5-4-2023

Insights into the Local Structural Impact of Ribonucleotide Inclusion in Duplexed DNA

Steven T. Brenden
sbrenden1@student.gsu.edu

Insights into the Local Structural Impact of Ribonucleotide Inclusion in Duplexed DNA

by

Steven T. Brenden

Under the Direction of Markus Germann, PhD

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2023

ABSTRACT

Ribonucleotide inclusion (rNMP) is the most prevalent form of DNA damage. The incorporation of a single RNP impacts the local structure and dynamics of the double-stranded DNA. Our previous work has shown that for an rG intrusion, the perturbation is strongly on the flanking nucleotide sequences. We have prepared a series of duplexed DNA oligonucleotides all containing a single rG or rC inclusion in different flanking sequence environments for enzymatic and structural studies. For the rG DNA substrates I have examined all possible nearest neighbor flanking sequence combinations. It was found that the reaction-efficiency-with RNase HII strongly depends on the flanking sequence by an up to 20-fold difference. NMR analysis was performed to determine the effects a ribonucleotide misincorporation has on a replication fork. The proximity of an RNP to a replication fork greatly distorts the fraction-south (fs) characteristics of the sugar residues and immediate phosphodiester backbone.

INDEX WORDS: DNA damage, NMR spectroscopy, Ribonucleotide inclusion, RNase HII, Replication fork

Insights into the Local Structural Impact of Neighboring Nucleotides in Ribo-damaged Duplex DNA

by

Steven T. Brenden

Committee Chair: Markus Germann

Committee: Gregory Poon

W. David Wilson

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

May 2023

DEDICATION

For Nora.

ACKNOWLEDGEMENTS

I would first like to express my deepest appreciation to my advisor and committee chair, Professor Markus Germann, without whose guidance and patience none of this would be possible. I am also extremely grateful to my committee members, Professor Gregory Poon, and Professor David Wilson, who regularly shared their knowledge and feedback along the way.

I would like to offer many thanks to my colleagues in the Germann lab for their help in proofreading, advice, and moral support.

Words cannot express my gratitude to my family, especially my parents, siblings. and grandparents for their belief in and support of me through the many challenges I have faced. Lastly, I would be remiss in not thanking my close friends Jared Smith and Ryan Dekker. Their energy and experiences always helped me calibrate for true north.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 DNA Damage

Deoxyribonucleic acid, DNA, is a polymeric molecule that makes up the genetic code for all of life. Through engagement with metabolic processes and environmental factors DNA is damaged quite frequently on a scale of 104 per cell per day. [1] Cellular mechanisms provide powerful tools to rapidly identify and correct localized DNA damage. DNA damage, if left unrepaired can lead to genetic dysregulation, disease, or cellular apoptosis. Many diseases may arise from unrepaired DNA remaining in the genome such as many cancers (including colon, breast, and ovarian), Alzheimer's disease, and premature or advanced aging. [2] Damaged DNA relies on error correction processes as it cannot simply be constructed without a template.

1.2 Types of DNA Damage

DNA damage occurs when either the sugar-phosphate backbone is disrupted, or the base is modified covalently. DNA may be damaged by a multitude of sources and instances. Such as the presence of reactive oxygen species, ultraviolet radiation, and toxins from the environment. [3] Ultraviolet radiation fuses two pyrimidines into a dimer distorting the strand and bulging out into what is known as a lesion. Another origin for lesion formation is the incorporation of a ribose sugar into the strand that should naturally only contain deoxyribose sugars. A ribonucleotide differs from a deoxyribonucleotide by only one atom. An oxygen present on the sugar ring at the 2′ position is a change drastic enough to alter the mechanical properties of a DNA backbone. [4] Ribonucleoside mono phosphates (rNMPs) are very often mistakenly incorporated into DNA and, in a properly functioning cell, are quickly identified and repaired. Many questions still exist about the biophysical dynamics taking place in the perturbed backbone of damaged DNA.

1.3 Ribonucleotide Incorporation

Ribonucleotide incorporation into DNA is the most common type of damage that can occur. Ribo-adulteration occurs during replication when a ribo sugar is mistakenly added to the growing chain of deoxy sugar nucleotides. [5] The DNA replication process is not interrupted when adding a ribo sugar into the backbone but rather continues thereby leaving an oxygen-containing 2′ carbon. The ambient concentration of ribonucleotide triphosphates (rNTPs) is much greater than that of deoxynucleotide triphosphates (dNTPs) at about 100:1, causing a greater number of ribo base pairing events and ultimately mistaken inclusion during replication. [5] Ribo damage may also occur when an Okazaki fragment primer on the lagging strand is not correctly targeted by ribonucleotide excision repair (RER). [6] This ribo lesion form of DNA damage affects the local structure of the DNA backbone, changes its strand elasticity, and can interfere with DNA repair mechanisms. [4][7] While the base order remains the same, unlike with a substitution mutation, the shape and geometry of DNA can change based on the nature of the damage. Ribo lesions disrupt the flexibility and geometry of the DNA backbone which is imperative for the proper winding, unwinding, and packaging. Further, it will affect the recognition of the DNA by various enzymes that may need to bind. More disruptively it makes the backbone sensitive to cleavage. The reactive oxygen on the 2′ may cause a nick and a disruption of the supercoiling. They also create a disturbance on the phosphate backbone as a bulging lesion and disrupt interactions. [8] The local structural perturbations inhibit DNA replication by reducing helicase activity and stall unwinding. Ribo guanine was previously thought to be the most commonly incorporated RNA base, but it is now known to be ribo adenine which is the most prevalent. [9]

1.4 Identification and Repair of DNA Damage

Understanding the mechanisms by which DNA damage is identified and repaired is of critical importance and may lead to future insights into the causation of cancer. DNA damage is most often corrected in a process known as nucleotide excision repair (NER). During NER endonucleases recognize and nick the DNA backbone of the problematic nucleotide. Then, subsequent proteins denature and remove the nicked strand containing the damage. Finally, the strand is repaired with polymerase complex activity. Another repair mechanism is known as base excision repair (BER) in which a single base is targeted by glycosylases and extracted. This process targets mutated or mismatched single bases and is not the same in scope as NER. One such very important NER enzyme is the endonuclease human RNase, a class three hydrolase. [10] RNases H are a sequence-modulated endonuclease that hydrolyzes the DNA phosphate backbone at a single ribonucleotide insertion. A mutation in the RNase H gene may cause ineffective proteins or an enzyme deficiency and can cause diseases such as Aicardi-Goutières syndrome (AGS). The symptoms of AGS affect newborn infants and include decreases in head growth, weakened muscle development, and a severe reduction in cognitive capabilities. Further, an RNase H2 knock-out yeast clone did not affect survivability, but a mouse knock-out proves to be lethal. [11] Over time, a buildup of rNMPs incorporated into DNA will increase the likelihood of nicking due to the presence of reactive oxygen on the 2′ carbon. Therefore, defective RNases H2 will lead to a weakening of the sugar-phosphate backbones and long-term genetic damage.

2 EXPERIMENTAL

2.1 Materials

Oligonucleotide sequences containing a single ribonucleotide inclusion were synthesized by Dharmacon Research incorporated and Integrated DNA Technologies (IDT). IDT Oligonucleotides were synthesized to 0.1 μmol scale. The Dharmacon oligonucleotides were synthesized to 0.2 μmol scale and contained a 2′-bis(2-Acetoxyethoxy) methyl (ACE) protecting group that was removed using their provided procedure. The IDT oligonucleotides were not protected. The oligonucleotide complements (DNA) were all purchased from IDT. Adenosine triphosphate (ATP) radioactive on the ɣ phosphate group (^{32}P) was purchased from Cambridge Isotope Laboratories, Inc and used for the 5′ labeling of all ribo containing oligonucleotides. Containing the specific activity of 3000 Ci / mmol. Table 1 contains the sequences that were 5′ labeled with ^{32}P, duplexed with their complement, and reacted with human RNase II at non saturating enzyme conditions. Creating a riboguanine oligo nucleotide with each combination of base 5′ and 3′ of the lesion creates 16 DNA duplexes. Additionally, oligonucleotides containing the T tail and flanked ribo guanine are presented creating 19 total ribo g tridecamers. The label column is the given identity per reactions such that the label consists of the 5′ nucleotide, the ribo ‘g’, the 3′ nucleotide, and finally the reaction number for internal record keeping. The complement labeling consists of the ribo adulterated strands 5′ to 3′ direct complement followed by the corresponding reaction number again for internal record keeping. The units for extinction coefficient (EXT. COEFF.) are Liters • (mol^{-1} • cm^{-1}). Extinction coefficients were calculated using the sums of the mononucleotides at 260 nm absorbance at 80 °C.

UV-Vis spec. For all samples concentration and amounts were confirmed through measurement on a Cary 100 Bio UV-Visible Spectrophotometer from Agilent.

2′-ACE Deprotection Buffer: 100 mM acetic acid adjusted to pH 3.8 with tetramethylethylenediamine (TEMED).

Enzymes: Oligonucleotides were labeled using T4 polynucleotide kinase from New England Biolabs. Pure RNase HII was purchased from NEB (NEB – M0887L, *E. coli*). Both enzymes were provided with 10x of their respective reaction buffers.

1x T4 Polynucleotide Kinase (PNK) Reaction Buffer: 70 mM Tris-HCl, 10 mM $MgCl_2$, 5 mM DTT, pH 7.6 @ 25 °C

Nonradioactive adenosine triphosphate (cold ATP): 25 mM stock from NEB

1x Thermopol reaction buffer: 20 mM Tris-HCl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton® X-100 pH 8.8 @ 25 °C

Denaturing polyacrylamide gels. 50 mL of continuous 15% polyacrylamide gels were cast with 8 M urea to ensure denaturing conditions. Gels were mixed with 1X TBE (90 mM Tris-borate with 5 mM EDTA pH 8), 29:1 Bis-acrylamide solution, and 20 μL TEMED (catalyst, 99%). Once mixed, the gels were degassed under vacuum for 10 minutes. Finally, gels were cast with 200 μL fresh APS solution (initiator, 10%).

Native polyacrylamide gels. 15% Polyacrylamide gels were made following the same recipe but without Urea.

Imaging: Polyacrylamide gels were imaged using Fujifilm BASII photoreactive plate and read on a Fujifilm FLA-7000 laser-scanner and running their proprietary imaging software version 1.1. Equipment located in the Petit Science Center 5th floor core facility.

NMR: All NMR experiments were performed using either a Bruker Avance 600 spectrometer QXI probe head or a Bruker Avance 500 spectrometer with TBI probehead.

Table 1: Ribo G adulterated oligonucleotide sequences that were treated with Human RNase II.

