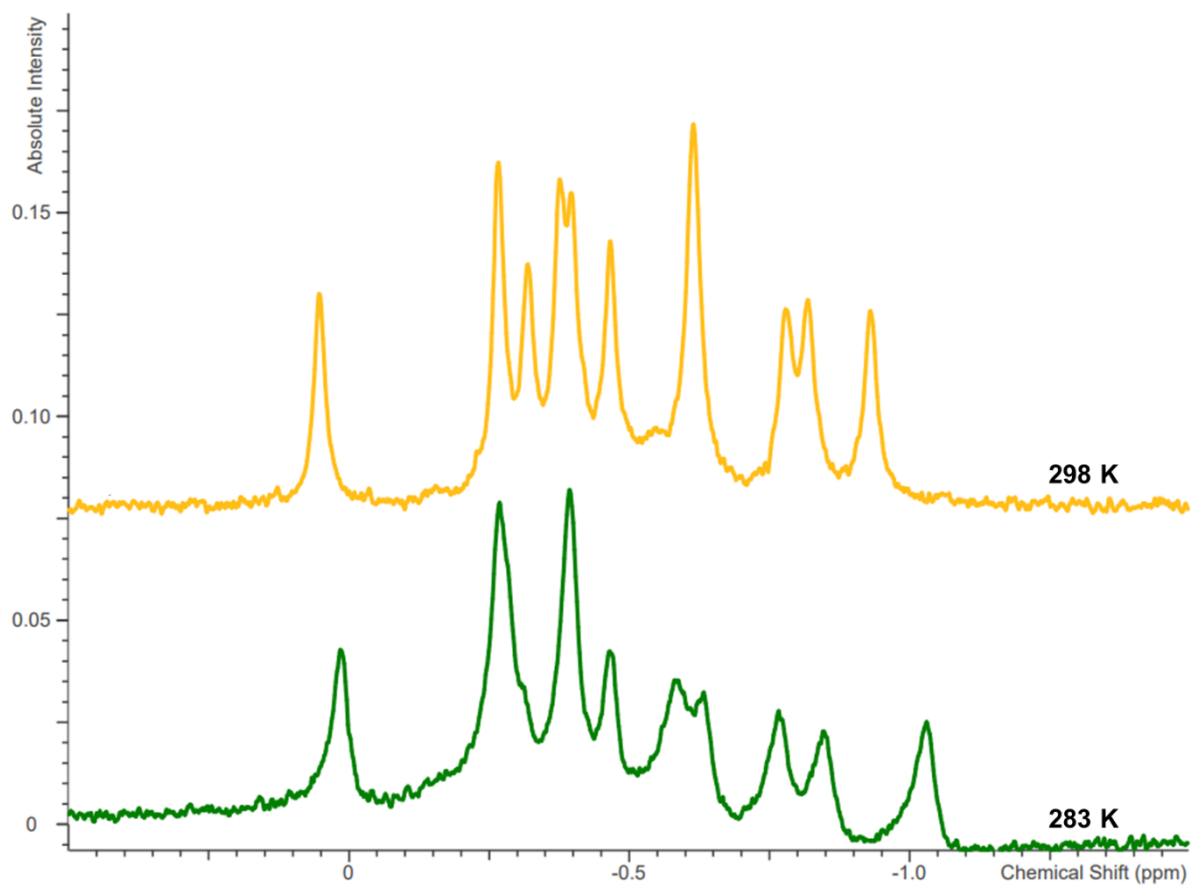
Appendix E-4: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer and 100 mM NaCl T9 Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.39	-0.38
G2pC3	-0.39	-0.40
C3pG4	-0.27	-0.32
G4pA5	-1.03	-0.93
A5pA6	-0.63	-0.61
A6pT7	-0.85	-0.82
T7pT8	-0.59	-0.61
T8pT9	-0.77	-0.78
T9pG10	0.02	0.05
G10pC11	-0.46	-0.47
C11pG12	-0.27	-0.27

Appendix E-5: T9 100 mM KCl Sample Condition H1′ Walk and H3′ Assignments at 298 K. The H1′ walk is on the left and the H3′ assignments are on the right.
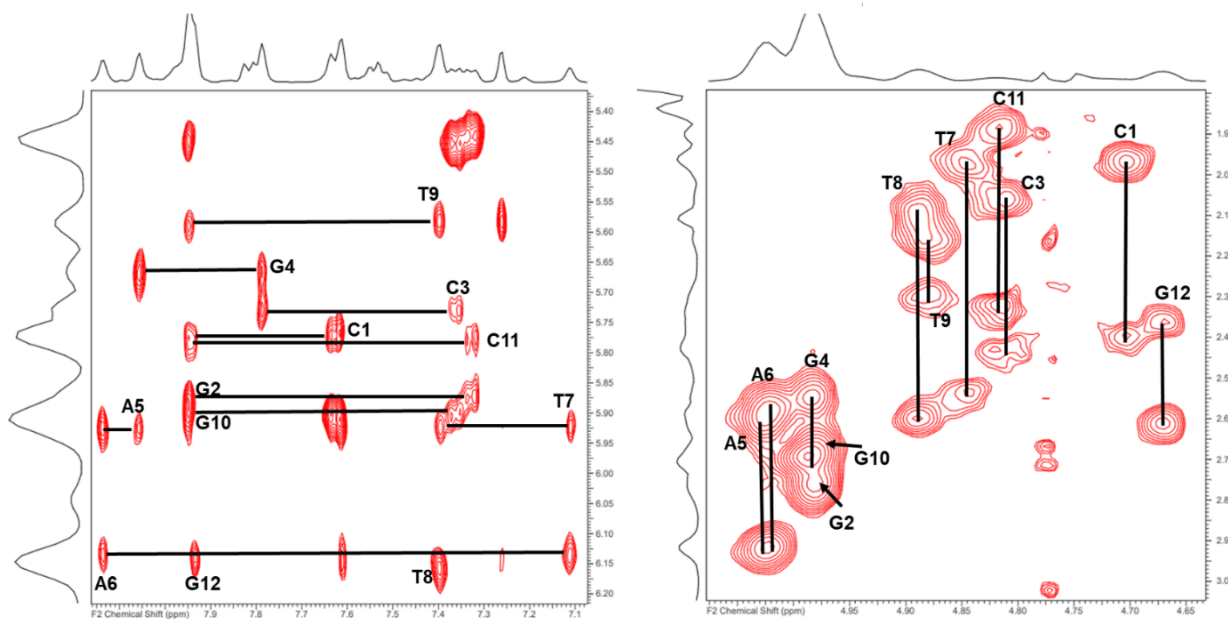
Appendix E-6: T9 100 mM KCl H2′ and H2″ Walks at 298 K. The H2′ walk is on the left and H2″ walk is on the right.
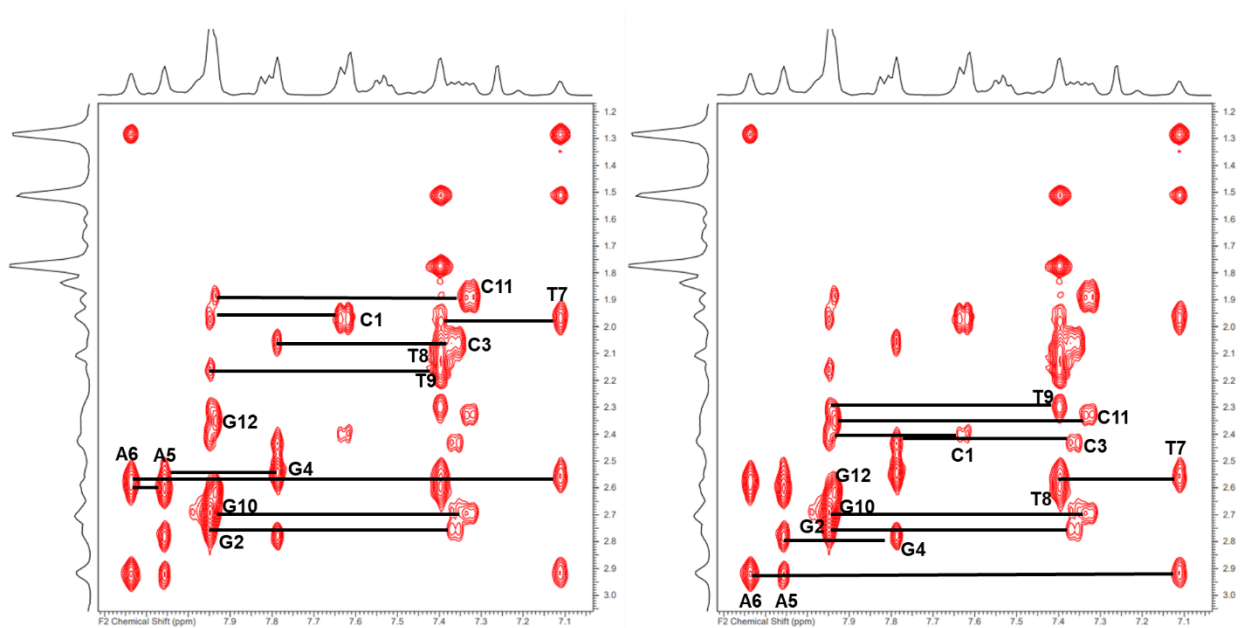
Appendix E-7: T9 100 mM KCl Sample Condition H4′ and HSQC Assignments at

