STRUCTURAL STUDY OF THE N^6-(DEOXY-D-*ERYTHRO*-PENTOFURANOSYL)-

2,6-DIAMINO-3,4-DIHYDRO-4-OXO-5-N-METHYLFORMAMIDO-PYRIMIDINE

DNA ADDUCT

By

Richard Kenichi Dempster

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

Master of Science

in

Chemistry

December, 2014

Nashville, TN

Approved:

Michael P. Stone, Ph.D.

Carmelo J. Rizzo, Ph.D.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

DNA Structure and Damage

Deoxyribonucleic acid (DNA), first discovered by Friedrich Miescher in 1871, is often referred to as the fundamental blueprint for all life found on earth. DNA is packed with information that replicates itself from its single strands, is transcribed to messenger RNA and codes for an organism's development from the embryo to full maturity. The right-handed double-helix duplex structure of DNA was elucidated by James Watson and Francis Crick in 1953. This structure includes an alternating phosphate and deoxyribose backbone that spans the entire length of the DNA and runs in opposite directions on either strand. Attached to the riboses are the nucleobases guanine, cytosine, adenine, and thymine. All 4 of these bases are heteroaromatic hydrocarbons, guanine and adenine being purines, and cytosine and thymine being pyrimidines. These bases typically form Watson-Crick pairs of purines to pyrimidines as guanine and cytosine share 3 hydrogen-bonds, and adenine and thymine share two hydrogen-bonds. The base pairs lay perpendicular to the direction of the backbones and pi-stack, further increasing the stability of the DNA structure.
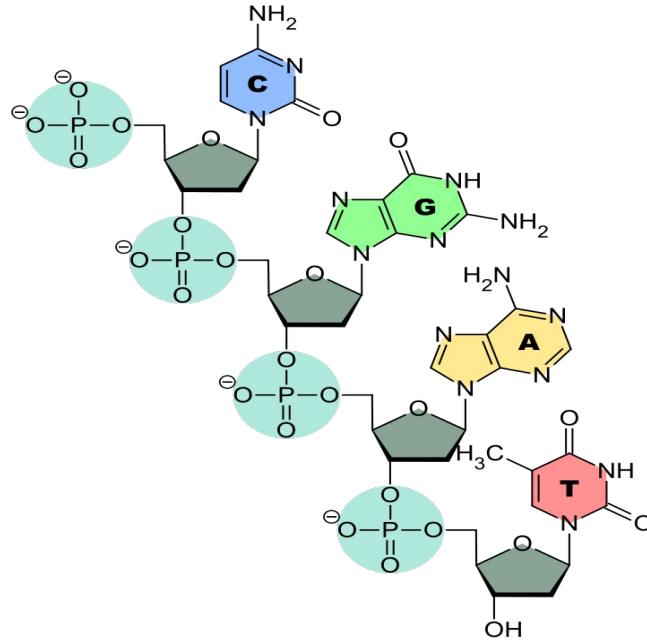
The coil of the DNA structure can vary in some organisms but B-DNA is accepted as the most common form, followed by Z-DNA and A-DNA. In B-DNA, parallel to the backbone of a strand of DNA, the base-pairs rotate at 35.9° increments, making a nearly full turn within 10 base-pairs. With the base-pairs spaced 3.32 Å apart, the 10 base-pair stretch is 33.2 Å in total. This rotation of the backbone and the directionality of the base-pairs to the backbone result in a 20 Å radius and the development of grooves in the overall structure. The two grooves, minor and major, are 12 Å and 22 Å wide respectively. This difference in width directly affects the accessibility of the DNA for proteins such as transcription factors that preferentially make contact with the major groove.
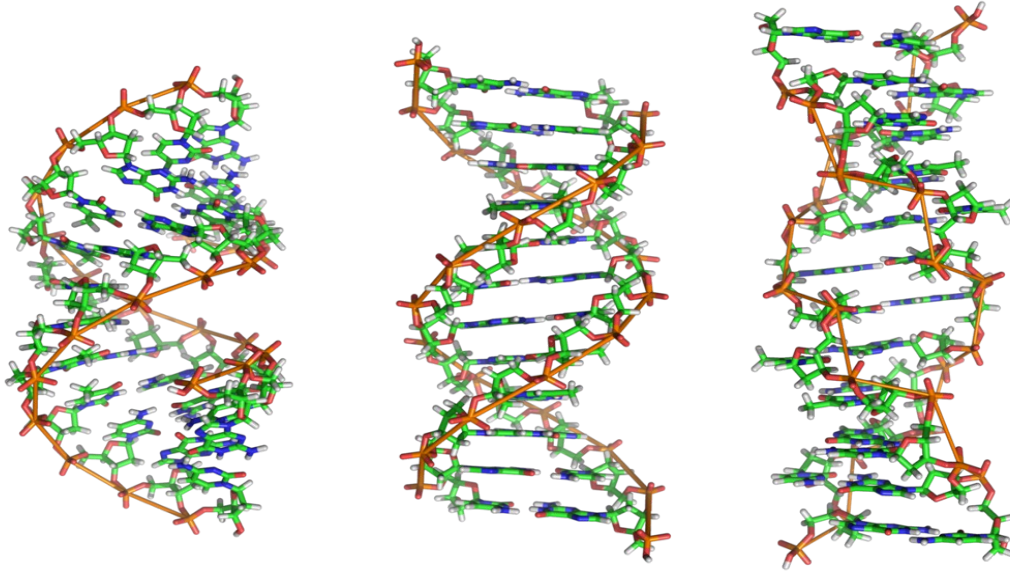
The other two forms of DNA have very different structures from DNA and can occur in alternative environments. Z-DNA is thought to occur in abnormally high salt concentrations or in sequences alternating bases pairs such as lengthy guanine-cytosine repeats. The most notable feature of Z-DNA is in its rotations, adopting the unique left-handed helix found in no other DNA structure. A-DNA can be found in dehydrated conditions such as in DNA crystals and forms a looser twist that is roughly 3° larger than typical B-DNA. This results in a very characteristic structure in which the two backbones run together along the periphery of the DNA in a rail-like arrangement.

DNA damage poses a very real and difficult problem for organisms that rely on its stability and reliability to provide crucial information necessary for cellular function. DNA can suffer damage from UV and other form of electromagnetic radiation that can give rise to abnormal structures such as thymine-thymine dimers. Alkylation and oxidation can occur on multiple locations of the nucleobases, either forming adducts that greatly perturb natural DNA structure or remove the nucleobase all together. Small and

large damage to DNA can result in a halt to transcription, or transcription errors that can lead to further developments such as cancers.