RIBO G ADULTERATED OLIGONUCLEOTIDES

EXT. COEFF. ($M^{-1} \cdot CM^{-1}$)	5′												3′	Label
121 700	C	G	A	T	G	G	A	rg	C	T	C	C	G	AgC-1
122 600	C	G	A	T	G	G	C	rg	A	T	C	C	G	CgA-2
124 000	C	G	A	T	C	C	rg	G	T	A	G	C	G	CgG-3
171 000	C	G	T	T	A	G	rg	C	C	T	G	C	G	GgC-4
132 100	C	G	A	T	C	A	rg	A	T	A	G	C	G	AgA-5
122 300	C	G	A	T	C	T	rg	T	T	A	G	C	G	TgT-6
126 900	C	G	A	T	C	G	rg	G	T	A	G	C	G	GgG-7
119 300	C	G	A	T	C	C	rg	C	T	A	G	C	G	CgC-8
119 600	C	G	A	T	C	T	rg	C	T	A	G	C	G	TgC-9
129 400	C	G	A	T	C	G	rg	A	T	A	G	C	G	GgA-10
129 600	C	G	A	T	C	A	rg	G	T	A	G	C	G	AgG-11
127 600	C	G	A	T	C	A	rg	T	T	A	G	C	G	AgT-12
126 800	C	G	A	T	C	T	rg	A	T	A	G	C	G	TgA-13
122 000	C	G	A	T	C	C	rg	T	T	A	G	C	G	CgT-14
124 900	C	G	A	T	C	G	rg	T	T	A	G	C	G	GgT-15
124 300	C	G	A	T	C	T	rg	G	T	A	G	C	G	TgG-16
	T tailed oligonucleotides													
111 700	T	T	T	T	T	T	A	rg	C	T	T	T	T	Tail AgC
110 200	T	T	T	T	T	C	rg	G	T	T	T	T	T	Tail CgG
107 800	T	T	T	T	T	G	rg	C	T	T	T	T	T	Tail GgC
	Complements													
119 300	C	G	G	A	G	C	T	C	C	A	T	C	G	TCG-1

119 900	C	G	G	A	T	C	G	C	C	A	T	C	G	GCT-2
119 500	C	G	C	T	A	C	C	G	G	A	T	C	G	GCC-3
121 800	C	G	C	A	G	G	C	C	T	A	A	C	G	CCG-4
118 800	C	G	C	T	A	T	C	T	G	A	T	C	G	TCT-5
127 000	C	G	C	T	A	A	C	A	G	A	T	C	G	ACA-6
116 600	C	G	C	T	A	C	C	C	G	A	T	C	G	CCC-7
122 200	C	G	C	T	A	G	C	G	G	A	T	C	G	GCG-8
124 300	C	G	C	A	T	G	C	A	G	A	T	C	G	ACG-9G
118 500	C	G	C	T	A	T	C	C	G	A	T	C	G	CCT-10
116 900	C	G	C	T	A	C	C	T	G	A	T	C	G	TCC-11
121 700	C	G	C	T	A	A	C	T	G	A	T	C	G	TCA-12
124 100	C	G	C	T	A	T	C	A	G	A	T	C	G	ACT-13
124 300	C	G	C	T	A	A	C	G	G	A	T	C	G	GCA-14
121 400	C	G	C	T	A	A	C	C	G	A	T	C	G	CCA-15
122 200	C	G	C	T	A	C	C	A	G	A	T	C	G	ACC-16
	T tailed complement													
149 600	A	A	A	A	A	A	G	C	T	A	A	A	A	ATCGA
148 100	A	A	A	A	A	G	C	C	A	A	A	A	A	AGCCA
148 700	A	A	A	A	A	C	C	G	A	A	A	A	A	ACCGA

Table 2: Ribo C adulterated duplex substrates evaluated with RNase HII.

RIBO C ADULTERATED OLIGONUCLEOTIDES														
EXT. COEFF.	**5′**												**3′**	**Label**
126 000	G	C	C	T	A	A	rc	A	G	A	T	G	G	AcA
122 500	G	C	C	T	A	C	rc	G	G	A	T	G	G	CcG
121 800	G	C	C	A	G	G	rc	C	T	A	A	G	C	GcC
	Complements													
118 500	C	C	A	T	C	T	G	T	T	A	G	G	C	TGT
120 200	C	C	A	T	C	C	G	G	T	A	G	G	C	GGC
115 500	G	C	T	T	A	G	G	C	C	T	G	G	C	CGG

Table 3: Replication Fork Analog, DNA control, and their complementary sequence.

	Replication Forks										
EXT. Coeff.	**5′**									**3′**	**Label**
87 500	C	A	T	C	G	T	G	C	T	T	RF1 D
87 500	C	A	T	C	G	T	G	rc	T	T	RF1 R
	Complement										
100 700	T	A	G	C	A	C	G	A	T	G	RF1 C

2.2 Methods

RNase free H_2O. Deprotection, labeling, and reactions were conducted using Millipore sub-micron ultrapure water containing 18.2 MΩ•cm resistivity (25 °C) contained in vessels thoroughly treated with RNaseZap™ from Invitrogen. This process limits the potential of non-desired cleavage about the ribo containing oligonucleotide strands due to contamination.

2′-ACE Deprotection. The synthesized RNA samples were protected by a 2′-bis(2-Acetoxyethoxy)methyl (ACE) group that must first be removed before carrying out reactions. Removal of the 2′-ACE requires the resuspension in 400 μL of pH 3.8 Deprotection Buffer and incubation of each sample at 60 °C for 30min. Following incubation, the samples are buffer exchanged over a 5 mL HI-TRAP desalting cartridge. Purities of the deprotected oligonucleotides were confirmed on denaturing polyacrylamide gels. The samples are now in RNase free H_2O and stored at 4 °C until ready for labeling. Samples are kept in -20 °C for long term storage.

5′ end labeling. Each oligonucleotide containing a single ribo intrusion must be labeled with a radioactive phosphate group for proper labeling on the Fujifilm photo reactive plate. Each sample were incubated for 30 min at 37 °C in the following reaction mixture: 1x PNK reaction buffer, cold adenosine triphosphate, ^{32}P labeled ATP (120 000 counts per minute), ribo adulterated oligonucleotides (50 pmol), bring to reaction volume of 20 μL with RNase Free H_2O and add T4 Polynucleotide Kinase (PNK, 2 units). The reaction was terminated with a 10 min incubation at 95 °C. Labeling reaction volumes are 20 μL per 50 pmol needed, keeping the final concentration at 2.5 pmol labeled oligonucleotide per 1 μL. This allows for ease of preparation for the RNase HII digestion.

RNase HII digestion. 10 pmol labeled oligonucleotides were duplexed with 11 pmol of their respective complement strands allowed to mix at room temperature. The duplexed strands were then incubated at room temp (25 °C) with 0.05 U RNase HII in 1x Thermopol reaction buffer (pH 8.8, 20 mM Tris-HCl, 10 mM ammonium sulfate, 10 mM potassium chloride, 10 mM magnesium sulfate, 0.1% Triton X-100) at a final reaction volume of 20 μL for 3 hours taking quenched samples every 1 hour. Reactions are quenched with equal reaction volume (20 μL) of Formamide.

Denaturing Gel Electrophoresis. Reaction samples were mixed 1:1 with 99% formamide for gel loading. Gels were run for 4 hours at 200 volts in 1x TBE buffer. Images were captured by developing the gels on Fujifilm BAS II photoreactive film.

Native Gel Electrophoresis. Samples were mixed with a 40% sucrose solution and loaded on the 15% native gel. Native gels were run at 100 V for 2 hours at room temperature.

mFold. Oligonucleotide structure approximations were gathered using the UNAfold webserver suite of tools.[12]

NMR Spectroscopy: For observation of exchangeable protons water samples were prepared (90% H_2O: 10% D_2O) in 50 mM NaCl, 10 mM sodium phosphate, and 0.1 mM EDTA at a pH of 6.6. All ^{1}H spectra were referenced to an internal 4,4-dimethyl-4-silapentane-1-sulfonice acid (DSS) addition. ^{31}P spectra were referenced to 85% H_3PO_4 capillary insert. Samples were confirmed to a concentration of 1 mM by UV spectroscopy and prepared to a 1:1 concentration with their complement thereby forming a duplex.

Imino proton spectra were acquired using a 1-1 pulse sequence at a range of temperatures from 278 K and every 5 degrees up to 318 K. ^{31}P spectra were obtained with the same gamut of temperatures and a proton decoupled pulse sequence. Imino proton were assigned by acquisition

of spectra using a 1-1 NOESY pulse program at 293 K (2k **X** 512, number of scans = 144) with a mixing time ($\boldsymbol{\tau}_M$) of 150 ms, relaxation delay of 3 μsec, and 1 second recycling delay.

Samples in D_2O were lyophilized and resuspended in 350 μL 99.996% D_2O. (pH* 6.23, 5 mm Shigemi microtube). A direct reading from an H_2O based glass electrode meter of pH in a mostly D_2O environment will read as more acidic. This reading is known as pH*. [13]

NOESY experiments were performed at 303 K (2k **X** 600, number of scans = 24) using mixing times of 75, 150, 400 ms. Finally, a 250 ms mixing time with a 1.5 second recycling delay was used for assignment.

COSY spectra were acquired using pre saturation solvent suppression (2k **X** 1300, number of scans = 24) at 303 K with recycling delay of 1.5 sec. Low flip angle COSY was run to determine fraction south J coupling characteristics. TOCSY spectra were also acquired at 303 K (2k **X** 600 number of scans = 16) with a mixing time of 250 ms. ^{31}P-^{1}H correlation experiments were run using the MWGCORRTP pulse program at 303 K (2k **X** 360, number of scans = 60) with a relaxation delay of 3 μsec, and 2 second recycling delay. Constant-time NOESY (CT-NOESY) was recorded as described previously. [7]

Assignment and integration of two-dimensional spectra were done using SPARKY. [14]

3 FLANKING NUCLEOTIDE SEQUENCES MODULATE RNASE H2/II ACTIVITY

Ribonucleotide incorporation into DNA impacts the local structure and dynamics and is therefore expected to affect the interaction with ligands and DNA servicing proteins. This should be particularly evident with minor groove binding compounds and proteins. Repair enzyme efficiency may be dependent on the pliability of DNA about the rNMP inclusion. To evaluate the degree to which the nucleotide sequence environment influences the efficiency of repair enzymes we devised an experiment to appraise neighboring nucleotides and their effect on enzyme activity.