298 K. The H4′ assignments are on the left and the HSQC assignments are on the right.
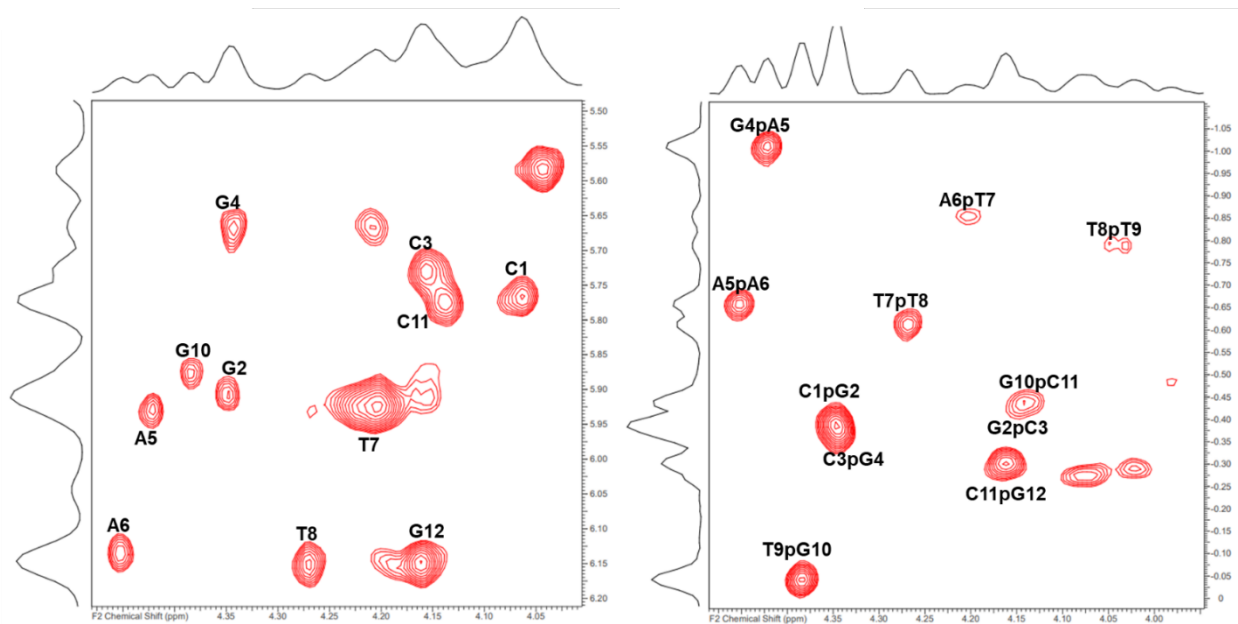
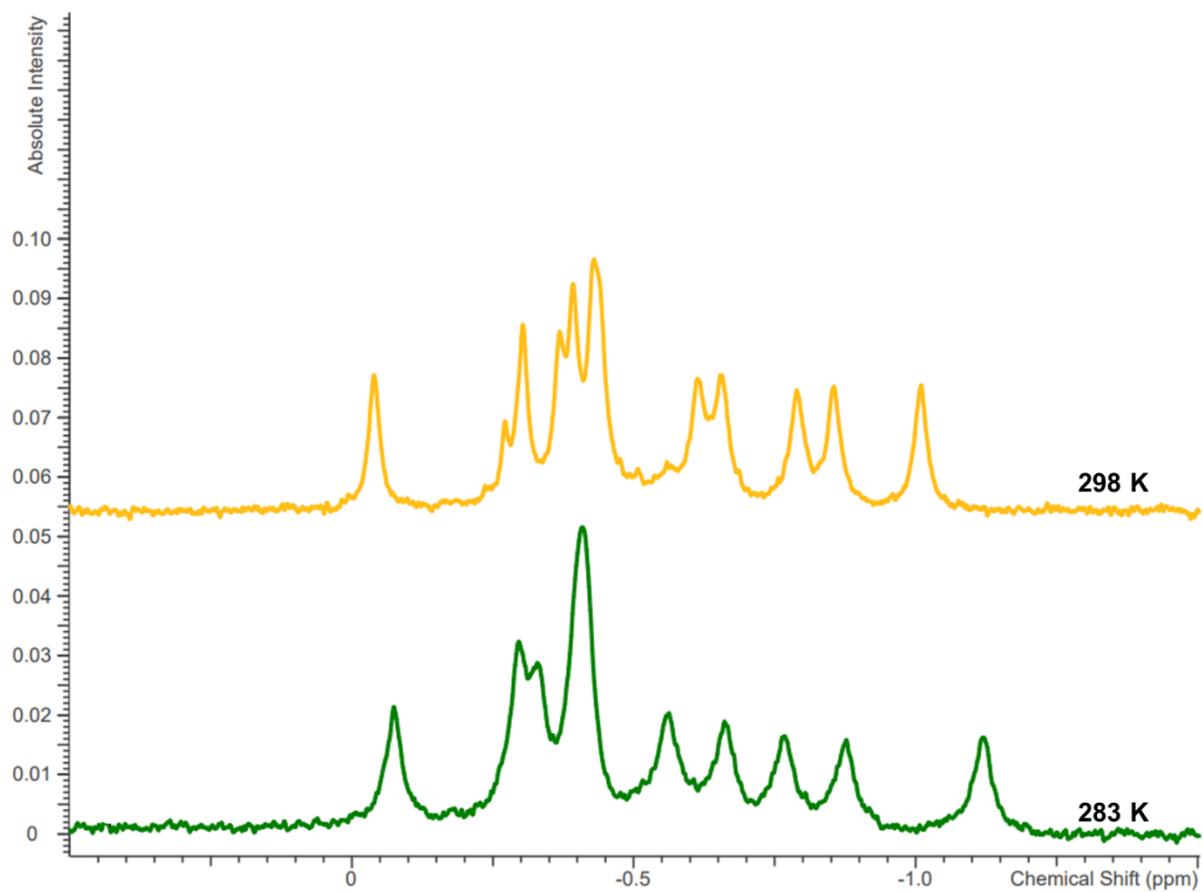**Appendix E-8: Stacked 1D ^{31}P Spectra at 283 and 298 K of T9 in 25 mM Phosphate Buffer and 100 mM KCl.**

Appendix E-9: Table of Proton Chemical Shift Values for 25 mM Phosphate Buffer and 100 mM KCl T9 Sample Condition at 298 K. All values are in parts per million (ppm).

Base	H1'	H2'	H2"	H3'	H4'	H6/H8	H5/methyl
C1	5.77	1.98	2.40	4.70	4.06	7.63	5.91
G2	5.90	2.75	2.75	4.98	4.35	7.95	
C3	5.73	2.06	2.43	4.81	4.16	7.36	5.45
G4	5.67	2.55	2.78	4.98	4.34	7.79	
A5	5.93	2.61	2.93	5.03	4.42	8.06	
A6	6.14	2.58	2.93	5.02	4.45	8.14	
T7	5.92	1.97	2.56	4.85	4.20	7.11	1.28
T8	6.15	2.09	2.60	4.89	4.27	7.40	1.51
T9	5.58	2.16	2.30	4.88	4.04	7.40	1.78
G10	5.88	2.70	2.70	4.98	4.38	7.95	
C11	5.78	1.90	2.33	4.82	4.14	7.33	5.44
G12	6.14	2.35	2.62	4.67	4.16	7.93	

Appendix E-10: Table of Proton Chemical Shift Values for 25 mM Phosphate Buffer and 100 mM KCl T9 Sample Condition at 283 K. All values are in parts per million (ppm).

Base	H1'	H2'	H2''	H3'	H4'	H6/H8	H5/methyl
C1	5.75	2.01	2.43	4.71	4.06	7.65	5.90
G2	5.91	n/a	2.76	4.99	4.36	7.97	
C3	5.71	2.07	2.45	4.82	4.15	7.37	5.46
G4	5.68	2.57	2.80	5.00	4.36	7.83	
A5	5.94	2.62	2.95	5.04	4.44	8.08	
A6	6.14	2.58	2.95	5.04	4.46	8.17	
T7	5.96	2.00	2.57	4.88	4.22	7.16	1.28
T8	6.17	2.10	2.60	4.89	4.29	7.44	1.51
T9	5.59	2.16	2.29	4.89	4.05	7.42	1.77
G10	5.87	2.69	2.69	4.99	4.40	7.97	
C11	5.76	1.91	2.35	4.83	4.15	7.35	5.45
G12	6.15	2.35	2.64	4.68	4.17	7.96	

Appendix E-11: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer and 100 mM KCl T9 Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.39	-0.38
G2pC3	-0.41	-0.42
C3pG4	-0.34	-0.38
G4pA5	-1.12	-1.01
A5pA6	-0.66	-0.66
A6pT7	-0.88	-0.86
T7pT8	-0.56	-0.61
T8pT9	-0.76	-0.79
T9pG10	-0.08	-0.04
G10pC11	-0.41	-0.44
C11pG12	-0.30	-0.30