MeFAPY-dG

N^6-(Deoxy-D-*erythro*-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-N-methylform-amidopyrimidine (MeFapy-dG) is a stable DNA adduct resulting from a ring-opening reaction of N^7-Me-dG with hydroxide ion. A number of alkylating agents can methylate deoxyguanosine, including dimethyl sulfate, N-methylnitrosamine, N-methyl-N-nitrosourea, N-methyl-N-nitroso-N-nitroguanidine, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, and S-adenosyl-L-methionine.[11-17] The adduct itself can theoretically assume a number of conformations due to its minimal steric restrictions that can be found in a similar adducts such as AFB-Fapy. There is the possibility of atropisomerization about the C5-N^5 bond, E:Z isomerization of the formyl group, and finally the α,β-anomerization of the sugar, presenting a minimum of 8 potential conformations to consider (**Figure 1**). The Fapy-dG adduct was initially synthesized and studied by Greenberg et al.[1-4] The MeFapy-dG sample currently under investigation was prepared by Carmelo Rizzo's group at Vanderbilt University. They were able to confirm the presence of the α and β anomers in a ratio of ~60:40 based on the UV_{260} absorption changes during a T_m determination.[5]
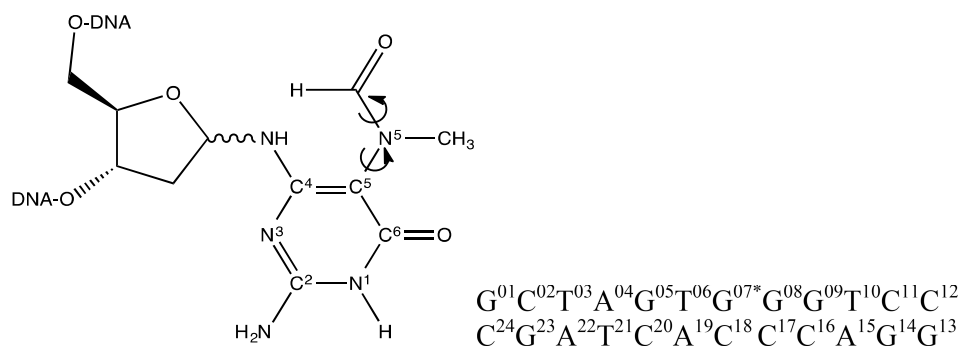
Figure1. MeFapy-dGuo anomeric/rotameric points.

Furthermore, Dr. Rizzo's group conducted an *in vitro* replication study on the MeFapy adduct in two sequences using Klenow Fragment of *E. coli* DNA polymerase I (Kf-⁻) and *Sulfolobus* solfataricus P2 DNA polymerase IV (Dpo4). They reported an error-free bypass and extension for the 5′-T-(MeFapy-dGuo)-G-3′ sequence and a statistically significant occurrence of dAdo misincorporation and a deletion product for 5′-T-(MeFapy-dGuo)-T-3′, suggesting a sequence dependence on the ability of the polymerase to replicate past this lesion.[6] A further investigation was done examining the replication past this lesion by replicative eukaryotic and translesion DNA polymerases. They were able to show that replicative eukaryotic DNA polymerases were strongly blocked by MeFapy and translesion polymerases largely bypassed the lesion without error.[7] Finally, a study was done using 9 different oligonucleotide sequences that were replicated in African green monkey kidney (Cos-7) cell culture, differing only on the 5′ and 3'-sides directly next to the MeFapy lesion with one exception. Three generalizations were made based off of the study: error-free bypass predominated, G to T and G to C transversions were present in all sequences, G to A transitions were present in all but one sequence (the sequence that is being investigated in this thesis), and in most cases single and dinucleotide deletions were generated.[8]

5

A structure of β-cFapy was determined by Carlos de los Santos's group at SUNY Stony Brook. This structure has a lone hydrogen instead of a methyl group on N^5 and includes the substitution of the $O^{4'}$ with a methylene group. This eliminates the possibility o f α,β-anomerization and significantly reduces the complexity of the problem. They reported three major structures: Z-β-cFapy and 2 distinct E-β-cFapy's, differing in their torsion angles about the formyl group and N5.[9-10]

The focus of this thesis is study the structure of the N^6-(deoxy-D-*erythro*-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-N-methylform-amidopyrimidine (MeFapy-dG) adduct and gain insight into the differences of the conformations adopted by the adduct within the surrounding DNA sequence.

NMR Structural Determination

Nuclear magnetic resonance spectroscopy (NMR) is one of two popular approaches, the other being X-ray crystallography, used in structural studies to investigate biological molecules such as oligonucleotides. Generally in NMR spectroscopy, an external field along the z-axis is applied, which aligns the spins of the nuclei in the sample in solution so as to create a net magnetization. The net magnetization is perturbed by a pulse of radio-frequency electromagnetic-radiation applied perpendicular to the z-axis, which creates coherence in the xy plane. The nuclei re-emit the electromagnetic-radiation as they relax to their original spin states, again in the radio-frequency range of the spectrum. This radiation is detected in the time domain and Fourier transformed into a frequency domain spectrum that indirectly reveals information on the structural features of the compound of interest.

The more basic form of NMR is one-dimensional. The data are collected as a function of a single time domain. Commonly, protons (^{1}H) and ^{13}C are analyzed using this technique, which provides a set of signals across a relative energy scale, calibrated in parts-per-million (ppm) of the external magnetic field. Where the signal falls on this scale is dependent on its accessibility to external magnetic field. A nucleus can be shielded, that is, it can 'feel' the minimized effect of an external field caused by the opposing effects of locally produced fields. An increased shielding results in the corresponding signal appearing farther up-field, or lower in ppm. A decreased shielding effect results in the corresponding peak appearing farther down-field, or higher in ppm. The area under any given signal is dependent on the number of nuclei to which it corresponds and can be used to identify the source of the peak. Furthermore, the signals can be split into multiple peaks when the nucleus is J-coupled to other nuclei, usually within 3 direct bonds. However, because a large molecule has many protons, the number of peaks produced on a single spectrum becomes cumbersome and indiscernible. To alleviate this problem a new experiment was devised that utilizes two time domains or -dimensions.

The two-dimensional experiments most commonly used are correlative spectroscopy (COSY) and nuclear Overhauser effect spectroscopy (NOESY). However, there are many derivatives of both COSY and NOESY experiments. COSY is used to identify nuclei with spins that are J-coupled through-bond to each other. With regard to DNA, this technique is very useful for identifying the H5 and H6 protons of cytosine. NOESY is used to identify nuclei that are close in space to each other. The most visible dipolar interactions of these nuclei are generally below 5 Å of each other and the resulting signal integrations are proportionally dependent on the inverse-sixth power of the distance between the two nuclei. This simple proportionality makes NOESY a

powerful technique that can be used to assign protons along the length of an entire molecule.

In an oligonucleotide, the protons attached to the nucleotide bases are close enough in proximity to the protons of the corresponding ribose sugar and the 5′-neighboring ribose sugar. Therefore, the protons can be traced down each strand of the double helix from ribose to nucleotide base and back to the ribose from the 5′-end to the 3′-end. Once the protons have been identified along a single strand, the protons involved in the Watson-Crick hydrogen-bonding of the nucleotide bases, which exchange with water, can be observed by the use of appropriate solvent-suppression pulse sequences.