3.1 RNases H

Human ribonucleases (RNases H) come in two main types. Type 1 RNases (H1/I) are a protein containing a single domain and perform excision on RNA/DNA hybrids that contain at least 4 ribonucleotide residues in sequence. They are most commonly employed at the RNA primer of an Okazaki fragment. [15] An Okazaki fragment is a deoxyribose oligonucleotide (about 200 bp in length) built on an RNA primer (about 20 bp in length) and ligated 5′ to 3′ into a new strand of complementary DNA. Okazaki fragments are employed to solve a natural hiccup that occurs during the replication process. As helicase moves along and unwinds the DNA strand a replication fork is created. Like a fork in a road the replication fork separates one strand of duplex DNA into two single strands of DNA. Since DNA can only be ligated in the 5′ to 3′ direction, one of the strands can be processed continuously as the replication fork moves along (leading strand) but the other strand must be ligated in chunks (lagging strand). The Okazaki fragment is the ‘chunk’ of newly synthesized complementary DNA, and the RNA primer must be removed before it becomes a part of the DNA strand. RNase H1/I excises the ribonucleotides

from the strand and a follow up pass from DNA polymerase and finally DNA ligase completes the lagging strand synthesis.

Type 2 RNases (H2/II and H3/III) contain three subunits, one catalytic (RNase H2A) and two (RNases H2B/H2C) that serve for binding interactions with replication and repair complexes. [15] The two ancillary subunits (RNases H2B/H2C) are tightly intertwined and interact with RNase H2A away from the nucleic acid interface. RNases H2/II are unique in that they can nick the backbone of damaged DNA containing single ribonucleotide inclusions. In fact, it is currently the only known non-sequence specific enzyme to do so. RNase H3/III are also a type 2 RNase H and cleave multiple ribo nucleotides in sequence, similar to the mechanism of RNase H1/I, located on a bulge such as in a stem loop. [11] By convention the nomenclature of ribonucleases has maintained that Roman numerals are used to classify prokaryotic enzyme homologs (HI and HII) and Arabic numerals classify eukaryotic (H1 and H2). [16]

3.1.1 Recognition and hydrolysis

Both types of RNase make contact along the minor groove of duplexed DNA before reacting with the phosphate group 5′ of the ribo sugar. Glycine, arginine, and glycine on the b subunit bulge out and contact the DNA minor groove. [17] RNase H2/II recognizes the single ribonucleotide through a small matrix of interactions that include three conserved residues and a hydrogen bond forming tyrosine. The DNA-Protein interaction is stabilized by the critical tyrosine residue (Y163) that forms a hydrogen bond directly with the 2′-OH on the sugar ring of the damaged nucleotide. Once bound, reactions occur at an active site comprised of aspartate and glutamate residues. Commonly known as a DEDD motif, the cleavage reaction mechanism is a two-Mg^{2+}-ion catalysis of the phosphate group 5′ of the ribo sugar. [11] Divalent magnesium (Mg^{2+}) or manganese (Mn^{2+}) is required for catalysis. RNase H2 is more active in the presence of

Mg^{2+} than Mn^{2+}. Two metal ions form coordinated bonds about the scissile phosphate (5′ of the OH containing ribo sugar) and allow for nucleophilic attack from a deprotonated and activated water molecule. [17]

Reactions catalyzed by RNase H2/II occur at the phosphate 5′ of the ribose sugar. The reaction products are a nicked DNA backbone with a 3′-OH group where the phosphoryl bond to the ribonucleotide 5′ once was and a 5′ phosphate group attached to the ribonucleotide. The DNA (with the nicked backbone, figure 3) remains duplexed due to the hydrogen bonding of the base pairs not being disrupted, therefore denaturation does not occur.

RNase H2 has been crystallized with a duplexed DNA substrate bound in its pocket and shows a direct binding at the ribo sugar along the minor groove. Figure 1, below, depicts the binding of an oligonucleotide substrate to an RNase H2 active site. [17] Figure 1A shows an enzyme with an applied surface while the duplex DNA is a simplified ribbon cartoon. In this zoomed out arrangement one can clearly see the security with which RNase H2 approaches the DNA minor groove. The DNA is double stranded, base paired and demonstrates B-DNA twist. In Figure 1B the DNA ribbon model is exchanged for a ball and stick to illustrate the positioning of the ribonucleotide atoms within the pocket. Note that the 2′-OH points inward toward the pocket forming a hydrogen bond with Y163. A green dash is placed from the 2′-OH to Y163 for clarity. Figure 1C displays only the critical active site residues and the ribonucleotide as ball and stick. Aspartate (D18, D107, D124), glutamate (E19), and tyrosine (Y163) side chains are labeled, and a green dash is added to the 2′-OH to show its distance to the tyrosine. Metal and water ions are removed for clarity. A light green circle is added to the scissile phosphate group to highlight its location as the target of the reaction. The deoxyribonucleotide 5′ of the target ribonucleotide is displayed to emphasize base stacking and present depth behind the scissile

phosphate group (light green circle). The DNA ribbon highlights purine bases in red and pyrimidine in blue. This model was created using Discovery studio suite using the data gathered from Chon H., *et al*. [17, 19]

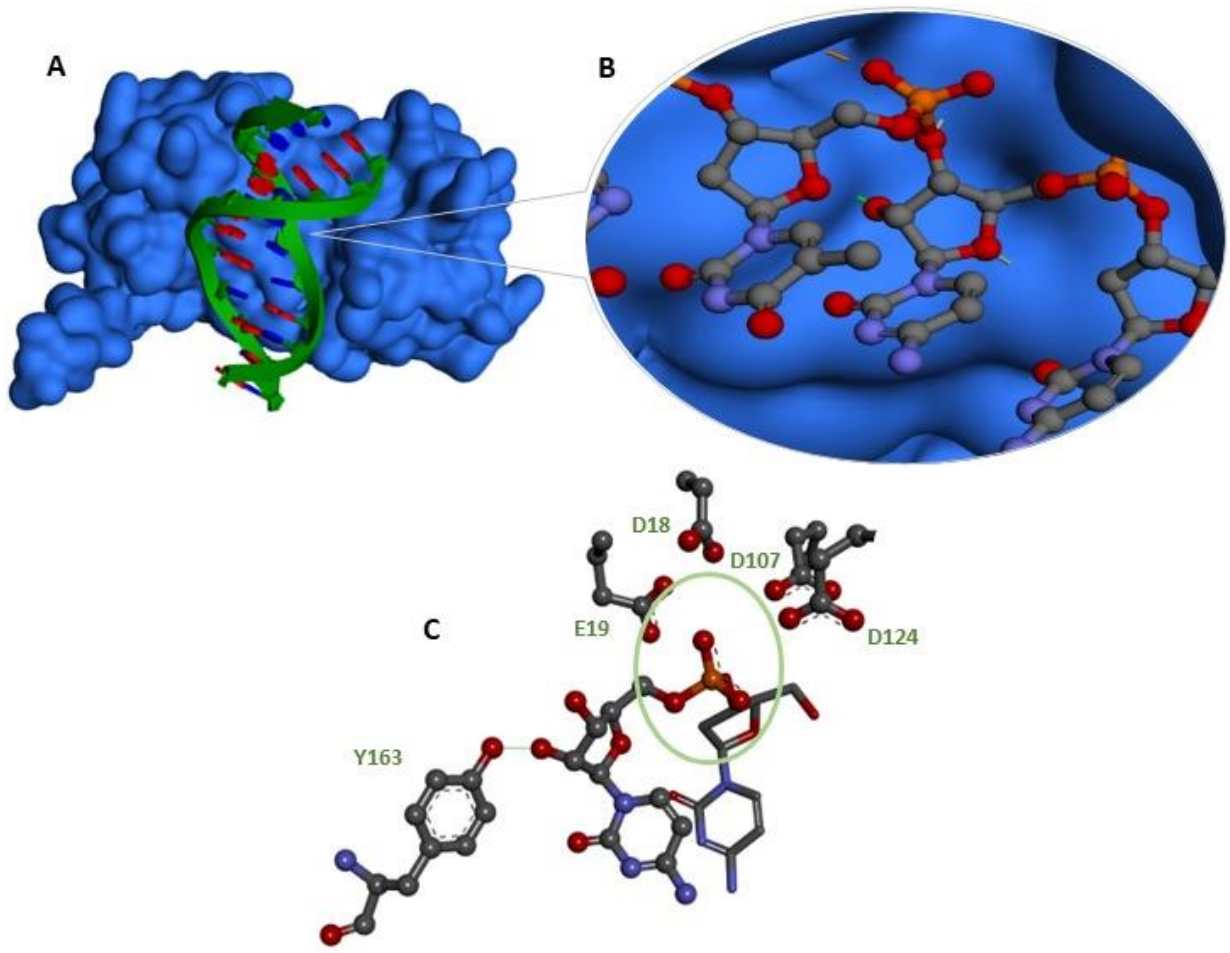


Figure 1: Representation of the binding of RNase H2. (A) The catalytic domain of RNase H2 binding with the minor groove of ribo adulterated DNA. (B) Zoom of the binding pocket at the ribose sugar. Green dashes mark the closest distances to the enzyme pocket residues. (C) The ribose sugar surrounded by the sidechains of the active site residues (D18, E19, D107, D124). Green oval denotes where cleavage takes place. Y163 ensure recognition of the hydroxyl on the ribose sugar before cleavage can begin.

3.2 Results

3.2.1 Ribo G adulterated DNA

Previous work has shown that the flanking sequences affect the local structure and dynamics of a ribo g lesion. This suggests that the flanking sequence may also modulate the recognition and processing by NER enzymes such as RNase H2/II. I have investigated all possible nearest neighbor combinations of ribo g damaged DNA duplexes. The purity of all oligonucleotides was established by denaturing gel electrophoresis prior to testing. Figure 2 depicts the nearest neighbor combinations for the ribo g duplexes that were investigated. AgC and CgA have the ribo g placed 8 bases from the 5′ end, while all the rest place the ribo sugar 7 bases from 5′. This causes a 7-mer cleavage product of AgC and CgA and a 6-mer product of the remaining oligonucleotides. RNase HII was shown to have potential activity on some single stranded oligonucleotides. These single stranded substrates were investigated in detail.