Appendix E-12: Stacked 1D ^{31}P Spectra at 283 and 298 K of T9 in 25 mM Phosphate Buffer, 100 mM NaCl, and 5 mM MgCl₂.
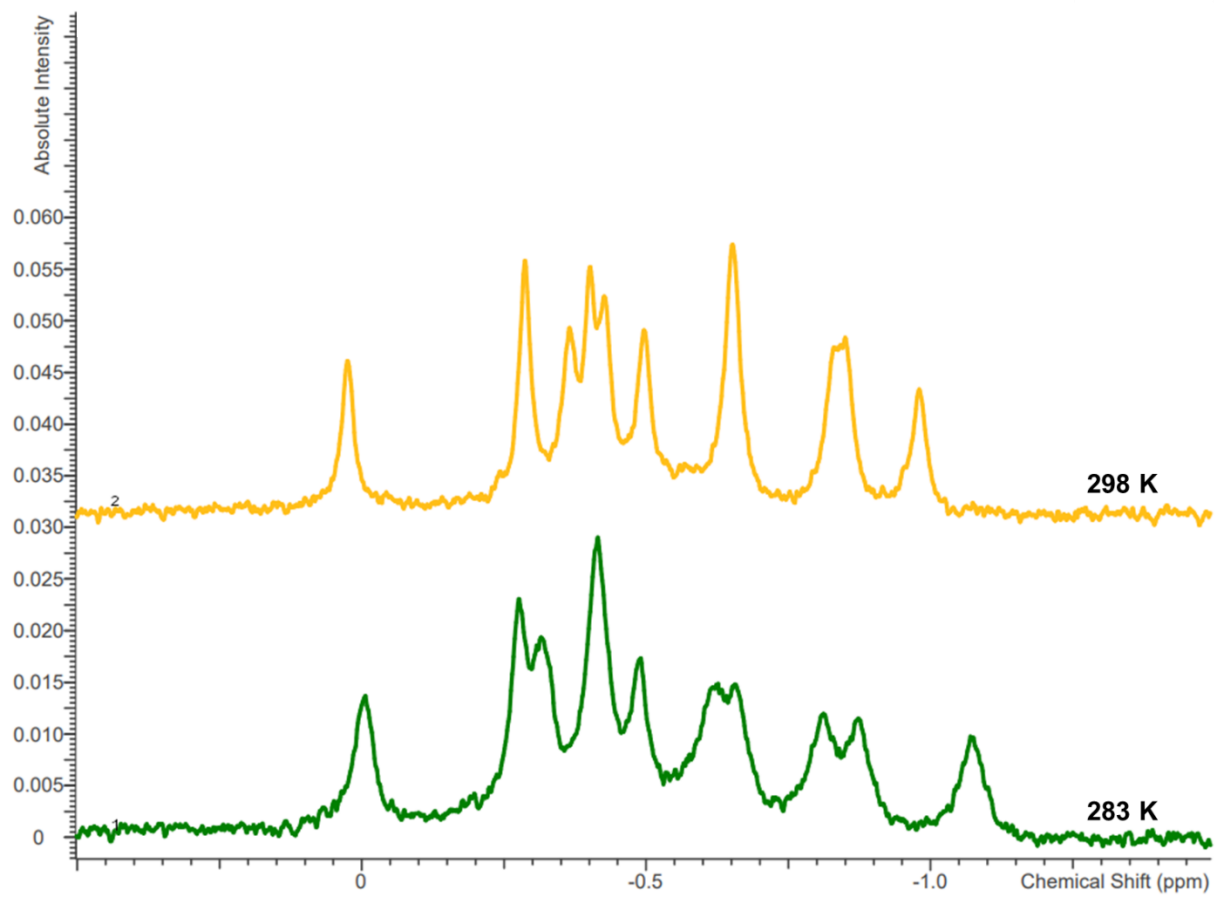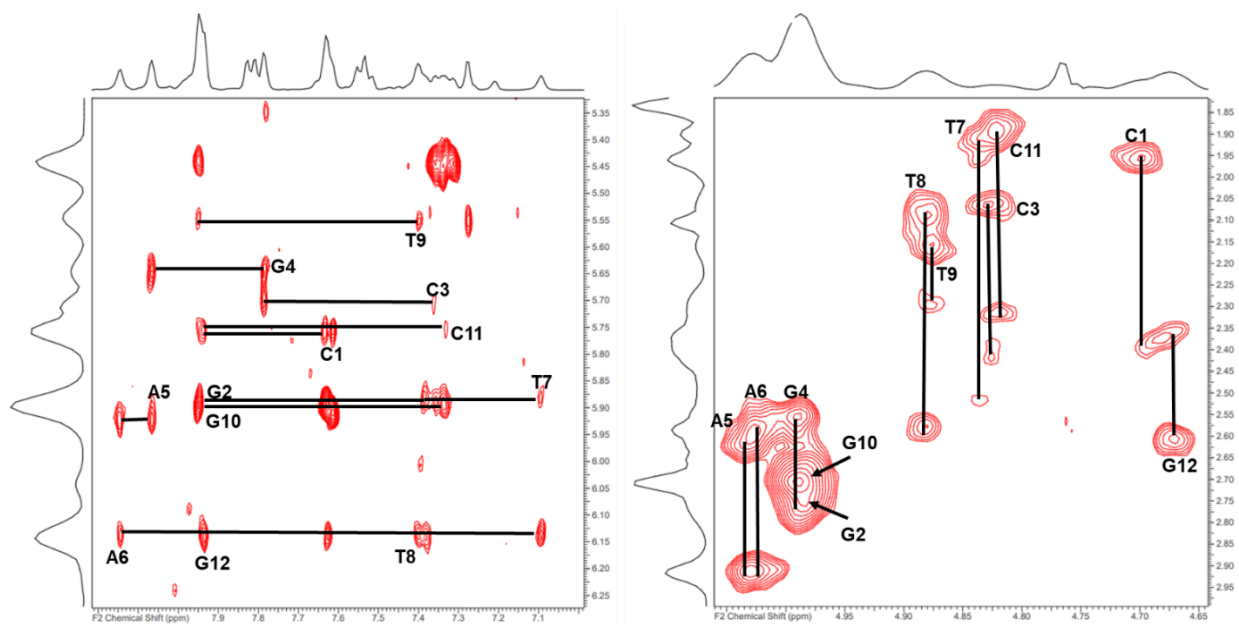
Appendix E-13: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer, 100 mM NaCl, and 5 mM MgCl₂ T9 Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

| Phosphate Step	283K	298K
C1pG2	-0.41	-0.40
G2pC3	-0.41	-0.43
C3pG4	-0.32	-0.37
G4pA5	-1.07	-0.98
A5pA6	-0.66	-0.65
A6pT7	-0.87	-0.84
T7pT8	-0.62	-0.65
T8pT9	-0.81	-0.84
T9pG10	-0.01	0.02
G10pC11	-0.49	-0.50
C11pG12	-0.28	-0.29

Appendix E-14: T9 20 mM MgCl₂ Sample Condition H1′ Walk and H3′ Assignments at 298 K. The H1′ walk is on the left and the H3′ assignments are on the right.
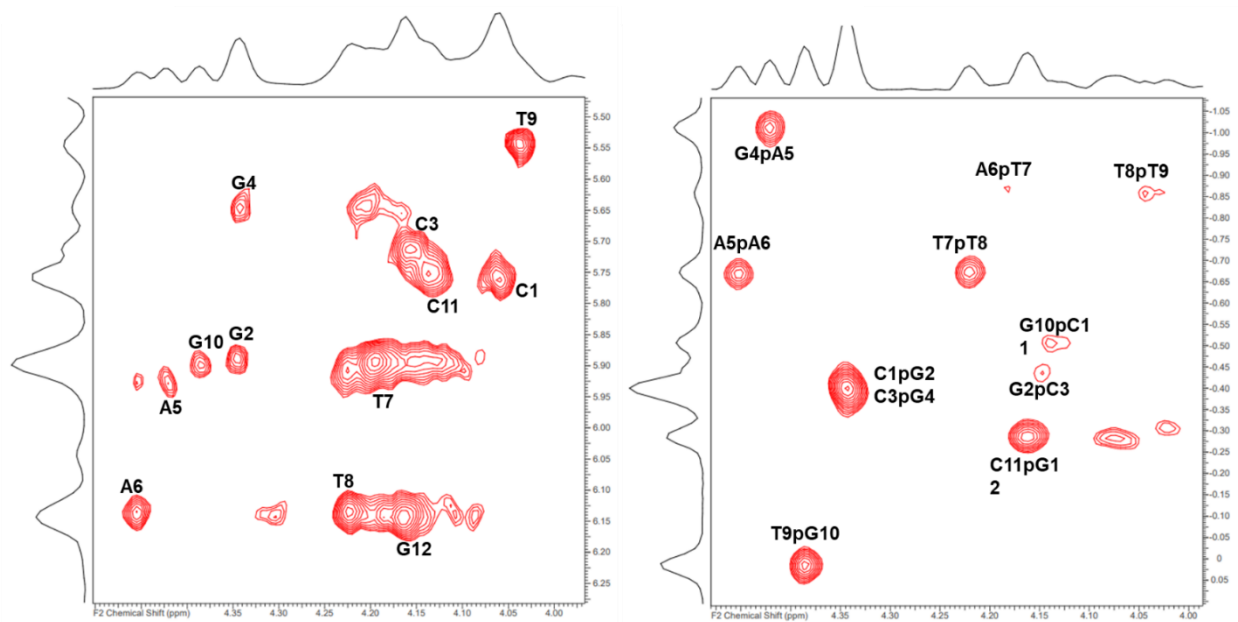
Appendix E-15: T9 20 mM MgCl₂ Sample Condition H2′ and H2″ Walks at 298 K.