MATERIALS AND METHODS

Sample Origin

The unmodified oligonucleotide sequence, 5′-GCTAGTGGGTCC-3′, and complimentary strand were synthesized by Midland Reagent Company. The modified oligonucleotide sequence, 5′-GCTACTG*GGTCC-3′, was synthesized by Professor Carmelo J. Rizzo and co-workers at Vanderbilt University. The synthesis in **figure 2** involved methylation of the N2-dimethylaminomethylene-dGuo using methyl iodide. This was treated with sodium hydroxide and neutralized with hydrochloric acid to afford the ring-opened product then converted to a suitable phosphoramidite reagents for solid-phase synthesis.[5] This adduct was inserted into the sequence above.
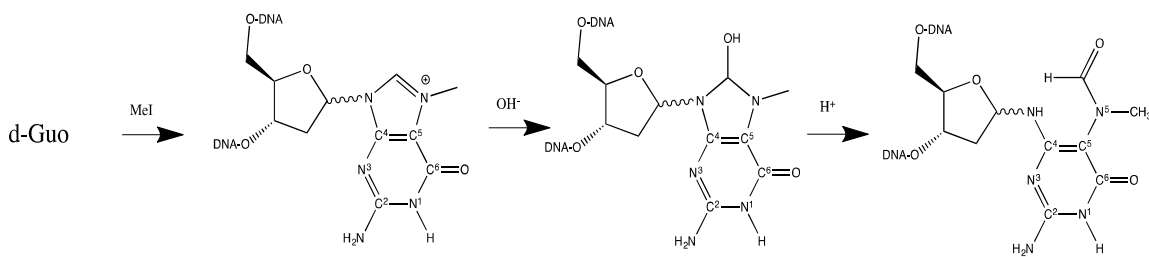

Figure 2. General synthesis of the MeFapy adduct from d-Guo.[5]

Sample Preparations

The unmodified sample strands were checked for purity by reverse-phase HPLC and the mass was confirmed using matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. The optical density and quantity of the sample was determined using a UV visible spectrometer at 260 nm with 1 µL of sample solution dissolved into 1 mL H_2O solution of 10 mM NaH_2PO_4, 100 mM NaCl, and 50 uM EDTA in a cuvette. After confirming the purity and masses, the two strands were annealed at room temperature for 10 minutes. The duplex sample was desalted on a G25 Sephadex column.

The modified sample was previously prepared and received by Liang Li in the Stone laboratory. The sample purity and mass was confirmed by reverse phase HPLC and MALDI-MS respectively followed by desalting on a G25 Sephadex column.

NMR Spectroscopy

Preparation of Unmodified Sample

The unmodified oligonucleotide was dissolved in 500 µL of 10 mM NaH_2PO_4, 100 mM NaCl, and 50 uM EDTA before being lyophilized to dryness. For the experiments conducted in deuterated solvent, the sample was deuterium exchanged with 99.9% D_2O (Cambridge Isotope Laboratories, Inc.) four times and dried via spin vacuum before being redissolved in 99.96% D_2O (Cambridge Isotope Laboratories, Inc.) This was then transferred to a 5mm NMR tube for analysis.

For the experiments conducted in non-deuterated solvent, the previous unmodified sample was removed from the NMR tube and dried via lyophilization. The sample was then dissolved in 450 µL of 99.9% H_2O and 50uL of filtered nanopure D_2O,

10

which was necessary to lock the spectrometer. This was transferred back to the 5 mm NMR tube for analysis.

Preparation of Modified Sample

The modified oligonucleotide was dissolved in 200 µL of 10 mM NaH_2PO_4, 100 mM NaCl, and 50 µM EDTA before being lyophilized to dryness. For the experiments conducted in deuterated solvent, the sample was deuterium exchanged with 99.9% D_2O (Cambridge Isotope Laboratories, Inc.) four times and dried via spin vacuum before being redissolved in 99.96% D_2O (Cambridge Isotope Laboratories, Inc.) This was then transferred to a 5 mm NMR tube for analysis.

For the experiments conducted in non-deuterated solvent, the previous modified sample was removed from the NMR tube and dried via lyophilization. The sample was then dissolved in 180uL of 99.9% D_2O and 50 µL of filtered nanopure H_2O of 18 M. This was transferred back to the 5mm NMR tube for analysis.

NMR Setup of Unmodified Sample

NOESY and COSY NMR spectra in deuterated solvent were acquired on the 600 MHz Bruker instrument at 5 ℃, 15 ℃, and 25 ℃ with a constant mixing time of 250 ms. A NOESY spectrum in non-deuterated solvent was acquired on the 900 MHz Bruker instrument at 5 ℃ with a constant mixing time of 250 ms. 1-D spectra in non-deuterated solvent were acquired at 5 ℃ - 60 ℃ in 5 ℃ increments with a constant mixing time of 200 ms with a 0.7 s relaxation delay.

NMR Setup of Modified Sample

NOESY and COSY NMR spectra in deuterated solvent were acquired on the 900 MHz Bruker instrument at 5 ºC, 15 ºC, and 25 ºC with a constant mixing time of 250 ms and varying mixing times of 40 ms, 60 ms, and 250 ms at 5 ºC and 25 ºC. NOESY spectra in non-deuterated solvent were acquired on the 900 MHz and 800 MHz Bruker instruments at 1 ºC, 5 ºC, 15 ºC, and 25 ºC with a constant mixing time of 250 ms. 1-D spectra in non-deuterated solvent were acquired at 5 ºC - 50 ºC in 5 ºC increments with a constant mixing time of 200 ms with a 0.7 s relaxation delay.

MeFAPY DUPLEX

Introduction

The following is an analysis of data collected for the unmodified and modified sequences with in 5′-d(G^{1}C^{2}T^{3}A^{4}G^{5}T^{6}X^{7}G^{8}G^{9}T^{10}C^{11}C^{12})-3′•

5′d(G^{13}G^{14}A^{15}C^{16}C^{17}C^{18}A^{19}C^{20}T^{21}A^{22}G^{23}C^{24})-3′. This includes all relevant COSY/NOESY data and corresponding discussion.

Analysis of Unmodified D$_2$O NOESY

The unmodified NOESY and COSY spectra in D$_2$O did not present any unusual behavior and showed full connectivity in the walking region. COSY showed 8 major peaks in the walking region corresponding to the 8 total cytosines in the duplex sequence (**Figure 3**). This figure shows the expanded walking region outlining the connectivity of the main strand and **Figure 4** shows the same expansion outlining the connectivity of the complimentary strand.
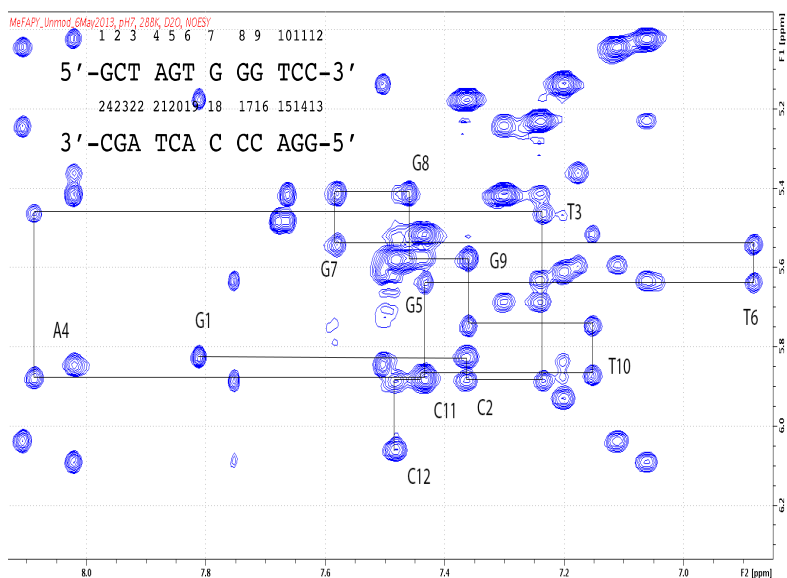
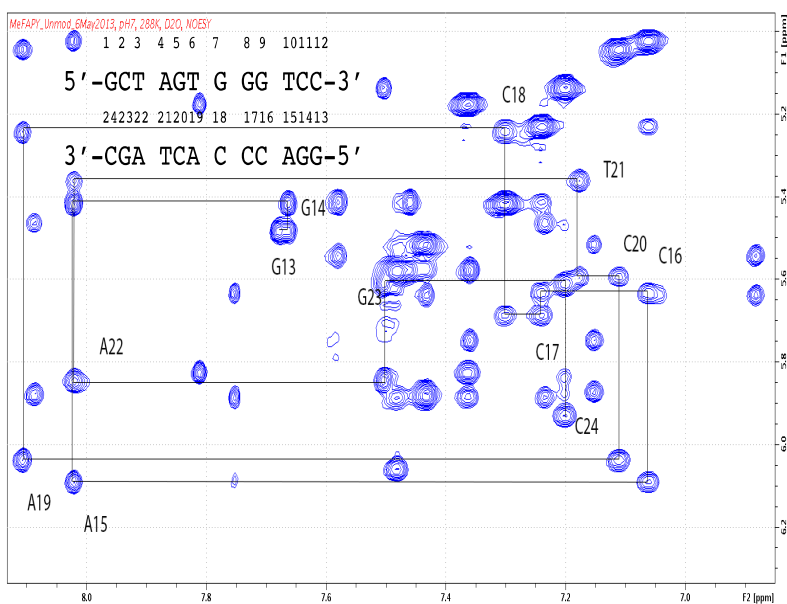Figure 3. Expanded D_2O NOESY of unmodified main strand.