Ribo G duplexes			
AgC	5′-C-G-A-T-G-G-A-rg-C-T-C-C-G-3′ 3′-G-C-T-A-C-C-T- C -G-A-G-G-C-5′	**TgC**	5′-C-G-A-T-C-T-rg-C-T-A-G-C-G-3′ 3′-G-C-T-A-G-A- C -G-T-A-C-G-C-5′
CgA	5′-C-G-A-T-G-G-C-rg-A-T-C-C-G-3′ 3′-G-C-T-A-C-C-G- C -T-A-G-G-C-5′	**GgA**	5′-C-G-A-T-C-G-rg-A-T-A-G-C-G-3′ 3′-G-C-T-A-G-C- C -T-A-T-C-G-C-5′
CgG	5′-C-G-A-T-C-C-rg-G-T-A-G-C-G-3′ 3′-G-C-T-A-G-G- C -C-A-T-C-G-C-5′	**AgG**	5′-C-G-A-T-C-A-rg-G-T-A-G-C-G-3′ 3′-G-C-T-A-G-T- C -C-A-T-C-G-C-5′
GgC	5′-C-G-T-T-A-G-rg-C-C-T-G-C-G-3′ 3′-G-C-A-A-T-C- C -G-G-A-C-G-C-5′	**AgT**	5′-C-G-A-T-C-A-rg-T-T-A-G-C-G-3′ 3′-G-C-T-A-G-T- C -A-A-T-C-G-C-5′
AgA	5′-C-G-A-T-C-A-rg-A-T-A-G-C-G-3′ 3′-G-C-T-A-G-T- C -T-A-T-C-G-C-5′	**TgA**	5′-C-G-A-T-C-T-rg-A-T-A-G-C-G-3′ 3′-G-C-T-A-G-A- C -T-A-T-C-G-C-5′
TgT	5′-C-G-A-T-C-T-rg-T-T-A-G-C-G-3′ 3′-G-C-T-A-G-A- C -A-A-T-C-G-C-5′	**CgT**	5′-C-G-A-T-C-C-rg-T-T-A-G-C-G-3′ 3′-G-C-T-A-G-G- C -A-A-T-C-G-C-5′
GgG	5′-C-G-A-T-C-G-rg-G-T-A-G-C-G-3′ 3′-G-C-T-A-G-C- C -C-A-T-C-G-C-5′	**GgT**	5′-C-G-A-T-C-G-rg-T-T-A-G-C-G-3′ 3′-G-C-T-A-G-C- C -A-A-T-C-G-C-5′
CgC	5′-C-G-A-T-C-C-rg-C-T-A-G-C-G-3′ 3′-G-C-T-A-G-G- C -G-A-T-C-G-C-5′	**TgG**	5′-C-G-A-T-C-T-rg-G-T-A-G-C-G-3′ 3′-G-C-T-A-G-A- C -C-A-T-C-G-C-5′

Figure 2: Oligonucleotide duplex substrates.

Figure 3 is a cartoon representation of the procedural flow of data gathering for this experiment. Radiolabeled oligonucleotides containing each of the four bases 5′ and 3′ of a single ribonucleotide inclusion are incubated with human RNase II and then analyzed using denaturing PAGE. The gel depicted is a theoretical representation showing the spacing and allocation for control samples. Though the gel at the time of running contains all three strands of DNA only the labeled 5′ phosphorus group will be visible on the light reactive film. Finally, reactivity is calculated by measuring the black pixel density of the bands and comparing the two values. Computer assisted densitometry is utilized to calculate the relative percent reacted. The graph represents the number of black pixels (independent) per position (dependent) on the film. The pixel positioning is determined by reception of the beta particles on the reactive Fujifilm and by the scanner resolution in pixels per inch (PPI). Images were captured at 25 μm per pixel resolution. B) example showing the black pixel bands and their placement. Vertical dotted lines are placed cutting off the dark band from the background of the scanned image and corresponding to the vertical dotted lines under the peaks in the above graph. The band table outputs the count data for each band separately. Both programs ImageQuant and ImageJ were used to calculate the density of black pixels in the band area.

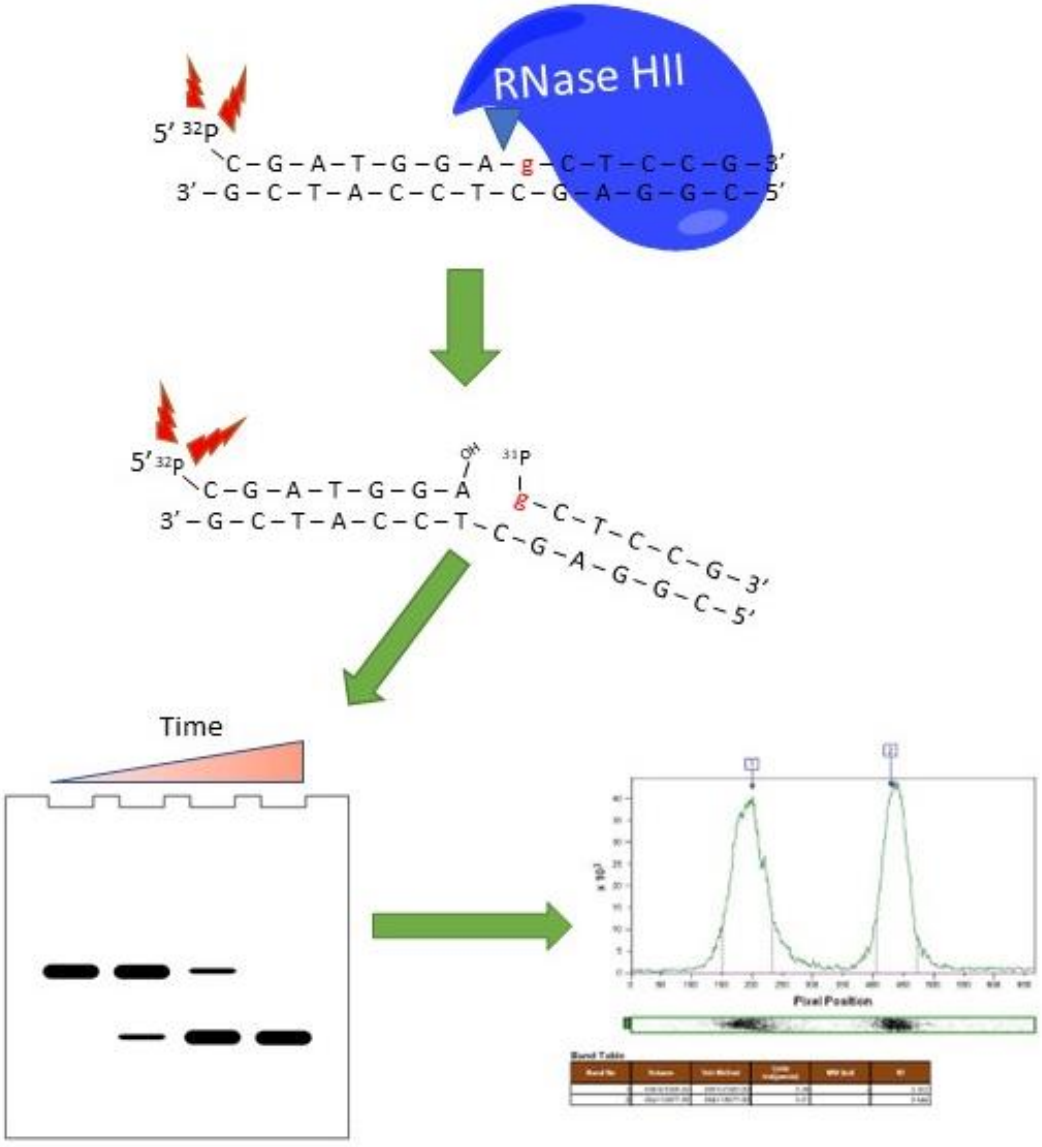


Figure 3: Schematic diagram of the enzymatic assay. ^{32}P radiolabeled duplex substrates are reacted with RNase HII (top and middle) and analyzed by denaturing gel electrophoresis (bottom).

3.2.2 Double Stranded ribo g reactions

Figure 4 investigates reactions that maintain the adenine 5′ of the ribo guanine and vary the bases 3′ side. For technical reasons the reactions were run at 25 °C to slow them down and to ensure stability of the formed duplexes. Demonstration of the sequence environment modulating

the enzymatic activity of RNase HII. The lanes of the gel are labeled from left to right across the top row with their corresponding sample label (1st row) and corresponding timepoint in hours (2nd row). The AgC duplex contains a 0hr time point that depicts proper radiolabeling and sample integrity. AgC reacted to 40% completion after just one hour. This reaction is reliably reused as a convenient control to monitor the reactivity of the other duplexes. After cleavage, the radiolabeled portion of the AgC oligonucleotide is 7 residues. This is one residue longer than the other AgX samples causing it to run slightly higher in the gel. The AgG and AgA oligonucleotides yield a reduction in reactivity to only 8% after 1 hour. While AgT (rightmost) appeared to have half reactivity yielding just 21% after 1 hour. This is a clear demonstration of how the sequence environment greatly modulates the enzymatic activity of RNase HII. AgC (leftmost) is the only sample with which we took a 0hr time point to show all proper radiolabeling and the sample is clean with no previous reaction. This sample is reliable and is used as a positive control based on consistent observations.

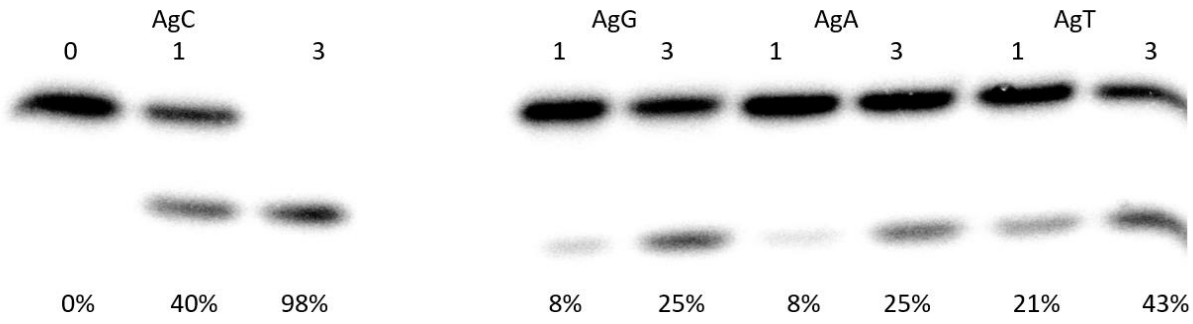


Figure 4: Gel electrophoresis of the 5′-AgX-3′ reactions.

Figure 5 is the next gel in the series, 5′-CgX-3′. AgC is used as a positive reaction control showing 42% cleaved after 1 hour, consistent with our previous result. The gels are loaded with the same amount of reacted DNA which causes the wells to form a more spherical shape as more

lanes (narrower wells) are added to include more oligonucleotides. CgC drops in reactivity to only 8% after 1 hour, similar to AgA and AgG. The 5′ C and 3′ purine samples cleavage drop even further to 4 and 2% respectively. CgA (like AgC) is a longer oligonucleotide after reaction with RNase H therefore it runs higher. CgT is the most reactive in the group cuts to 10% after one hour, on par with the other pyrimidine neighboring oligonucleotides.