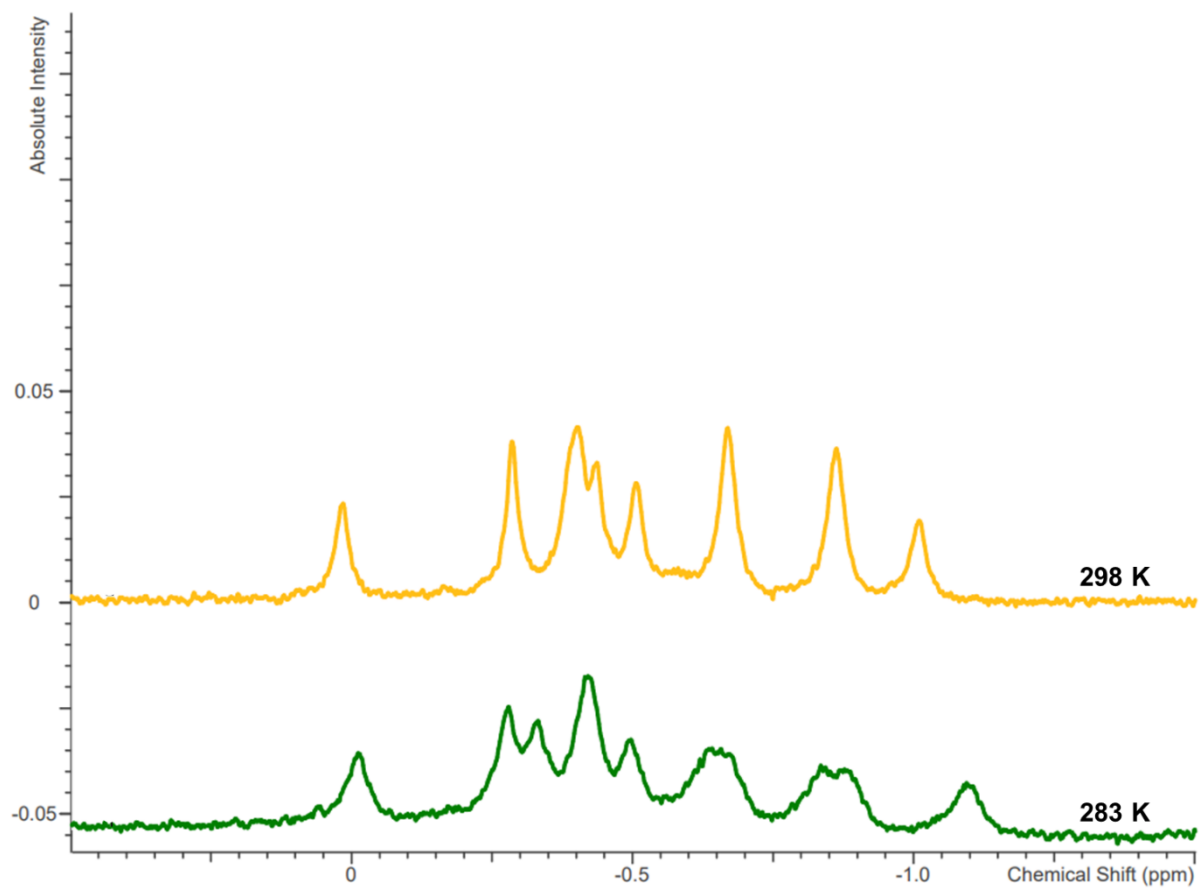The H2′ walk is on the left and the H2″ walk is on the right.
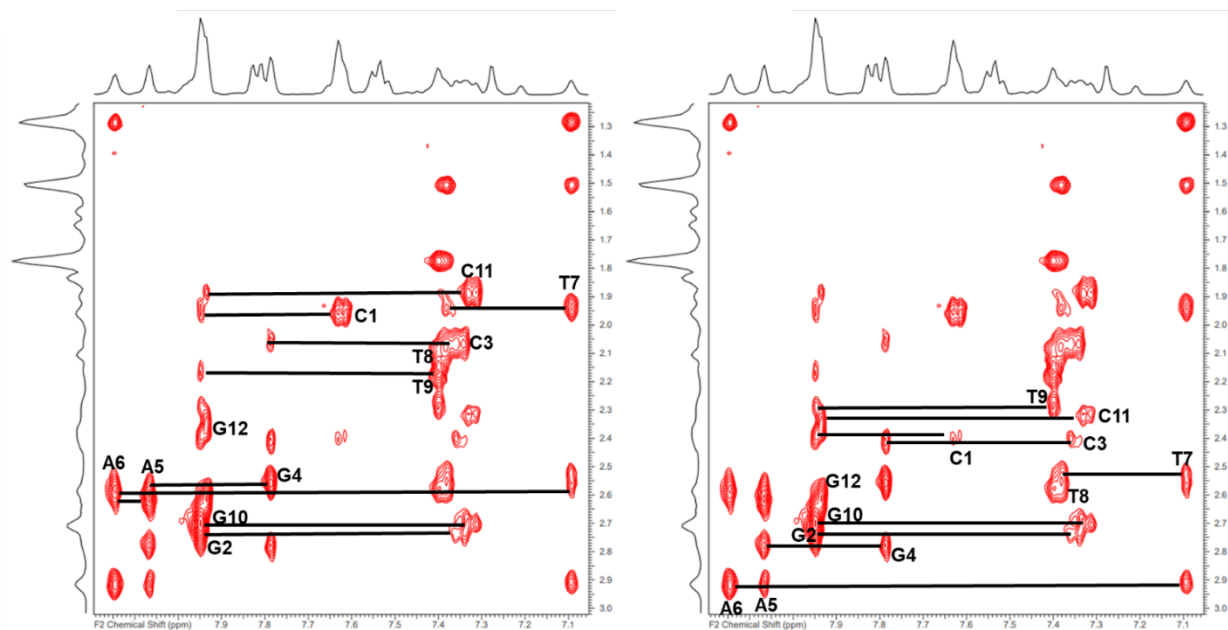
Appendix E-16: T9 20 mM MgCl₂ Sample Condition H4′ and HSQC Assignments at

298 K. The H4′ assignments are on the left and the HSQC assignments are on the right.

Appendix E-17: Stacked 1D ^{31}P Spectra at 283 and 298 K of T9 in 25 mM Phosphate Buffer, 100 mM NaCl, and 20 mM MgCl$_2$.

Appendix E-18: Table of Proton Chemical Shift Values for 25 mM Phosphate Buffer, 100 mM NaCl, and 20 mM MgCl₂ T9 Sample Condition at 298 K. All values are in parts per million (ppm).

Base	H1'	H2'	H2"	H3'	H4'	H6/H8	H5/methyl
C1	5.76	1.96	2.40	4.70	4.06	7.62	5.90
G2	5.89	2.73	2.73	4.97	4.35	7.95	
C3	5.71	2.07	2.41	4.83	4.16	7.35	5.45
G4	5.64	2.56	2.78	4.99	4.34	7.79	
A5	5.92	2.61	2.92	5.03	4.42	8.07	
A6	6.14	2.58	2.92	5.02	4.45	8.14	
T7	5.88	1.94	2.53	4.84	4.19	7.09	1.29
T8	6.14	2.08	2.56	4.88	4.22	7.38	1.51
T9	5.55	2.17	2.28	4.88	4.04	7.40	
G10	5.89	2.70	2.70	4.99	4.39	7.95	
C11	5.75	1.89	2.32	4.82	4.14	7.32	5.44
G12	6.14	2.36	2.61	4.67	4.16	7.94	

Appendix E-19: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer, 100 mM NaCl, and 20 mM MgCl₂ T9 Sample Condition. The values are in parts per million and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.42	-0.40
G2pC3	-0.42	-0.44
C3pG4	-0.33	-0.40
G4pA5	-1.10	-1.01
A5pA6	-0.65	-0.67
A6pT7	-0.88	-0.87
T7pT8	-0.65	-0.67
T8pT9	-0.84	-0.86
T9pG10	-0.01	0.02
G10pC11	-0.50	-0.51
C11pG12	-0.28	-0.28