Figure 4. Expanded D_2O NOESY of unmodified complimentary strand.

The unmodified NOESY in 9:1 H_2O/D_2O also did not present any unusual behavior and showed full connectivity in the walking region. All of the cytosine amine protons were assignable and showed cross-peaks with the corresponding imino peaks of the guanines. (**Figure 5**)

Analysis of Unmodified H₂O NOESY

The unmodified 1-D in 9:1 H_2O/D_2O at 5 ºC -60 ºC in 5 ºC increments suggested a typical degradation of the Watson-Crick base pairing and all imino peaks disappeared by 60 ºC. (**Figure 6**) As expected, the G-C base pairs were the strongest at higher temperatures due to the extra hydrogen bond in their Watson-Crick pairing. G^1 was relatively weak and disappeared first, most likely because it was part of a terminal base pair with more dynamic freedom. The same could be said for G^{13}; however, because it is most likely hidden under 3 other base pairs, its stability could not be definitively determined.
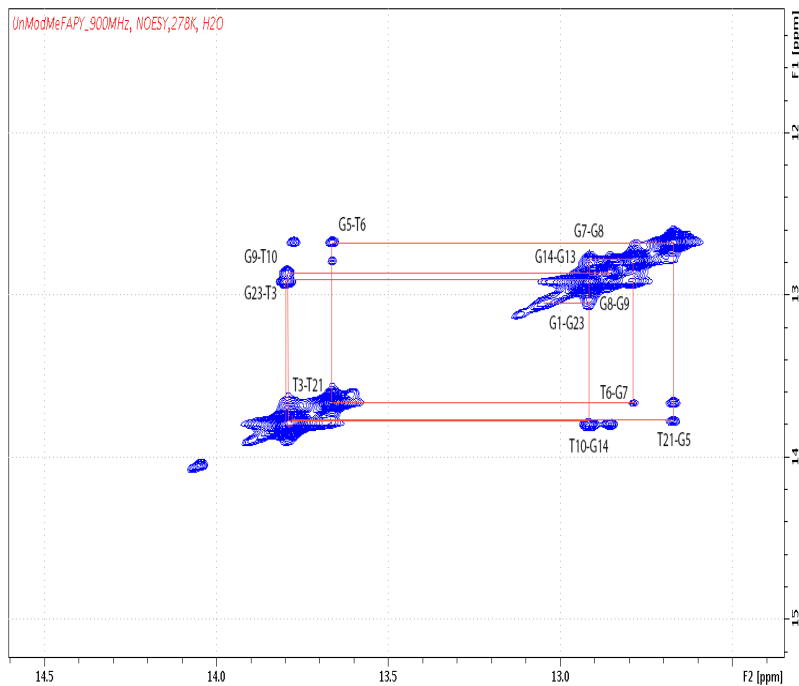

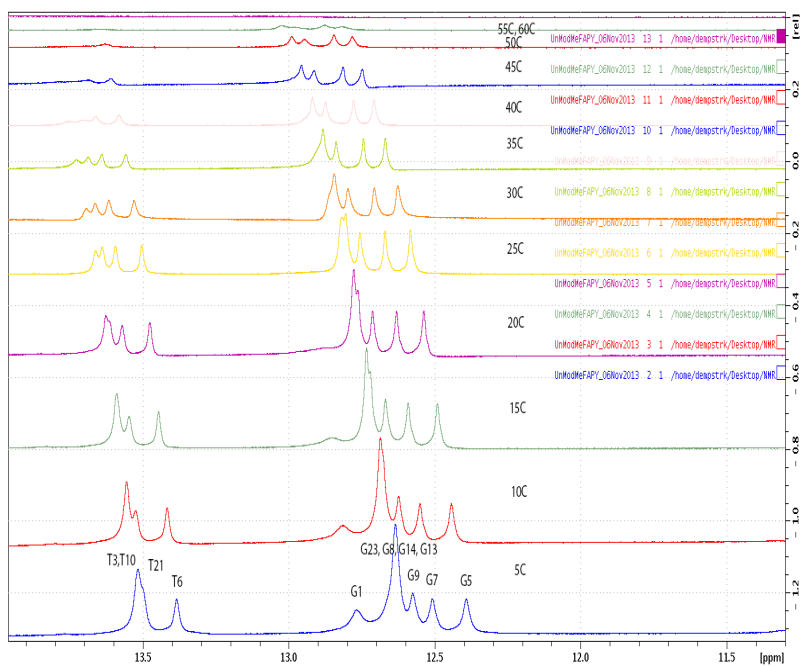Figure 5. Expanded H₂O NOESY of unmodified duplex strands in the imino region.

Figure 6. Expanded H$_2$O 1-D temps of unmodified duplex strands 5 ºC -60 ºC.

Analysis of Modified D₂O NOESY

The modified NOESY and COSY spectra in D_2O were much more complicated and suggested the presence of multiple conformations in the sample. **Figure 7** shows the expanded walking region outlining the connectivity of the main strand modified at G^7 and the complimentary strand. The main strand shows a break after T^6 and continues again at G^8. G^5, T^6, G^8, and G^9 seem to have 2 peaks each with full connectivity. The G^7 formyl proton seems to be completely missing in one conformation while possibly present in the second conformation. Unexpectedly, the suggested formyl peak is at ~6.9 ppm when, at best the peak would be expected to appear at ~8 ppm. The presence of seemingly only 2 conformations is surprising considering all of the conformational possibilities for this adduct. This suggests that there must be some forces such as hydrogen bonding or steric hindrance that are locking the adduct into a limited number of stable conformations. If more conformations are present in the spectrum, it is not apparent due to the weakness of the signals they emit.