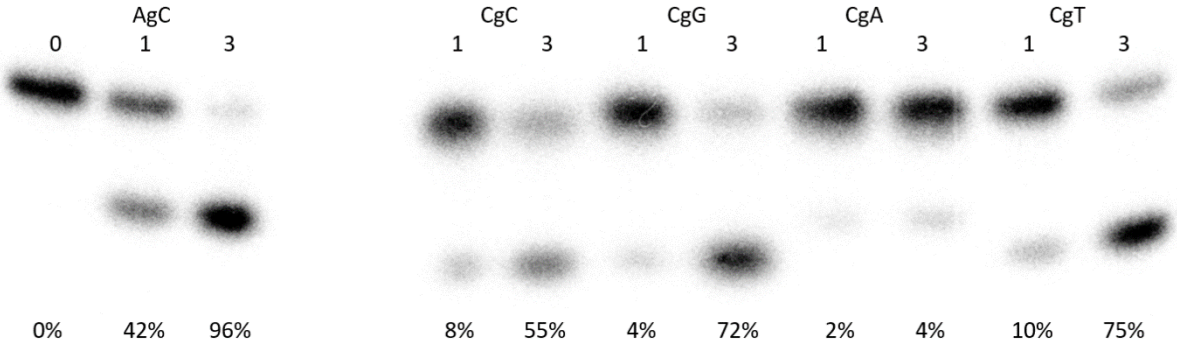


Figure 5: Gel electrophoresis of the 5'-CgX-3' reactions.

Figure 6 is a gel of the 5′-GgX-3′ reactions continuing the series. AgC reacted predictably to 41% complete after one hour, which is on par with previous controls. GgC reliably reacts at 34% near the 40% mark similar to AgC. GgT shows only a slight reduction to the control, while GgG and GgA are more greatly reduced down to 24% and 15%, respectively.

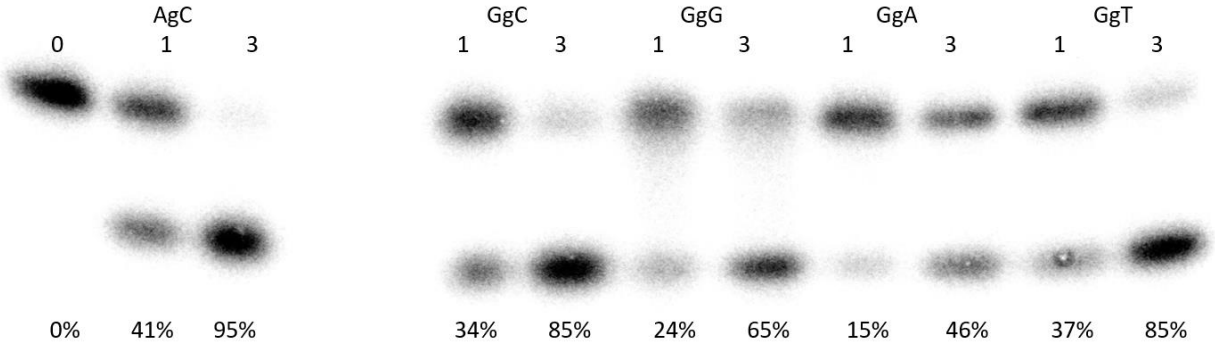


Figure 6: Gel electrophoresis of the 5'-GgX-3' reactions.

Figure 7 is the final gel of the series 5′-TgX-3′, finishing the 4X4 matrix of all possible neighboring base combinations. The AgC sample shows lighter bands because the previous radio labeling of the samples has lost intensity due to the ^{32}P half-life. The 3′ purine offers the greatest reduction in reactivity with TgG at 17% after one hour and TgA at just 9%. Each oligo in this series of reactions shows a reduction in reactivity from the AgC control sample.

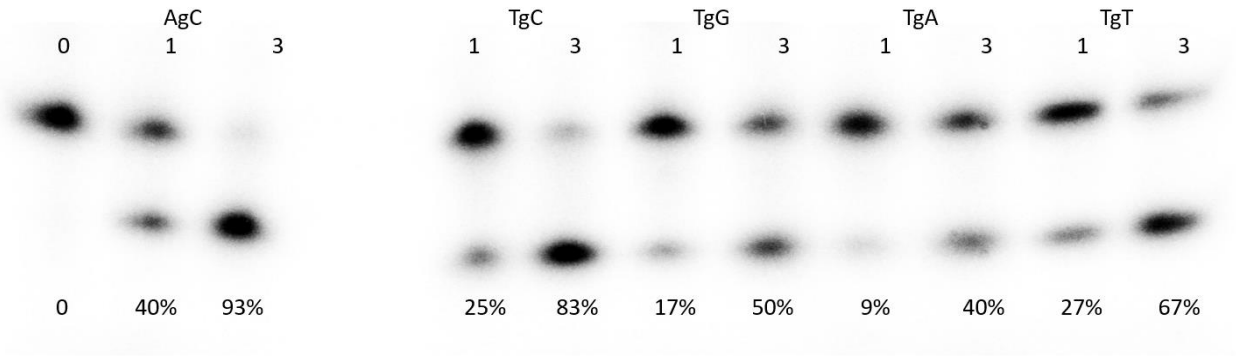


Figure 7: Gel electrophoresis of the 5′-TgX-3′ reactions.

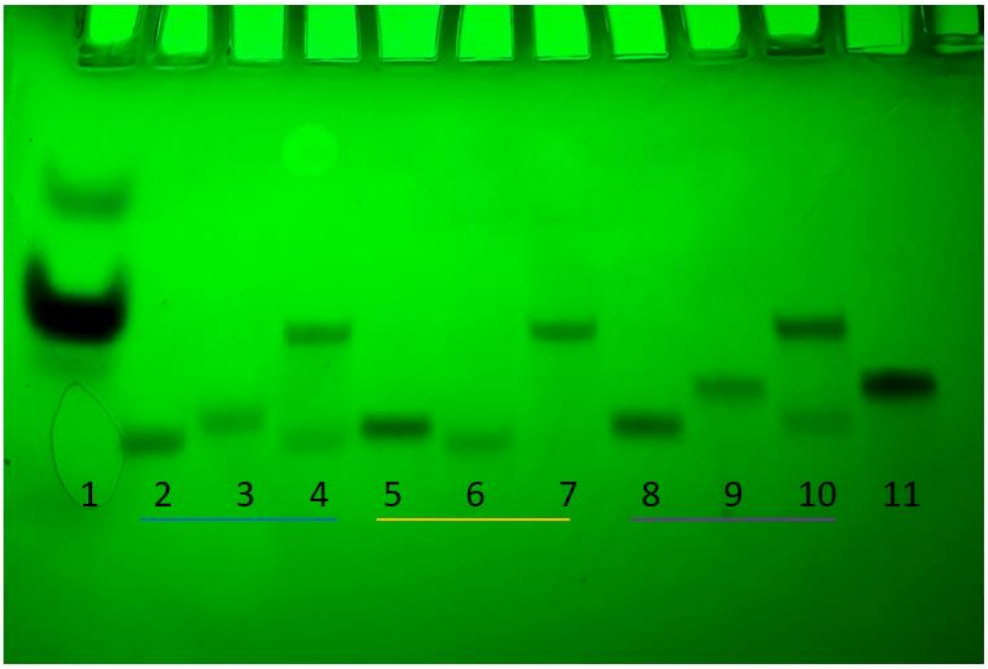
Table 5 is an organization of all the 1-hour at 25 °C timepoint reaction data collected. All of the 3′ purine strands show greatest reduction in reaction efficiency. While the 3′ pyrimidine strands offer the greatest reactivity when coupled with a 5′ purine such as with GgC and GgT. The lowest two reactivities are highlighted red, and the highest two reactivities are highlighted in green.

Table 4: Double stranded oligonucleotide reaction efficiency of the ribo G substrates.

	% Cut after 1 hour			
	XgA	**XgG**	**XgC**	**XgT**
AgX	8	8	40	21
GgX	15	24	34	37
CgX	2	4	8	10
TgX	9	17	25	27

3.2.3 Ribo c oligonucleotides

Selected flanking sequences were also explored for, the most prevalent, ribo c DNA damage to demonstrate modulation of RNase HII activity. Figure 8 is a 15% native acrylamide gel showing duplex formation of ribo c oligonucleotides. Each ribo damaged oligonucleotide set is outlined in color matching boxes below the gel to clarify. Lane 1 contains a simple bromophenol blue marker. Lane 2 is the DNA complementary strand of the ribo damaged CcG oligo (CGG), while lane 3 contains the ribo damaged CcG strand. Lane 4 contains both oligonucleotides with the complement in slight excess to ensure duplex formation. Likewise, lanes 5 and 6 are the complement and ribo damaged strand GcC respectively and lane 7 is their duplex with no excess. Lanes 8, 9, and 10 are the final ribo C samples that were ultimately replaced with an AcA substrate. Finally, Lane 11 is an Eco R1 self-complementing oligonucleotide run as a control. It's noted that while each ribo-incorporated oligonucleotide runs at a different height, their complements and duplex forms do not.



2= 5′-CCATC**CGG**TAGGC-3′
3= 5′-GCCTA**C**c**G**GATGG-3′
4= 5′-CCATC**CGG**TAGGC-3′
3′-GGTAG**G**c**C**ATCCG-5′

5= 5′-GCTTA**GGC**CTGGC-3′
6= 5′-GCCAG**G**c**C**TAAGC-3′
7= 5′-GCTTA**GGC**CTGGC-3′
3′-CGAATC cGGACCG-5′

8= 5′-CCATGG**AGC**TCGC-3′
9= 5′-GCGA**G**c**T**CCATGG-3′
10= 5′-CCATGG**AGC**TCGC-3′
3′-GGTACC**T**c**G**AGCG-5′

Figure 8: Native polyacrylamide gel showing duplexing of ribo c oligonucleotides and their complements. Bands are made visible by UV shadowing.

The goal of the ribo C oligonucleotides was to simply compare their reactivities to the ribo G and not complete every combination. Figure 9 is a gel comparing the reactivity of GcC to that of GgC and the control AgC. A slight reduction in reaction efficiency was observed when compared to GgC, a reliably highly reactive oligonucleotide.