Appendix E-20: Stacked 1D ^{31}P Spectra at 283 and 298 K of T9 in 25 mM Phosphate Buffer and 200 mM NaCl.
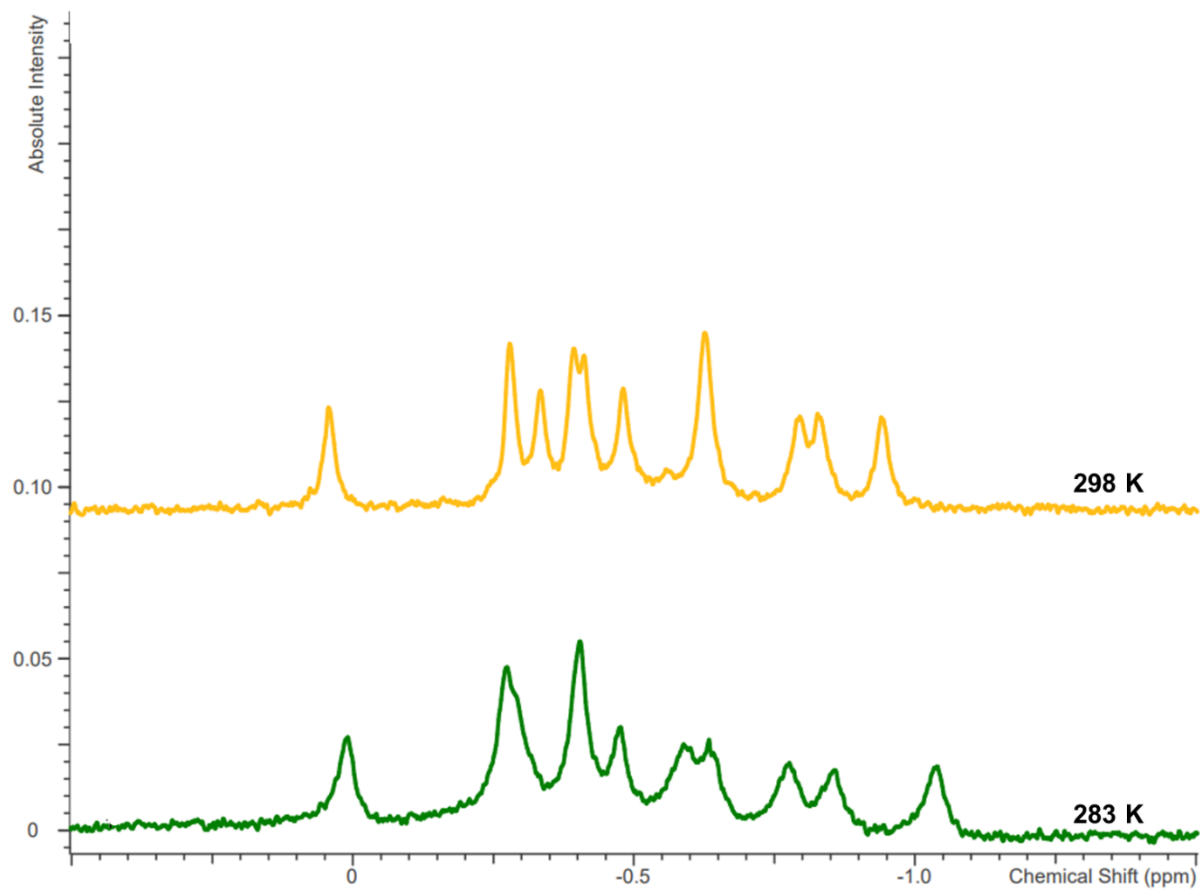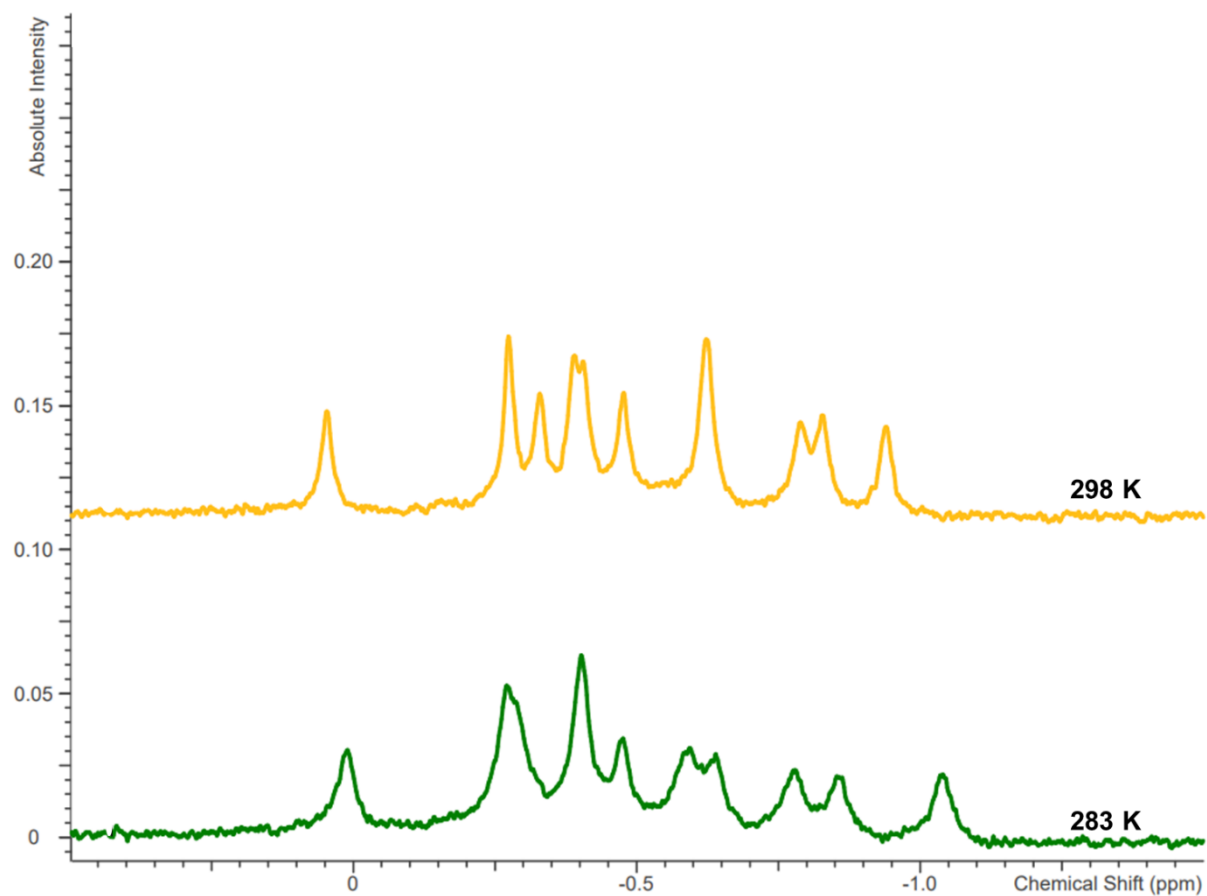
Appendix E-21: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer and 200 mM NaCl T9 Sample Condition. All values are in parts per million and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.40	-0.39
G2pC3	-0.40	-0.41
C3pG4	-0.27	-0.33
G4pA5	-1.04	-0.94
A5pA6	-0.63	-0.63
A6pT7	-0.86	-0.83
T7pT8	-0.59	-0.63
T8pT9	-0.78	-0.80
T9pG10	0.01	0.04
G10pC11	-0.48	-0.48
C11pG12	-0.27	-0.28

Appendix E-22: Stacked 1D ^{31}P Spectra at 283 and 298 K of T9 in 25 mM Phosphate Buffer and 300 mM NaCl.

Appendix E-23: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer and 300 mM NaCl T9 Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

| Phosphate Step	283K	298K
C1pG2	-0.40	-0.39
G2pC3	-0.40	-0.41
C3pG4	-0.27	-0.33
G4pA5	-1.04	-0.94
A5pA6	-0.64	-0.62
A6pT7	-0.86	-0.83
T7pT8	-0.59	-0.62
T8pT9	-0.78	-0.79
T9pG10	0.01	0.05
G10pC11	-0.48	-0.48
C11pG12	-0.27	-0.27