Figure 8 shows the expanded walking region outlining the connectivity of the complimentary strand. The complimentary strand does not break but does show the doubled peak pattern seen in the main strand. The doubled peaks are seen for C^{17}-C^{20}. The degree of doubling was not as pronounced as in the main strand where the peaks were completely distinguishable, and resulted in severe overlapping.
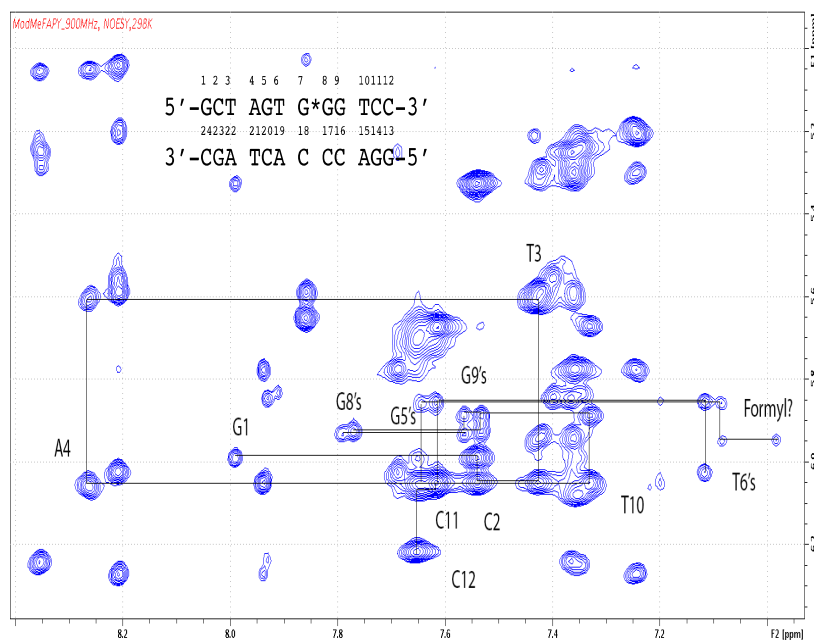
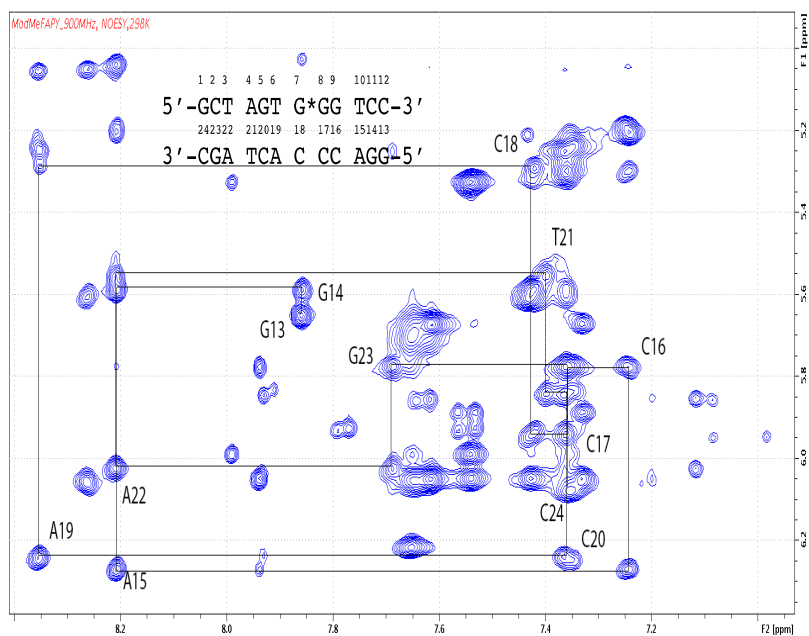Figure 7. Expanded H$_2$O NOESY of unmodified main strand.

Figure 8. Expanded H$_2$O NOESY of unmodified complimentary strand.

The next **figures 9-11** are expansions of the base region just below the walking region. These were taken at 5 ºC, 15 ºC, and 25 ºC respectively. The cross-peak with the hypothesized formyl group appears at 15 ºC and gets stronger at 25 ºC while the

corresponding formyl group gets weaker. This seems to suggest the interconversion of one conformation with another. However, due to the lack of cross peaks at this new location, it is difficult assign this peak with certainty. The only other cross peak was found in the ribose 2',2'' region and has no further cross-peaks associated with it (**Figure 12**). In this figure, one can clearly see the T^{6A} and T^{6B} columns at 7.12 ppm and 7.08 ppm respectively. These peaks continue to show NOE's with the T^6 sugar protons throughout the spectrum. (**Figure 13**)

Figure 9. Expanded D_2O NOESY of modified duplex strands in the base region at 5°C.