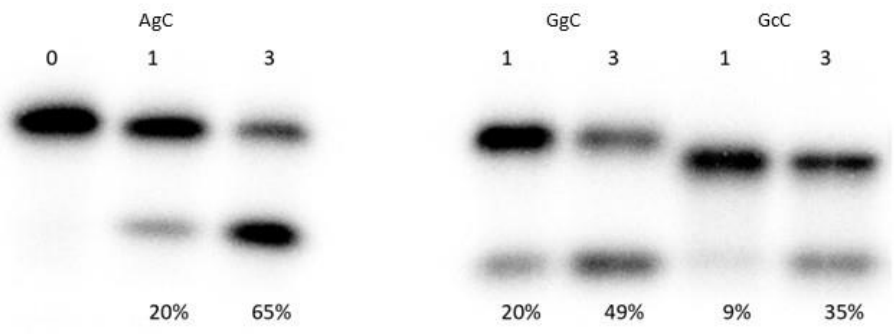


Figure 9: Gel electrophoresis of the GcC substrate reaction.

Figure 10 is the reaction of a 5′ and 3′ purine with AcA. AcA showed to be similarly as reactive as AgA with no apparent reduction and even an increase in reactivity. Figure 11 is the reaction of the inverse of GcC with the CcG oligonucleotide. CcG is much less reactive than GcC. In fact, CcG is least reactive which is expected when compared to the trend observed in table 4 above.

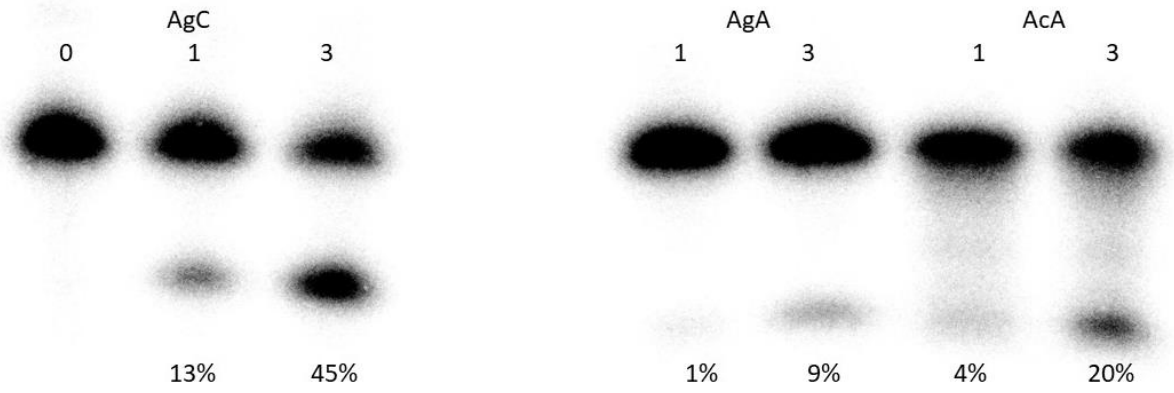


Figure 10: Gel electrophoresis of the AcA substrate reaction.

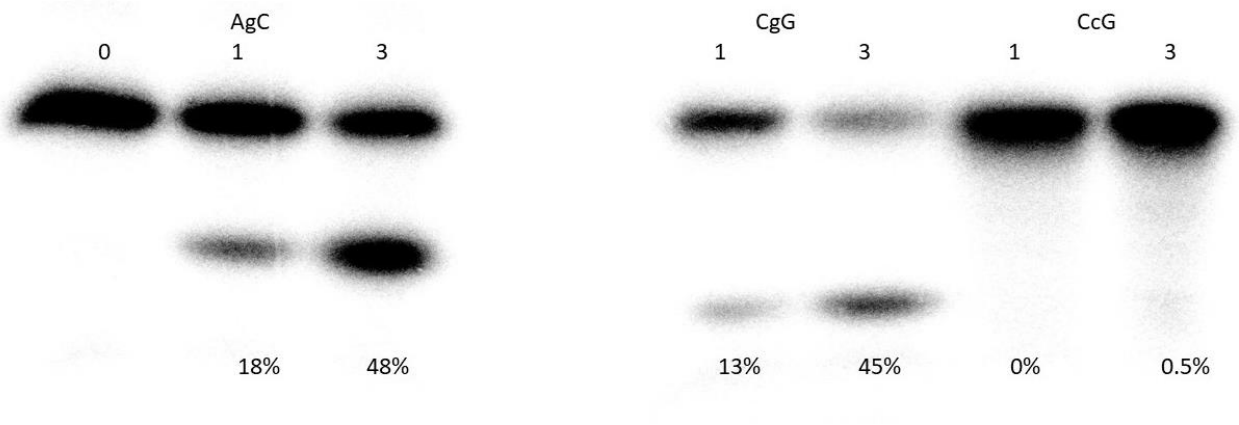


Figure 11: Gel electrophoresis of the CcG substrate reaction.

3.2.4 single stranded ribo g reactions

RNase HII activity on single stranded oligonucleotides was investigated after discovery of potential substrates. Reactions were run to identify which oligonucleotides exhibited substrate characteristics. Figure 12 is one such gel demonstrating single stranded reactivities. Three-hour timepoints were run to ensure reaction completeness. AgC, GgC, and CgG oligonucleotides react as single stranded substrates which appears to be novel. CgA, AgA, and TgT offer no reactivity. The remaining 13 oligonucleotides are also not substrates for RNase HII. Interestingly, as can be seen in figure 13, it was also observed single stranded reactivity for the GcC oligonucleotide though at a much greater reduction. None of the other ribo c oligonucleotides showed any reactivity.

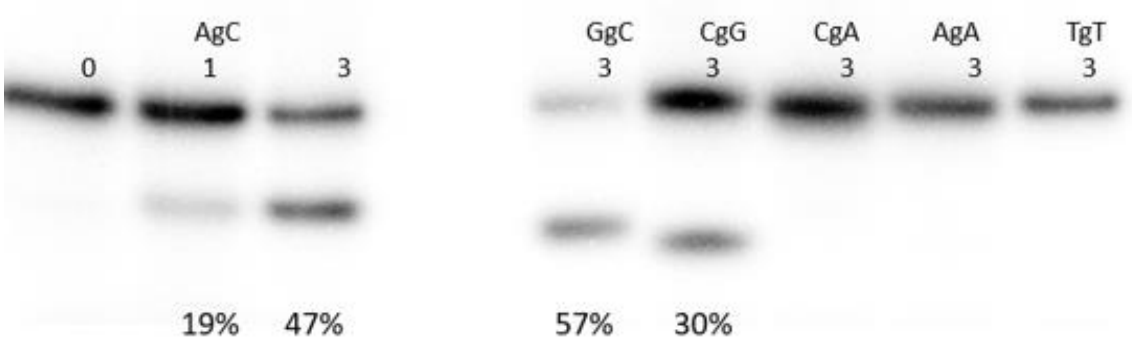


Figure 12: Gel electrophoresis comparing single stranded reactivities.

In order to examine the unexpected reactivity, the most reactive single stranded oligonucleotide, AgC, was investigated in detail. AgC must either be a substrate as a single stranded entity or alternatively it could form a duplex structure. This could happen in either an intra or intermolecular fashion. mFold predicts, figure 14, a semi-stable hairpin structure where the ribo g residue would be in a hairpin loop. This would not appear to be a good substrate for RNase HII. Alternatively, these reactive sequences could form homodimers.

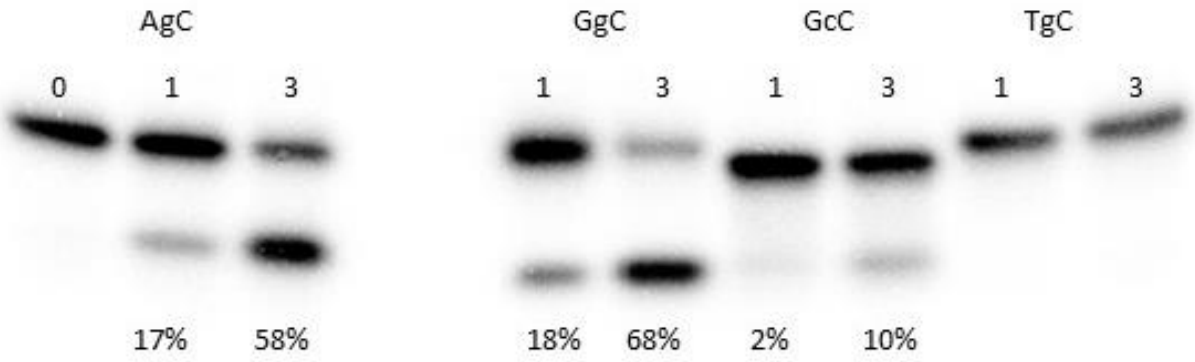
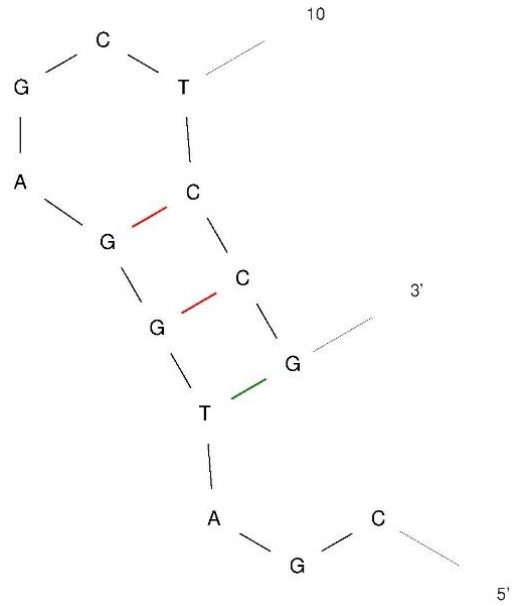


Figure 13: Gel electrophoresis comparing the reactivity of a single stranded ribo C oligonucleotide.



Thermodynamics of Folding	
ΔG	0.49 kcal/mol at 37 °C
T_m	27.1 °C

Figure 14: Predicted mFold secondary structure of AgC oligonucleotide and thermodynamics table. Calculated at 10 µM concentration with 50 mM Na^+ and 10 mM Mg^{++}.

Utilizing the UNAFold Web Services homodimer simulations were run on the three reactive ribo g oligonucleotides. As it turns out, AgC is able to form a stable homodimer with ten base pairs. Figure 16 is a representation of the possible base pairing in the homodimerization of AgC. This is supported by DINAMelt which indicates significant stability for AgC, GgC, and CgG homoduplexes with AgC being the most stable, Table 6.

Figure 15: Example homodimerization of the AgC oligonucleotide.

To check for homodimerization a native gel was run on single stranded oligonucleotides. This demonstrated unusually slow mobility of the AgC single strand which supports homodimer formation for the most stable homodimer (Table 5). Figure 16 is the native UV shadowing gel in which AgC migrates much more slowly representative of dimer formation. Bromophenol blue, ‘BB’, was used as a visual tracker and molecular weight marker as it reliably migrates at the approximate size of a 13 – 15 nt oligonucleotide in a 15% acrylamide gel.