Appendix E-24: T9 pH 5 Sample Condition H1′ Walk and H3′ Assignments at 298

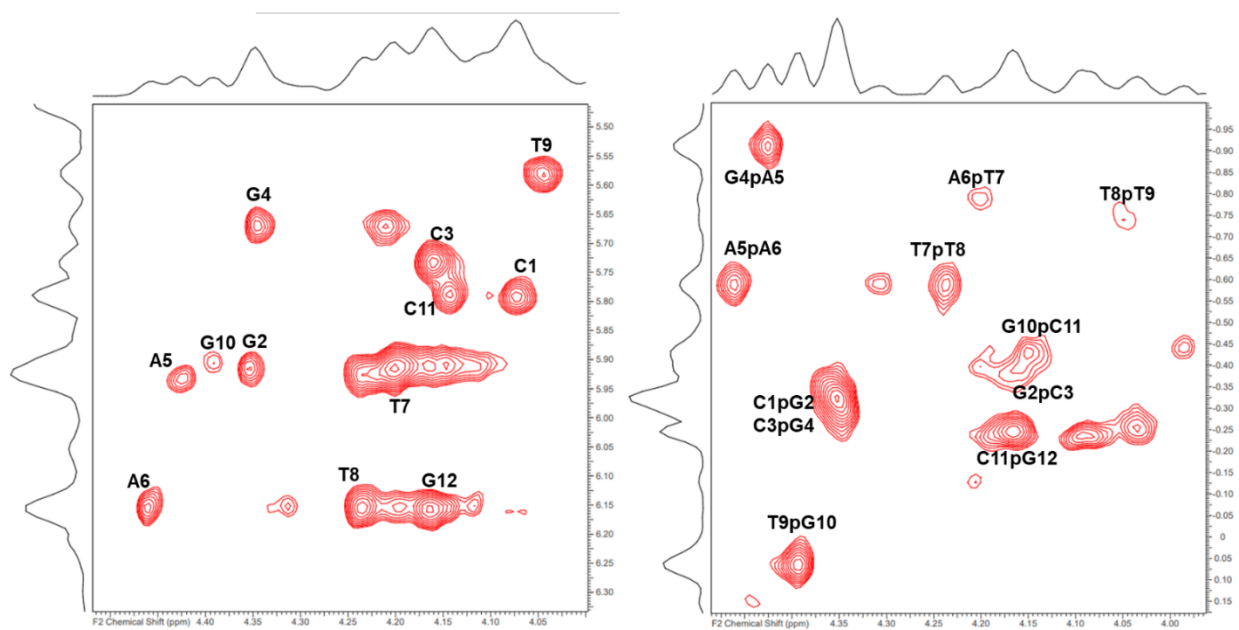**K.** The H1′ walk is on the left and the H3′ assignments are on the right.
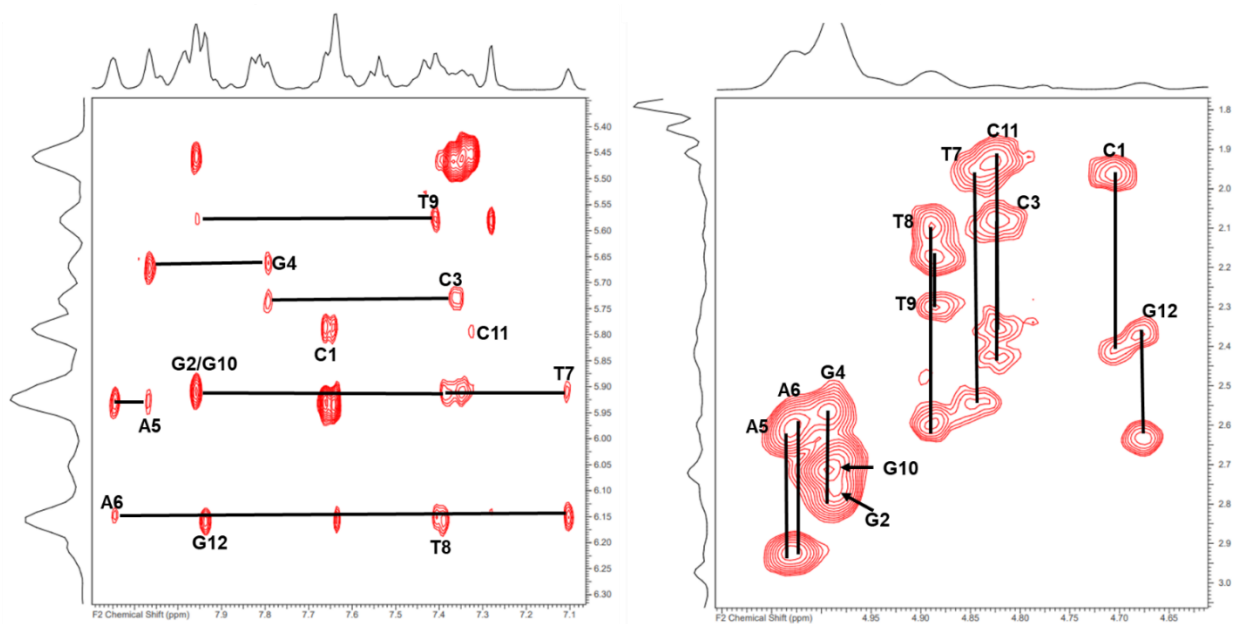
Appendix E-25: T9 pH 5 Sample Condition H2′ and H2″ Walks at 298 K. The H2′

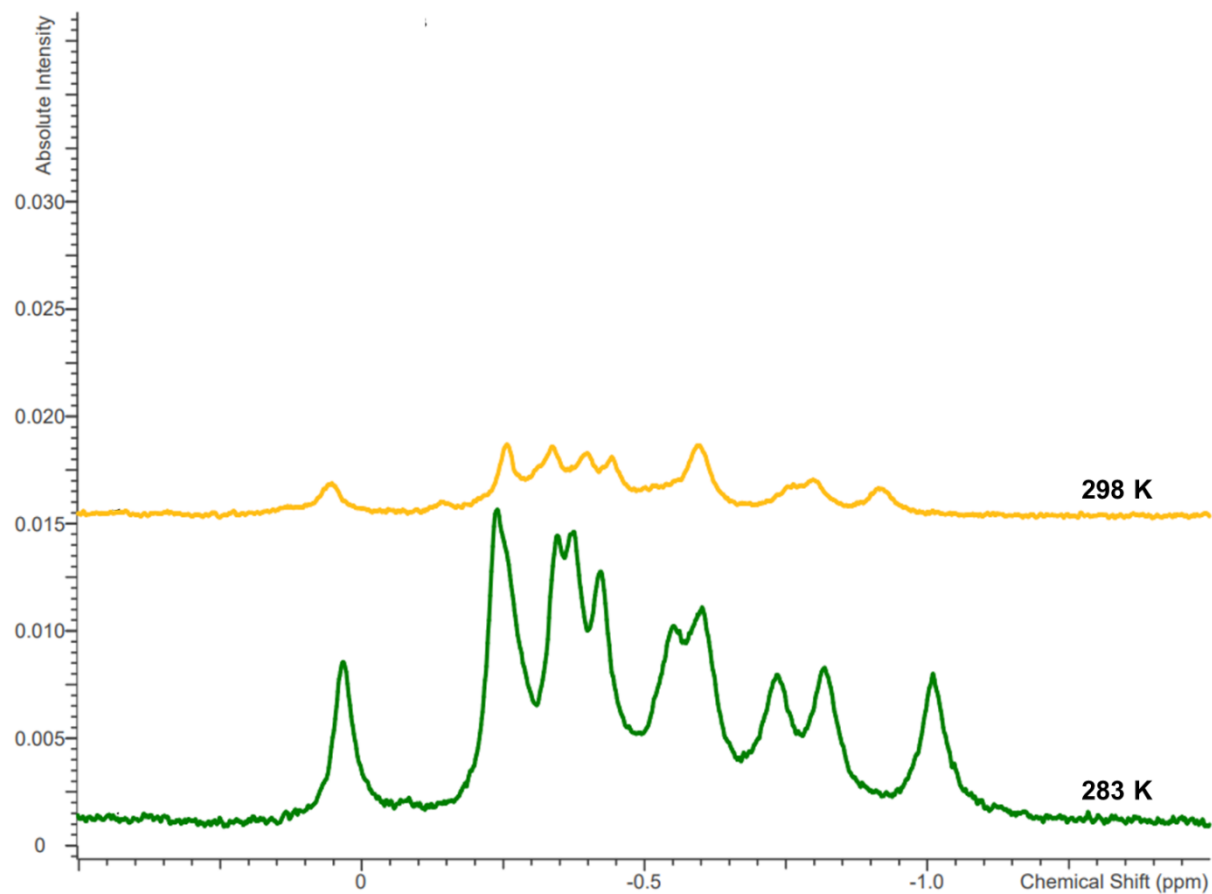walk is on the left and the H2″ walk is on the right.
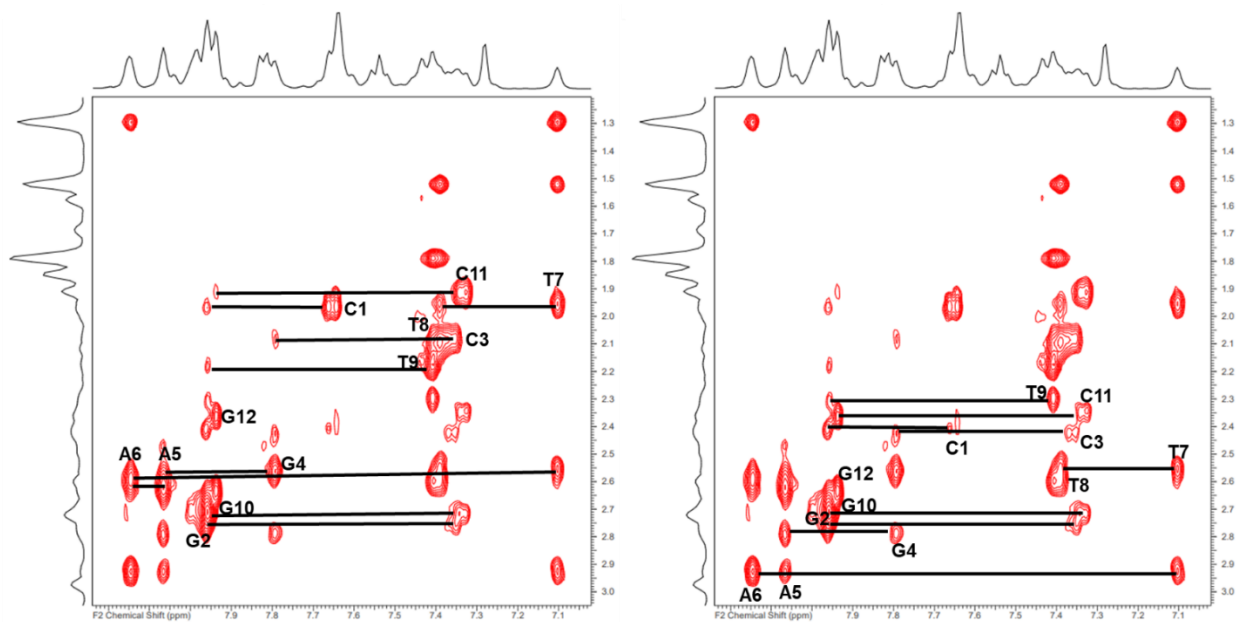
Appendix E-26: T9 pH 5 Sample Condition H4′ and HSQC Assignments at 298 K.