Figure 10. Expanded D$_2$O NOESY of modified duplex strands in the base region at 15 $^{\circ}$C.
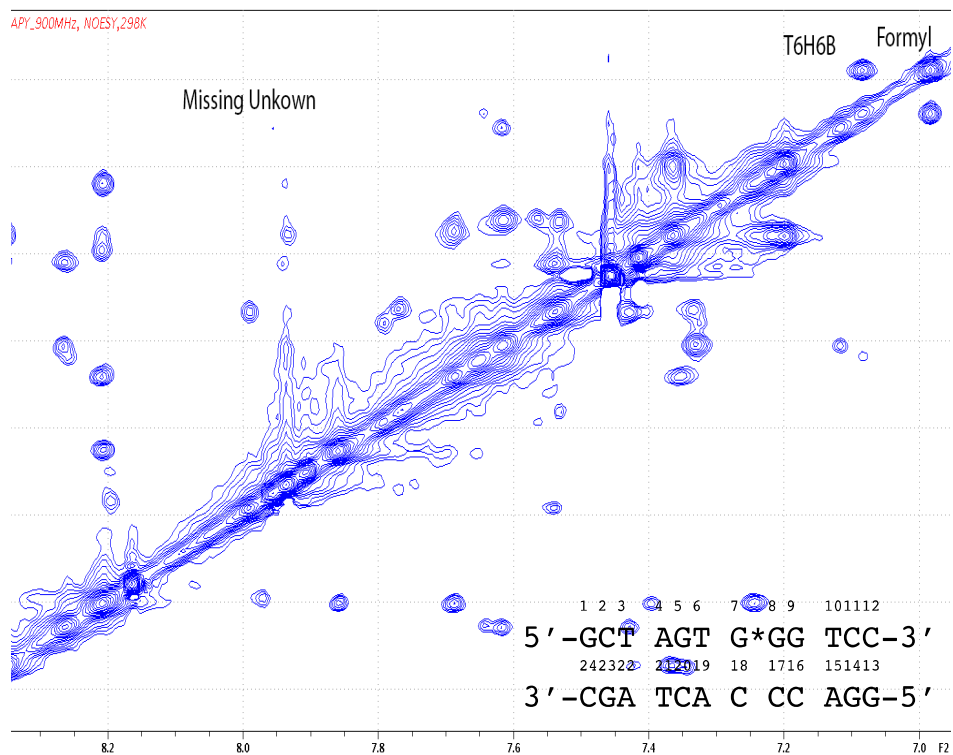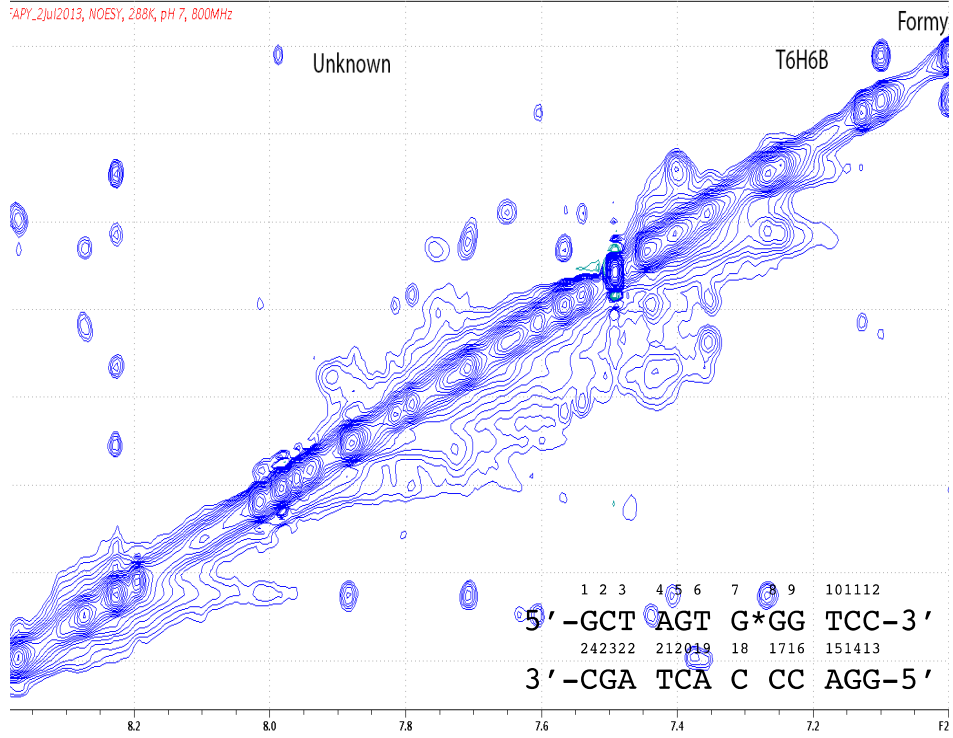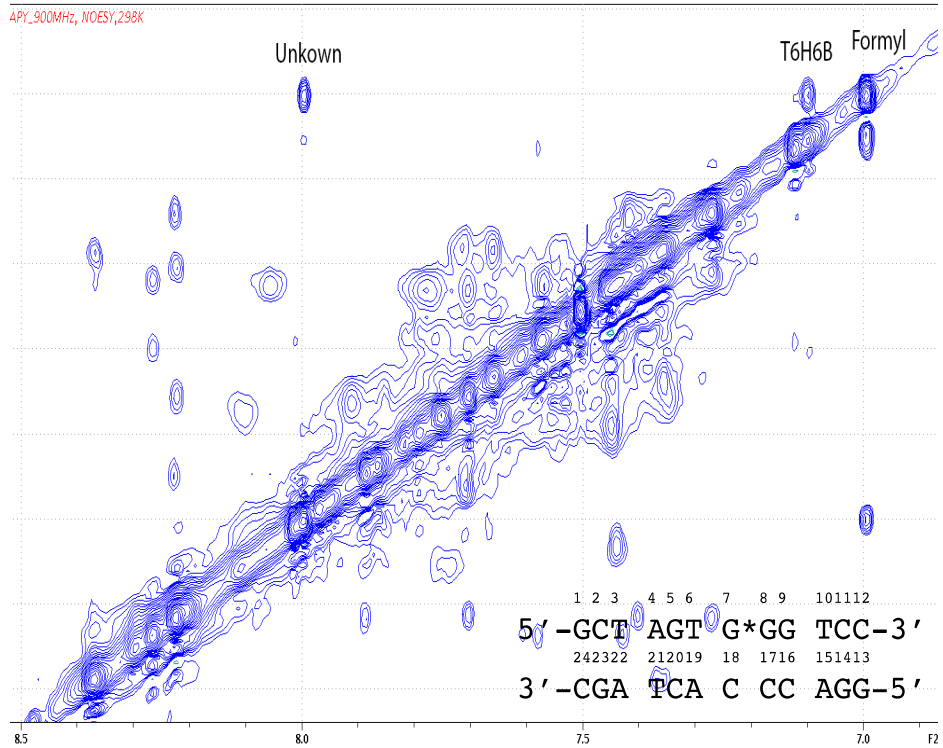

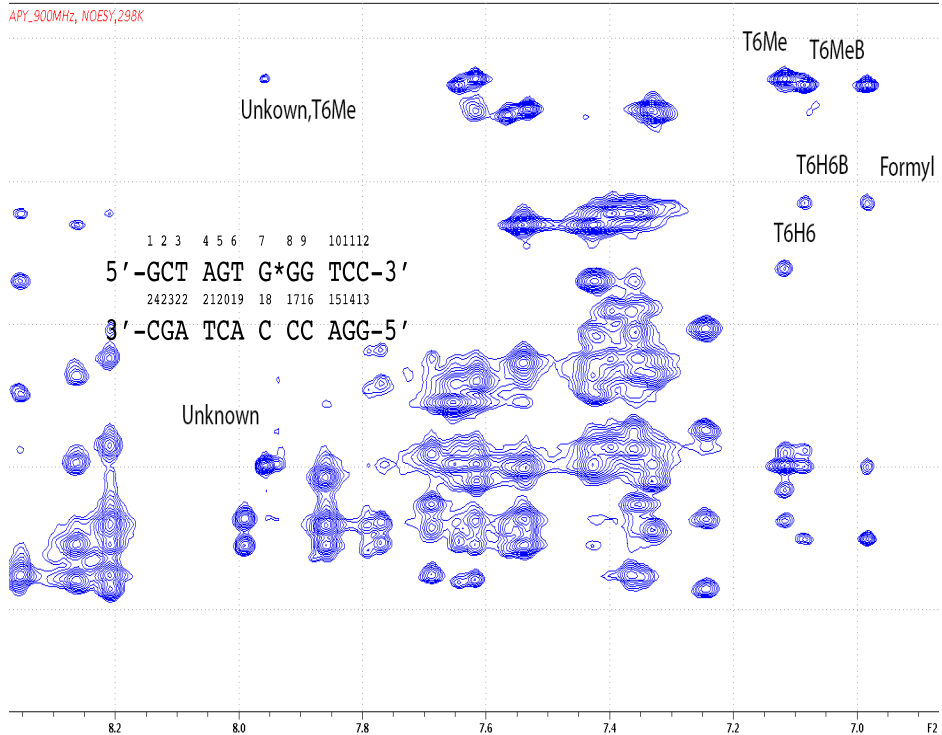Figure 11. Expanded D$_2$O NOESY of modified duplex strands in the base region at 25 $^{\circ}$C.