Table 5: Homodimer stability predicted by web server ‘DINAMelt’.

Homodimer stability prediction					
AgC		**GgC**		**CgG**	
T_m	41.4 °C	T_m	30.8 °C	T_m	34.1 °C
ΔG	5.5 kcal/mol	ΔG	3.8 kcal/mol	ΔG	3.3 kcal/mol
Conditions					
[M]	.01 mM	[Na^+]	50 mM	[Mg^{++}]	10 mM

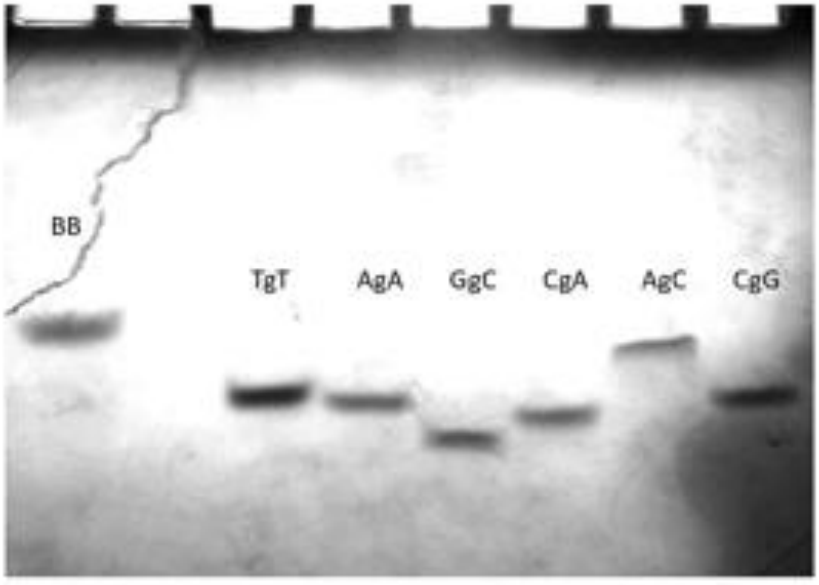


Figure 16: Native UV shadowing gel electrophoresis of single stranded oligonucleotides.

This suggests that if the reactions were run at higher temperatures, we may expect a reduction for the least stable homodimer forms even though the enzymatic activity would be higher at 37 °C. Indeed, that is what we observe in figure 15, as CgG is no longer a substrate and the GgC reaction is much lower.

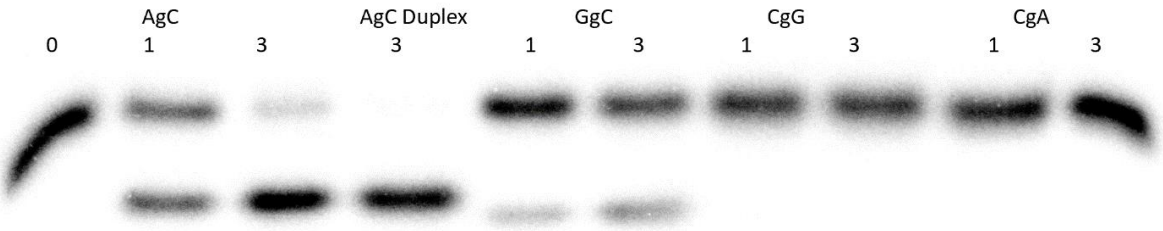


Figure 17: Gel electrophoresis of Ribo G single stranded samples reacted at 37 °C.

As a final test to examine the substrate quality of the AgC sequence element, T-tailed substrates were designed. These substrates cannot form a hairpin or homodimer structure. In addition, an A tailed complement was all designed for the TAgCT sequence to have a double stranded control. None of these single stranded sequences were substrates for the RNase HII while the duplex structure formed from TAgCT, and its A-tailed complement was reactive.

Table 6: Sequences of T tail oligonucleotides

	Sequence													Label
5′	T	T	T	T	T	A	rg	C	T	T	T	T	T	3′	TAgCT
5′	T	T	T	T	T	C	rg	G	T	T	T	T	T	3′	TCgGT
5′	T	T	T	T	T	G	rg	C	T	T	T	T	T	3′	TGgCT

TAgCT Duplex

5′ – T – T – T – T – T – A – rg – C – T – T – T – T – T – 3′
3′ – A – A – A – A – A – T – C – G – A – A – A – A – A – 5′

Figure 18: T tail AgC duplexed with its A tail complement. T tail oligonucleotides were synthesized to ensure no secondary structure formation.

The complement of AgC tail consists of a poly A tail oligonucleotide. If our hypothesis holds then the DNA should be reactive as both a duplex and single stranded. Figure 19 is the gel of single stranded oligonucleotides AgC and the T-tail variants. As predicted the T-tail

oligonucleotides show no reactivity at room temperature compared to the complementary strand lacking AgC sample.

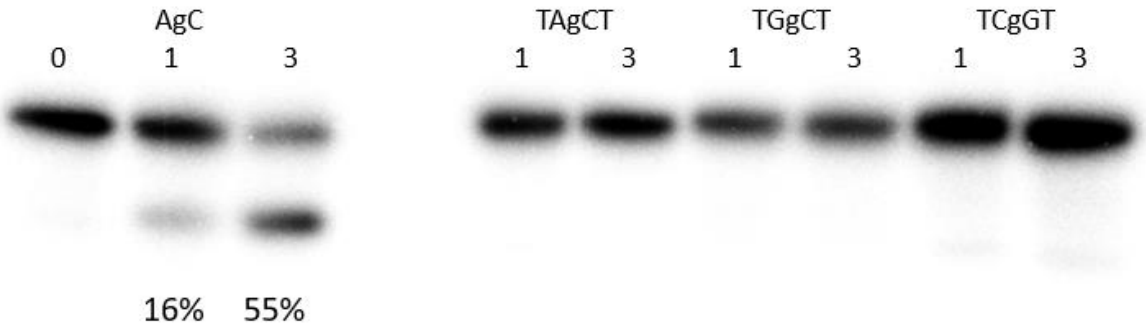


Figure 19: Gel electrophoresis of T tail substrates with an AgC control. Single stranded T tail oligonucleotides show no reactivity.

The T-tail AgC was then reacted at a higher temperature. Figure 20 is the gel of the reactions at 37 °C with minute timepoints taken for more granular data gathering. The 30-minute AgC duplex and AgC single stranded timepoints reacted similarly to our previous 60-minute observations. AgC duplex reacts to completion by the 60-minute timepoint.

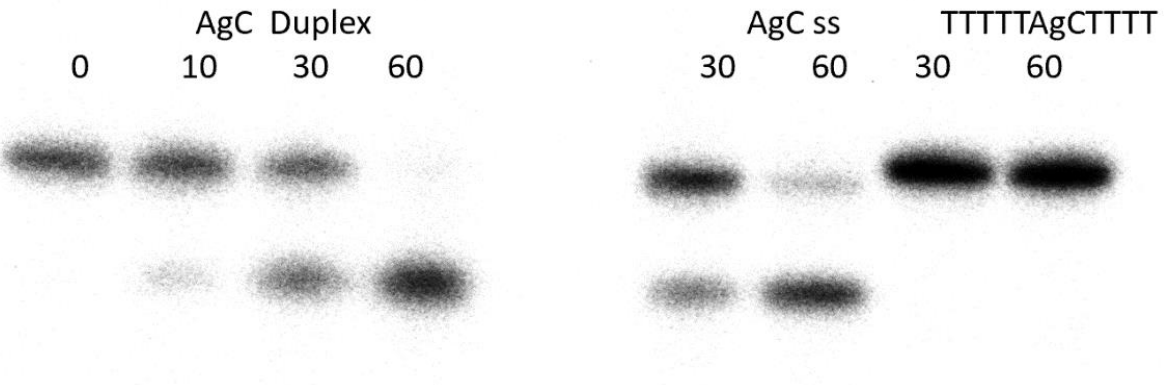


Figure 20: Gel electrophoresis comparing T tail substrate with the AgC single stranded substrate and the AgC duplex control. Reactions carried out at 37 °C to ensure total possible reactivity. Reactivity is completely stopped with the addition of the poly-T tail.

The T-tail AgC was duplexed with its A-tail complement and a single reaction was run to confirm the gain of reactivity once duplex was reestablished. The reaction in figure 21 was run with double the amount of enzyme to ensure complete reactivity and observe slight reactivity, if any exists, in TAgCTss. No reactivity was present. The TAgCT oligonucleotide runs lower after cutting due to the placement of the ribo guanine within the strand. Reactions were run at room temperature.

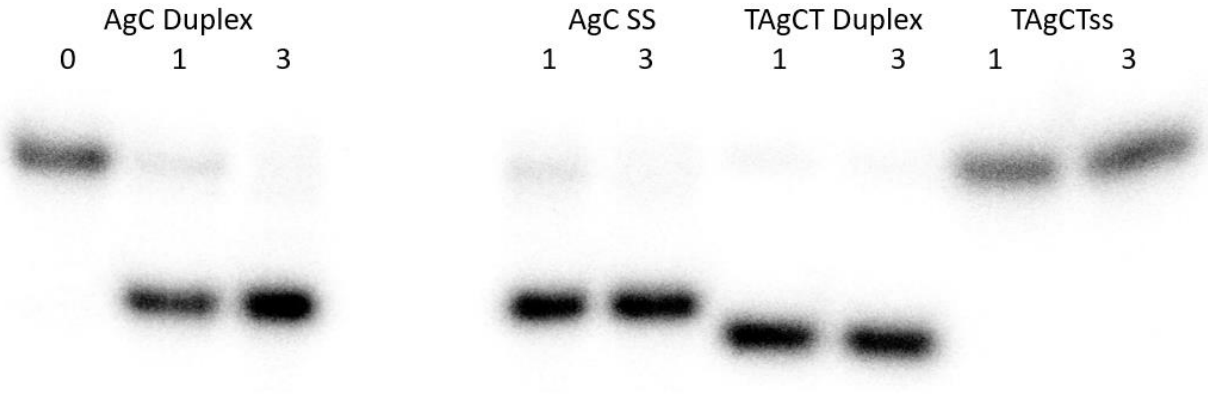


Figure 21: Gel electrophoresis of the T tail AgC substrate with its A tail complement. Once duplexed reactivity is completely restored.

3.3 Discussion: Flanking nucleotide sequences modulate RNase HII activity.

Certain sequences containing ribo-lesion DNA damage are targeted for incision by RNase HII more rapidly than other sequences. AgG and AgA are cut at approximately 20% the rate of the AgC control sample. After organizing the reaction data into a 4x4 matrix a clear trend emerges. Oligonucleotides containing a purine 5′ of the ribo guanine lesion and a pyrimidine 3′ exhibit the highest rate of reactivity. While oligonucleotides con a 5′ pyrimidine and 3′ purine reacted at the lowest rate. With as much as a 20-fold change in reactivity from 2% to 40% occurring by swapping the placement of the adenine and the cytosine in AgC and CgA.