The H4′ assignments are on the left and the HSQC assignments are on the right.

Appendix E-27: Stacked 1D ^{31}P Spectra at 283 and 298 K of T9 in 25 mM Phosphate Buffer with a pH of 5.

Appendix E-28: Table of Proton Chemical Shift Values for 25 mM Phosphate Buffer with a pH of 5 T9 Sample Condition at 283 K. All values are in parts per million (ppm).

Base	H1'	H2'	H2''	H3'	H4'	H6/H8	H5/methyl
C1	5.77	2.02	2.44	n/a	4.07	7.66	5.93
G2	5.91	2.78	2.78	n/a	4.36	7.98	
C3	5.71	2.10	2.45	n/a	4.16	7.37	5.46
G4	5.69	2.59	2.82	n/a	4.36	7.83	
A5	5.94	2.64	2.96	n/a	4.45	8.08	
A6	6.15	2.61	2.96	n/a	4.47	8.18	
T7	5.93	1.98	2.58	n/a	4.22	7.14	1.29
T8	6.16	2.12	2.61	n/a	4.25	7.43	1.52
T9	5.57	2.18	2.30	n/a	4.05	7.43	1.78
G10	5.91	2.72	2.72	n/a	4.41	7.98	
C11	5.76	1.95	2.37	n/a	4.16	7.36	5.46
G12	6.16	2.36	2.66	n/a	4.18	7.96	

Appendix E-29: Table of Proton Chemical Shift Values for 25 mM Phosphate Buffer with a pH of 5 T9 Sample Condition at 298 K. All values are in parts per million (ppm).

Base	H1'	H2'	H2"	H3'	H4'	H6/H8	H5/methyl
C1	5.79	1.96	2.41	4.71	4.07	7.65	5.93
G2	5.91	2.76	2.76	4.98	4.35	7.96	
C3	5.73	2.09	2.43	4.82	4.16	7.36	5.47
G4	5.66	2.57	2.79	4.99	4.35	7.79	
A5	5.93	2.62	2.93	5.04	4.42	8.07	
A6	6.15	2.59	2.93	5.03	4.46	8.14	
T7	5.91	1.96	2.55	4.84	4.20	7.11	1.30
T8	6.16	2.09	2.60	4.89	4.24	7.39	1.52
T9	5.58	2.18	2.30	4.89	4.04	7.41	1.79
G10	5.91	2.71	2.71	4.98	4.39	7.96	
C11	5.79	1.92	2.35	4.82	4.14	7.33	5.45
G12	6.16	2.36	2.64	4.68	4.16	7.94	

Appendix E-30: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer with a pH of 5 T9 Sample Condition. All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.34	-0.33
G2pC3	-0.37	-0.38
C3pG4	-0.26	-0.33
G4pA5	-1.01	-0.92
A5pA6	-0.60	-0.59
A6pT7	-0.82	-0.80
T7pT8	-0.54	-0.59
T8pT9	-0.72	-0.76
T9pG10	0.04	0.06
G10pC11	-0.41	-0.44
C11pG12	-0.23	-0.25

Appendix E-31: Stacked 1D ^{31}P Spectra at 283 and 298 K of T9 in 25 mM Phosphate Buffer with a pH of 9.
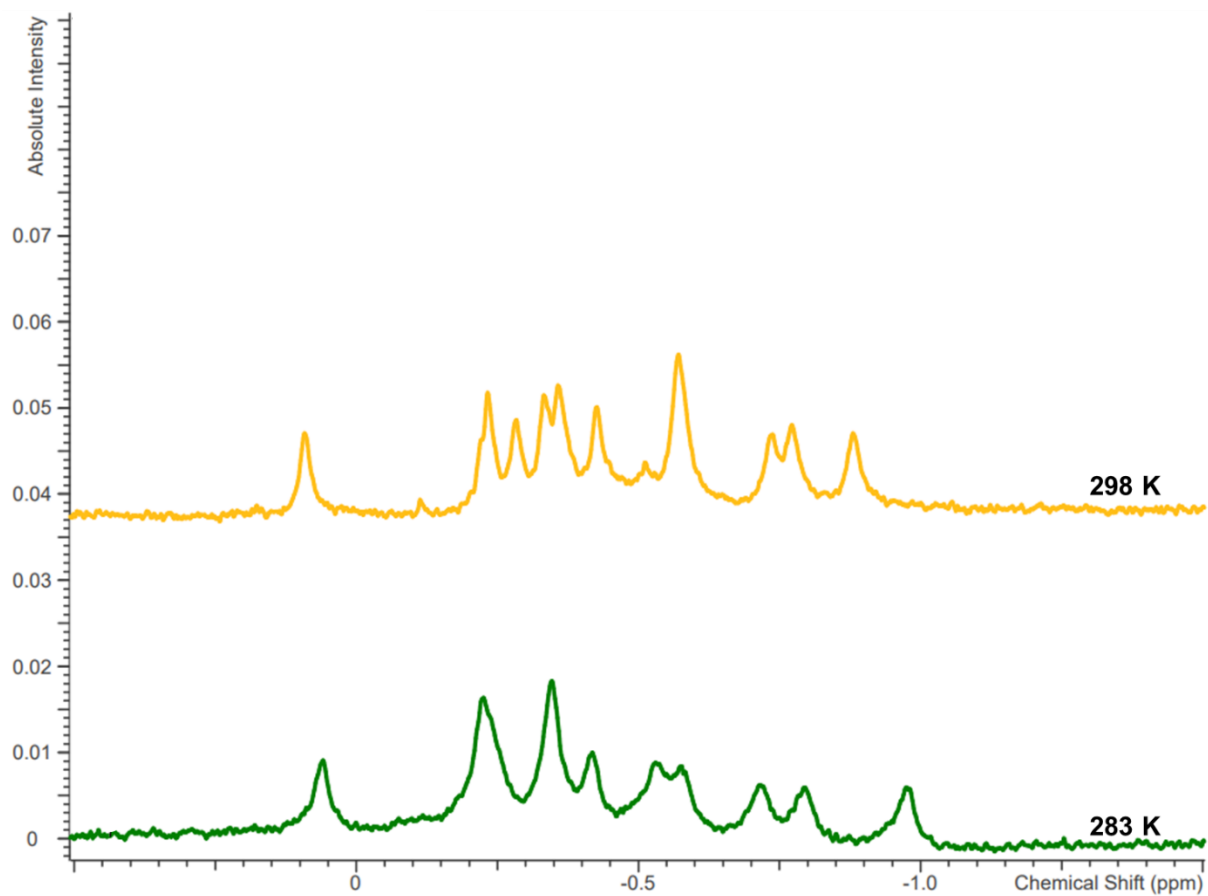
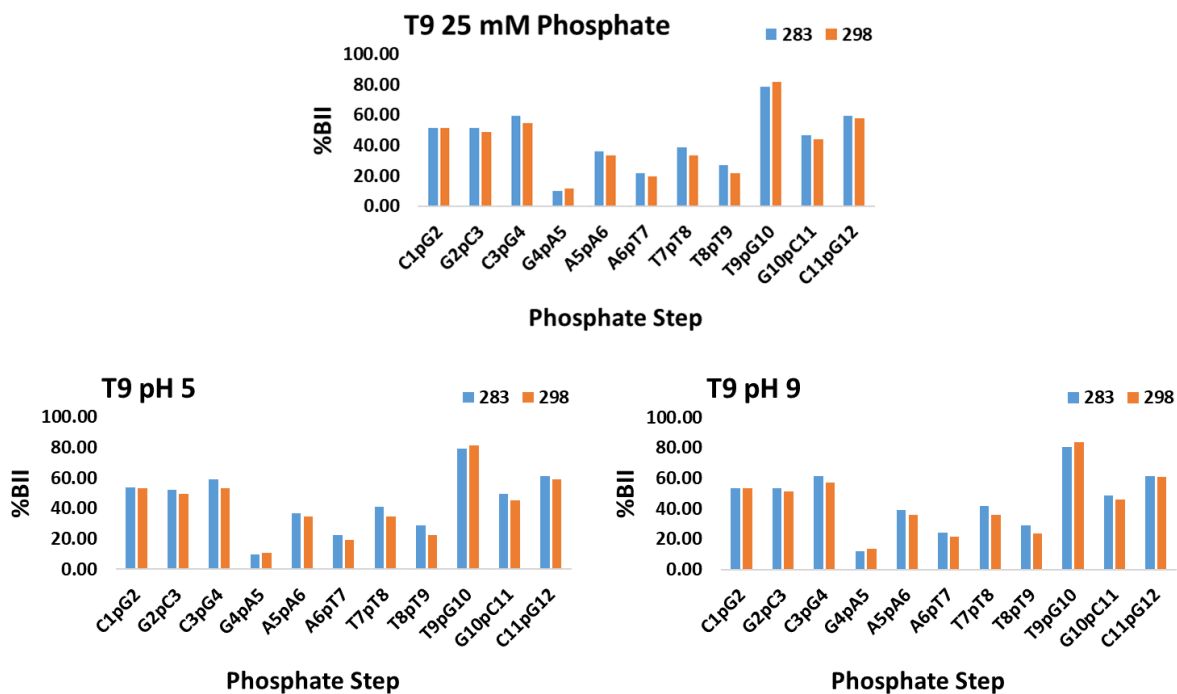**Appendix E-32: Table of Phosphorus Chemical Shift Values for 25 mM Phosphate Buffer with a pH of 9 T9 Sample Condition.** All values are in parts per million (ppm) and were referenced with a phosphoric acid coaxial insert, at 283 and 298 K.