Figure 12. Expanded D₂O NOESY of modified duplex strands in the 2',2'' region at 25 ºC.
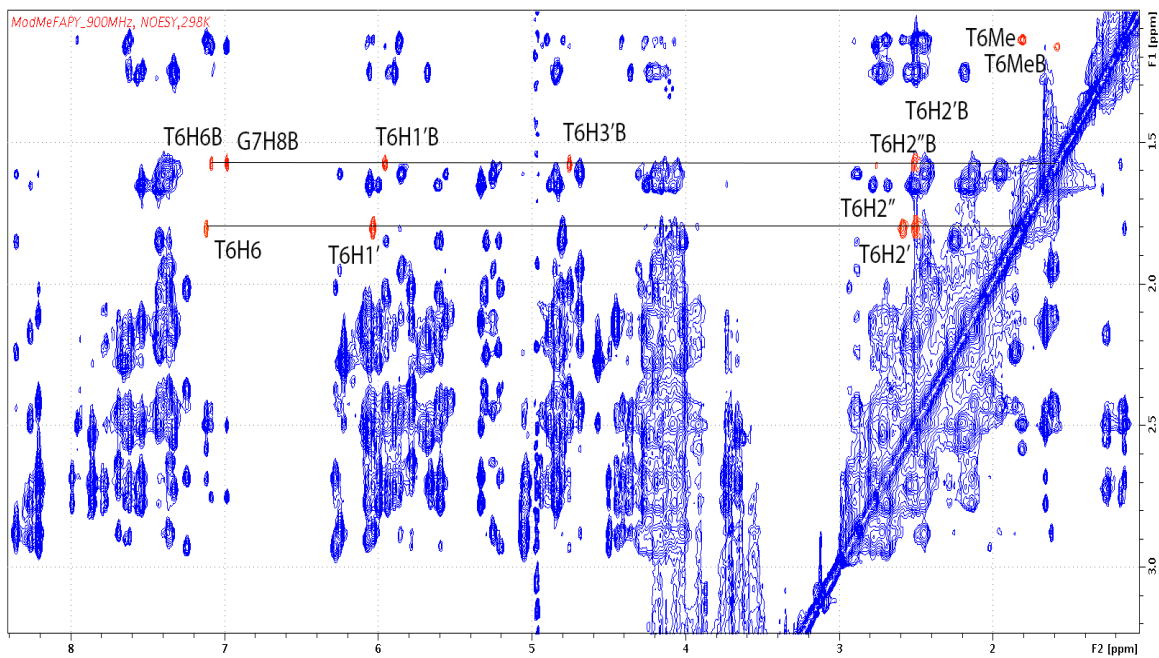
Figure 13. Connectivity of the formyl groups A and B to corresponding T6 methyl and sugar protons.

21

Analysis of Modified H₂O NOESY

The spectra of the modified duplex in 9:1 H_2O/D_2O at 5 ℃ - 50 ℃ in 5 ℃ increments suggested a typical degradation of the Watson-Crick base pairing and all imino peaks disappeared by 50 ℃. This would suggest that the adduct does interfere with the stability of the duplex, despite its distance from the center of the duplex. In this case, it seemed that the T-A base pairs were maintained even at higher temperatures, but this may be a result of 3 of the base pairs stacking. Furthermore, it was expected that the imino peaks would lose their intensities as temperature was increased, but the peaks were in fact sharper at higher temperatures with minimal loss of signal. (**Figures 14-15**) It was not possible to assign all of the peaks due to the complexity discussed above.

Figure 14. Expanded H₂O NOESY of modified duplex strands in the imino region.

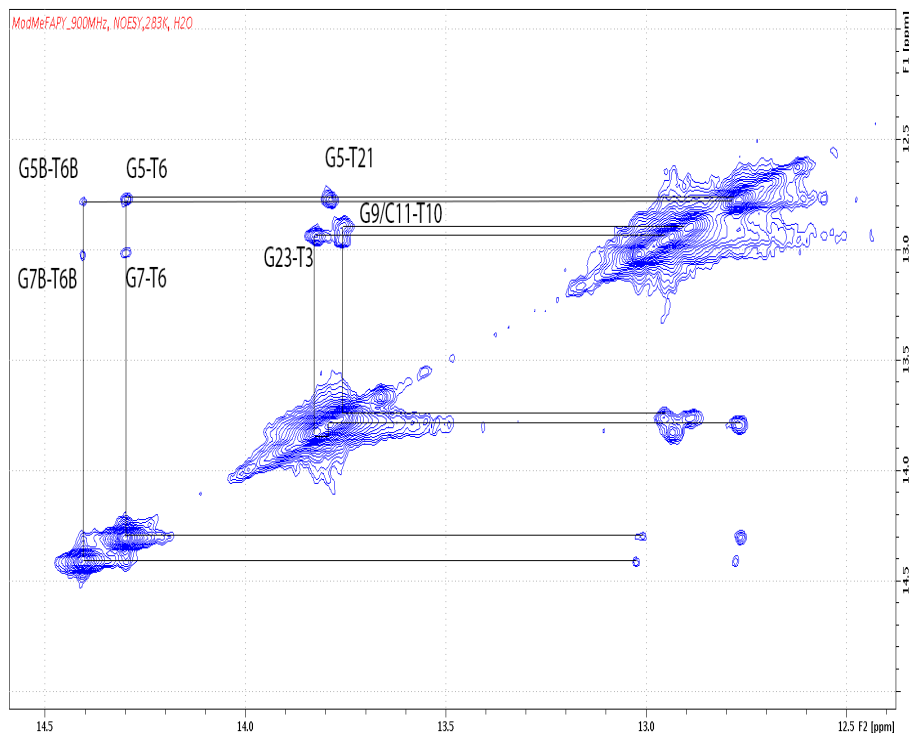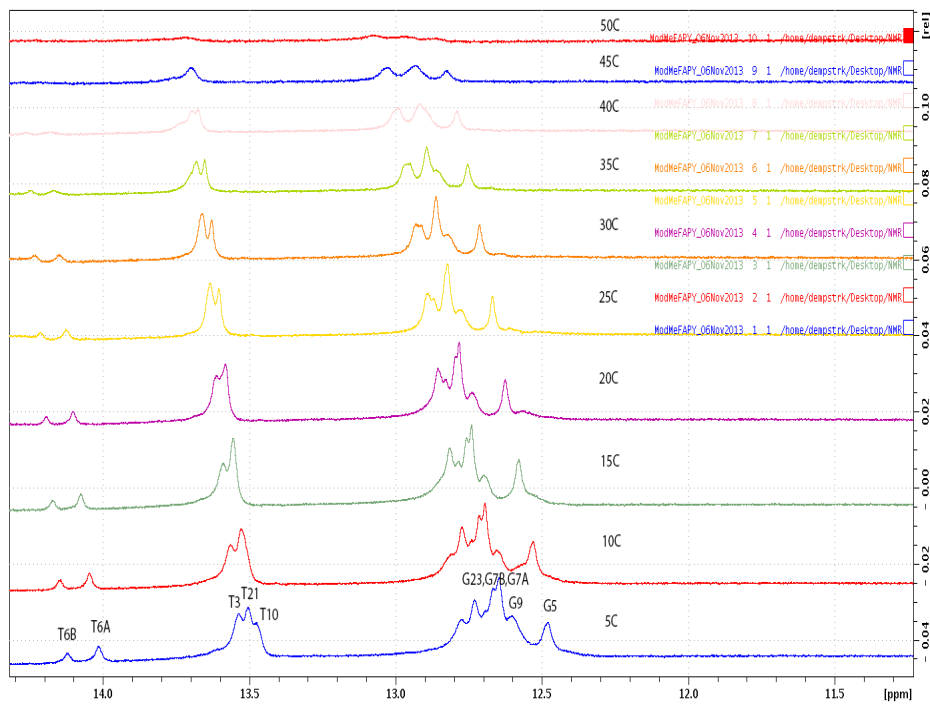Figure 15. Expanded H$_2$O 1D temps of modified duplex strands 5 ºC - 50 ºC.

CHAPTER IV

FUTURE DIRECTIONS

Several items of business remain to be addressed. The melting temperature of the adduct in this sequence has not yet been determined and will most likely be determined using UV spectroscopy. The purity of the samples should be reconfirmed with an orthogonal method such as capillary gel electrophoresis (CGE).

The MeFapy-dG adduct seems to assume at least 2 major conformations and due to the adduct's effects on neighboring bases, causes much complexity in the resulting spectra. First, in light of the temperature 1D results, several NOESY spectra in water at higher temperatures with an increased number of scans should be acquired. This may give sufficient resolution in the imino region that will allow assignment of several key cross-peaks. Furthermore, higher temperatures may result in better peak resolution in the other parts of the spectrum, and more useful information regarding the unknown peak highlighted in **Figures 9-11** may be obtained.

It may also be necessary to use different solvent systems for this particular project. The pH of the buffer solution at different temperatures is also a variable that can be explored. At a lower pH value there is a concern for the chemical stability of the duplex, while at higher pH values there is concern for the maintenance of the duplex conformation, so this approach may be limited. Of course, the biological significance would also be in question if the pH were to be altered too significantly. Finally, if these

approaches do not help, incorporating the adduct in a different sequence should be considered. There is a very real possibility the adduct peaks that are essential to understanding the sequence are hidden under NOE's of other bases. If there is no other way to distinguish the adduct peaks, the conformational information that can be gleaned for this project would probably be minimal at best.

Assuming that it is possible to find conditions in which meaningful NOE's are identified, restrained molecular dynamics calculations should be performed. This would include combining the integration of the volumes of all usable peaks with a hybrid matrix produced by the MARDIGRAS software. The resultant data would be used as restraints in the rMD calculations to produce a 3-D structure. If this cannot be done, it may be possible to at least accumulate useful information regarding the types of conformations that adduct is assumes, and use this to hypothesize what the duplex might look like.

REFERENCES

1) Kazuhiro Haraguchi, Michael O. Delaney, Carissa J. Wiederholt, Aruna Sambandam, Zsolt Hantosi, and Marc M. Greenberg. (2002) "Synthesis and Characterization of Oligodeocynucleotides Containing Formamidopyrimidie Lesions and Nonhyrolyzable Analogues." *J. Am. Chem. Soc.* 124, 3263-3269.

2) Kazuhiro Haraguchi, Michael O. Delaney, Carissa J. Wiederholt, Aruna Sambandam, Zsolt Hantosi, and Marc M. Greenberg. (2001) "Synthesis and Characterization of Oligodeocynucleotides Containing Formamidopyrimidie Lesions (FapydA, FapydG) at Defined Sites." *Nucl. Acids Res., Supplement* 1, 129-130.

3) Nirmala Krishnamurthy, Kazuhiro Haraguchi, Marc M. Greenberg, and Sheila S. David. (2008) "Efficient Removal of Formamidopyrimidines by 8-Oxoguanine Glycosylases." *Biochemistry* 47, 1043-1050.

4) Marc M. Greenberg, Zsolt Hantosi, Carissa J. Wiederholt, and Christopher D. Rithner. (2001) "Studies on N4-(2-Deoxy-D-pentofuranosyl)-4,6-diamino-5-formamido pyrimidine (FapydA) and N6-diamino-5-formamido-4-hydroxypyrimidine (FapydG)." *Biochemistry* 40, 15856-15861.

5) Plamen P. Christov, Kyle L. Brown, Ivan D. Kozekov, Michael P. Stone, Thomas M. Harris, and Carmelo J. Rizzo. (2008) "Site-Specific Synthesis and Charaterization of Oligonucleotides Containing an N^6-(deoxy-D-*erythro*-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-N-methyl-formamidopyrimidine Lesion, the Ring-Opened Product from N7-Methylation of Deoxyguanosine." *Chem. Res. Toxicol.* 21, 2324-2333.

6) Plamen P. Christov, Karen C. Angel. F. Peter Guengerich, and Carmelo J. Rizzo. (2009) "Replication Past the N5-methyl-Formamidopyrimidine Lesion of Deocyguanosine by DNA Polymerases and an Improved Procedure for Sequence Analysis of in Vitro Bypass Products by Mass Spectrometry." *Chem. Res. Toxicol.* 22, 1086-1095.

7) Plamen P. Christov, Kinrin Yamanaka, Jeong-Yun Choi, Kei-ichi Takata, Richard D. Wood, F. Peter Guengerich, R. Stephen Lloyd, and Carmelo J. Rizzo. (2012) "Replication of the 2,6-Diamino-4-hydroxy-N5-methyl Formamidopyrimidine (MeFapy-dGuo) Adduct by Eukaryotic DNA Polymerases." *Chem. Res. Toxicol.* 25, 1652-1661.

8) Lauriel F. Earley, Irina G. Minko, Plamen P. Christov, Carmelo J. Rizzo, and R. Stephen Lloyd. (2013) "Mutagenic Spectra Arising from Replication Bypass of the 2,6-Diamino-4-hydroxy-N5-methyl Formamidopyrimidine Adduct in Primate Cells." *Chem. Res. Toxicol. 26, 1108-1114.*

9) Mark Lukin, Tatiana Zaliznyak, Sivaprasad Attaluri, Francis Johnson, and Carlos de los Santos. (2012) "Solution Structure of Duplex DNA Containing a b-Carba-Fapy-dG Lesion." *Chem. Res. Toxicol.* 25, 2423-2431.

10) Mark Lukin, Conceicao A. S. a. Minetti, David P. Remeta, Sivaprasad Attaluri, Francis Johnson, Kenneth J. Breslauer, and Calos de los Santos. (2011) "Novel Post-Synthetic Generation, Isomeric Resolution, and Chracterization of Fapy-dG within Oligodeoxynucleotides: Differential Anomeric Impacts on DNA Duplex Properties." *Nucl. Acids Res.*, 39, 5776-5789.

11) Louis R. Barrows and Peter N. Magee. (1982) "Nonenzymatic Methylation of DNA by S-adenosylmethionine in Vitro." *Carcinogenesis* 3, 349-351.

12) B. Rydberg and T. Lindahl. (1982) "Nonenzymatic Methylation of DNA by the Intracellular Methyl Group Donor S-adenocyl-L-methionine is a Potentially Mutagenic Reaction." *The EMBO Journal*, 1, 211-216.

13) Stephen S. Hecht. (1998) "Biochemistry, Biology, and Carcinogenicity of Tobacco-Specific N-Nitrosamines." *Chem. Res. Toxicol.,* 11, 560-603.

14) Kent S. Gates, Tony Nooner, and Sanjay Dutta. (2004) "Biologically Relevant Chemical Reactions of N7-Alkylguanine Residues in DNA." *Chem. Res. Toxicol.* 17, 7, 840-856.

15) Lawrence A. Loeb and Bradley D. Preston. (1986) "Mutagenesis by Apurinic/Apyrimidinic Sites." *Ann. Rev. Genet.* 20, 201-230.

16) D. T. Beranek, C.C. Weis, F. E. Evans, C. J. Chetsanga, and F. F. Kadlubar. (1983) "Identification of N5-methyl-N5-Formyl-2,5,6-triamino-4-Hydroxypyrimidine as a Major Adduct in Rat Liver DNA After Treatment with the Carcinogens, N, N-Dimethylnitrosamine or 1,2-Dimethylhydrazine." *Biochem. Biophys. Res. Commun.* 110, 635-631.

17) Fred F. Kadlubar, David T. Beranek, Constance C. Weis, Frederick E. Evans, Ray Cox, and Charles C. Irving. (1984) "Characterization of the Purine Ring-Opened 7-Methylguanine and Its Persistence in Rat Bladder Epithelial DNA after Treatment with the Carcinogen N-Methylnitrosourea" *Carcinogenesis* 5, 5, 587-592.

18) B. Singer and J. T. Kusmierek. (1982) "Chemical Mutagenesis." *Ann. Rev. Biochem.* 52, 655-693.