Phosphate Step	283K	298K
C1pG2	-0.35	-0.33
G2pC3	-0.35	-0.36
C3pG4	-0.23	-0.28
G4pA5	-0.98	-0.88
A5pA6	-0.57	-0.57
A6pT7	-0.79	-0.77
T7pT8	-0.53	-0.57
T8pT9	-0.72	-0.74
T9pG10	0.06	0.09
G10pC11	-0.42	-0.43
C11pG12	-0.23	-0.23

Appendix E-33: %BII Plots for all T9 Salt Conditions at 283 and 298 K. These values were calculated via Equation 1. The blue bars correspond to 283 K and the orange correspond to 298 K.
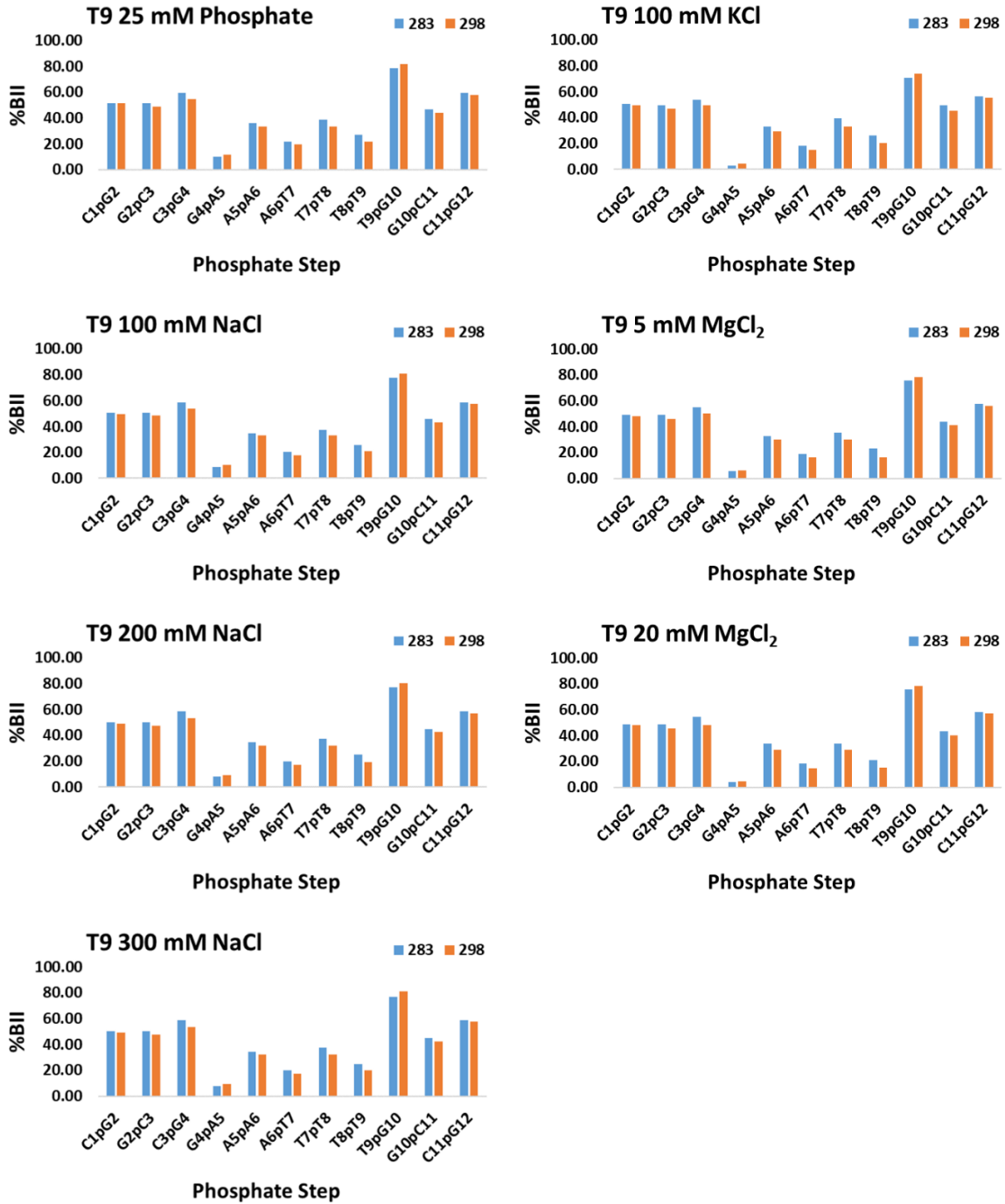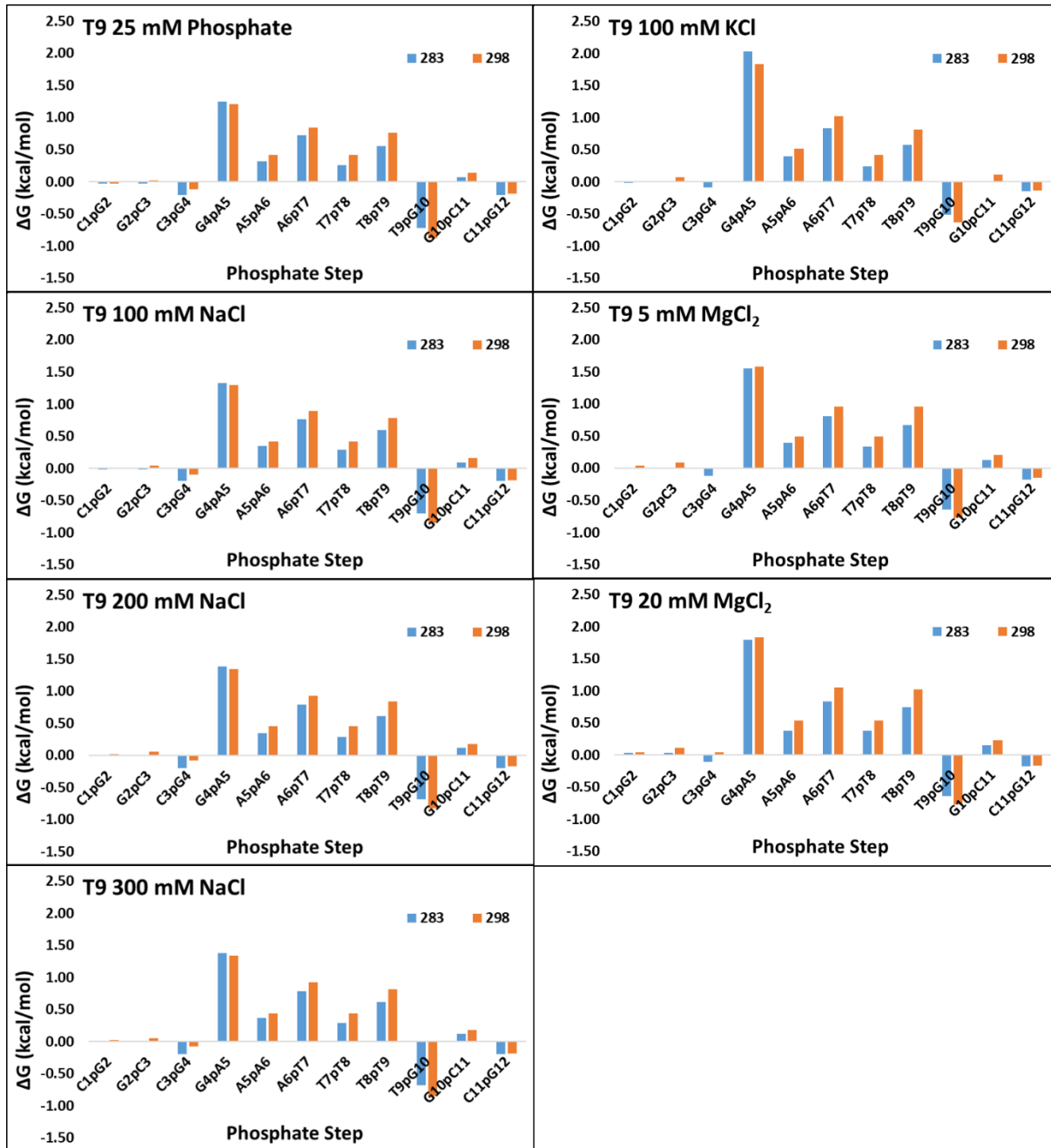
Appendix E-34: %BII plots for all T9 pH Conditions at 283 and 298 K. These values were calculated via Equation 1. The blue bars correspond to 283 K and the orange correspond to 298 K.

Appendix E-35: ΔG plots for all T9 Salt Conditions at 283 and 298 K. These values were calculated via Equation 2. The blue bars correspond to 283 K and the orange correspond to 298 K.

Appendix E-36: ΔG plots for all T9 pH Conditions at 283 and 298 K. These values were calculated via Equation 2. The blue bars correspond to 283 K and the orange correspond to 298 K.