Meanwhile, all of the 5′ and 3′ purine (AgA, AgG, GgG, and GgA) reaction rates were slowed almost equally as much as the 5′ and 3′ pyrimidine (CgC, CgT, TgC, and TgT) reactions. The symmetrical reaction site oligonucleotides (AgA, GgG, CgC, and TgT) reacted to either 8% or 25%. Further research will include the entire gamut of ribonucleotide inserts, the flanking sequences effect, and mutated bases such as inosine. The undertaking herein encompasses all possible nearest neighbor nucleotide combinations for a 13-mer oligonucleotide containing riboguanine. Recent data suggests, however, that the most prevalent form of ribonucleotide incorporation may be in the form of ribocytosine inclusion. [17]

3.3.1 Ribo c vs ribo g

The presence of ribocytosine does not disrupt the sequence environment dependence of RNase HII and the reaction rate is still modulated by the flanking sequences. For the ribo C substrates we investigated the reaction rates were generally lower than that for the ribo G substrates with the same flanking sequences. The only increase in reactivity occurred with the AcA oligonucleotide. Further investigation is necessary but preliminary data suggests the neighboring nucleotide trend is affected by the base of the ribo lesion being targeted by RNase HII. Additionally, we learned that RNase HII does not act on single-stranded substrates. Such that a single ribonucleotide present in a DNA sequence will not be reacted upon unless duplexed. Our AgC oligo reacted without the need of a complement due to some secondary duplex structure being formed. Implementation of the T tail oligonucleotides was useful in shutting down single stranded reactions. This further highlights the substrate specificity for RNase HII.

These data shed more light on the importance of biophysical factors surrounding the repair efficiency of ribonuclease enzymes. This research has the potential to make an impact on laboratory assay development. Reactions catalyzed with RNase H2/II leave a 3′-OH group and a

5′ phosphate group. This characteristic provides some laboratory utility as the 3′-OH ends are extendable under polymerase chain reaction (PCR). Integrated DNA Technology (IDT) has developed a version of PCR known as RNase H-dependent PCR (rhPCR) in specially designed kits to target rare variant sequences. Their well-established rhPCR technique received a patent (US9434988b2) in 2016 and aims to solve a tricky problem of primer dimerization. Employing a primer containing a single ribonucleotide and a cleavage treatment with a thermophilic RNase H2, *Pyrococcus abyssi* (P.a. RNase H2, which is simply added to the PCR master mix) one can reduce primer dimerization and more precisely target / amplify difficult-to-observe sequences. The process leads to the removal of primer-dimer formation by simply adding an additional heated step.

IDT gives an example of P.a. RNase H2 enabling an HCV amplicon PCR product that would not normally appear due to the common occurrence of primer dimer formation blocking amplification. [20] Another use case promoted by IDT is discrimination against closely related genes where the specificity of priming sites is crucial. They demonstrate through a SYBR Green fluorescence assay that improper primer binding due to sequence similarities can be completely removed. We have shown why their reaction efficiency numbers may be further optimized as the reaction efficiency of RNases H are modulated by the neighboring nucleotides causing an up-to 20-fold decrease in cleavage.

Primer construction may take into consideration a 5′ purine 3′ pyrimidine advantage in reaction efficiency. In the paper (Figure 2, B) researchers experiment with the substrate ′S-rC 14-1-15 a 30-mer oligonucleotide with a reactive site containing the sequence 5′-TGcAG-3′, with the lowercase c representing the ribonucleotide insert. Making the single ribonucleotide insert a

guanine and moving it to a location within the sequence that conforms to the 5′ purine 3′ pyrimidine motif would likely see an increase in PCR efficiency.

4 IMPACT OF SINGLE RIBONUCLEOTIDE WITHIN PROXIMITY OF A REPLICATION FORK ANALOG

The structure and flexibility of DNA has been shown to be important for enzyme recognition and these effects were investigated further. In chapter 1 the local sequence environment surrounding a single ribonucleotide inclusion was shown to determine the DNA recognitions and processing efficiency of RNase HII. In this chapter the structural consequences of the presence of a single ribonucleotide in a replication fork analogue are observed using state-of-the-art nuclear magnetic resonance techniques.

4.1 Helicases

DNA helicases are essential enzymes that play a crucial role in replication as the unwinder of duplexed DNA strands. They are present in all prokaryotic and eukaryotic life and are all related to a central progenitor enzyme. This persistence throughout earth's history lends to their importance throughout evolution and is why most human helicases are homologues of prokaryotic helicases. A class 3 hydrolase composed of six subunits, helicase binds and slides along unwinding the duplexed DNA strands into two single strands. Helicases are members of the replisome and are classified as belonging to one of six superfamilies. The superfamily to which a helicase is assigned (I to VI) depends on the presence of specialized domains. [21] There are many different helicases with unique use-cases but the most well studied are those belonging to superfamily II. This superfamily contains the RecQ helicase which is present in both *E. coli* and humans and is responsible for unwinding DNA such that it can be properly repaired.

The breaking of the great many hydrogen bonds and unstacking in duplexed DNA requires significant energy input as it is energetically unfavorable. ATP is hydrolyzed by helicases into ADP + Pi and the energy from this hydrolysis is required for separating the strands. Using this ATP driven "grip and pull" method of moving the unwound leading strand through the central hole of helicase, the binding and fitment is very important.[22] Further, helicases exhibit directionality as they may move in either 3′ to 5′ or 5′ to 3′ depending on their superfamily. RecQ moves 3′ to 5′ similar to other helicases from the superfamily II.

A perturbed DNA backbone caused by a ribo lesion may lead to a reduction in helicase activity. Helicase binds at a replication origin site and the presence of a single ribonucleotide on the leading strand impedes its function. [23] An origin of replication is a specific sequence present in a genome that whose existence signals facilitator proteins to bind when cellular conditions are met for replication. The facilitator proteins once bound initiates helicase binding and activity. Prokaryotic plasmids contain only one origin of replication while Eukaryotic genomes contain many. Origin sites are typically AT base pair rich. Following up on research by Dr. Brosh, *et al* (Figure 25) it has been shown that the proximity of a ribonucleotide perturbation to a replication fork on the leading strand causes a localized distortion that inhibits RECQ1 helicase activity. [24] Helicase unwinding studies on the phenomena show enzymatic activity that is negatively affected by the incorporation of a ribo sugar into the leading strand. [25]

4.1.1 Replication Fork

The site of the ssDNA/dsDNA junction is known as the replication fork, and it is moving from the origin site in the direction of replication. A replication fork may be created during replication or also during DNA repair. The replication fork analog "RF1" samples were created with and without a ribo adulteration 2 nucleotide bases up from the fork. This proximity to the

junction was shown by Dr. Brosh to have the greatest impact on helicase unwinding. The presence of a single oxygen atom is all that prevents the helicase from operating at full efficiency. The data in Figure 22 was provided by Dr. Brosh at the NIH and illustrates the impact of a single oxygen atom on a strand of replicating DNA.

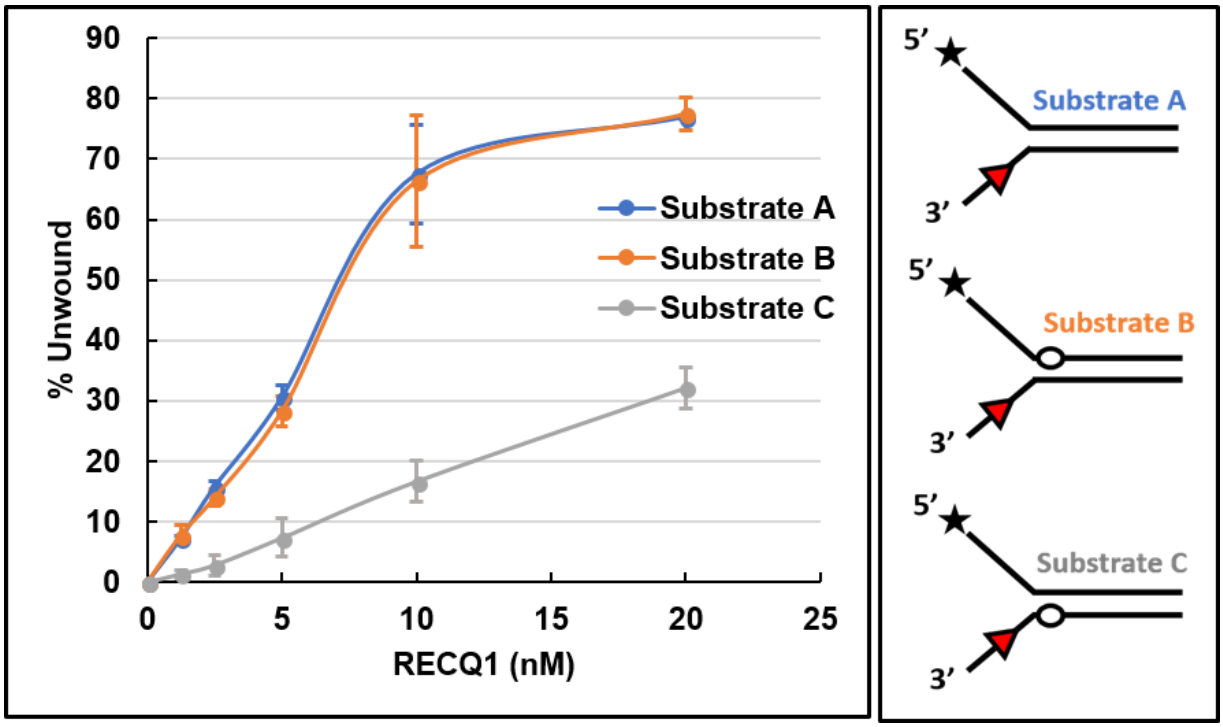


Figure 22: Figure from Dr. Brosh of the unwinding inhibition of RECQ1 caused by the presence of a single ribonucleotide insert. Helicase activity is disturbed when a ribo lesion is in line with the movement of the helicase. The location of the ribonucleotide insert is denoted by the white circle. Substrate A is the control without a ribolesion, substrate B has a ribolesion on the opposite strand, and substrate C has the ribolesion on the strand targeted by RECQ1.

Figure 23 shows the base pairing in our duplexed oligonucleotide. The ssDNA/dsDNA junction occurs at the ninth nucleotide residue (T9). The strands were given numbering priority based on the leading strand having greater molecular weight with the ribo adulteration.

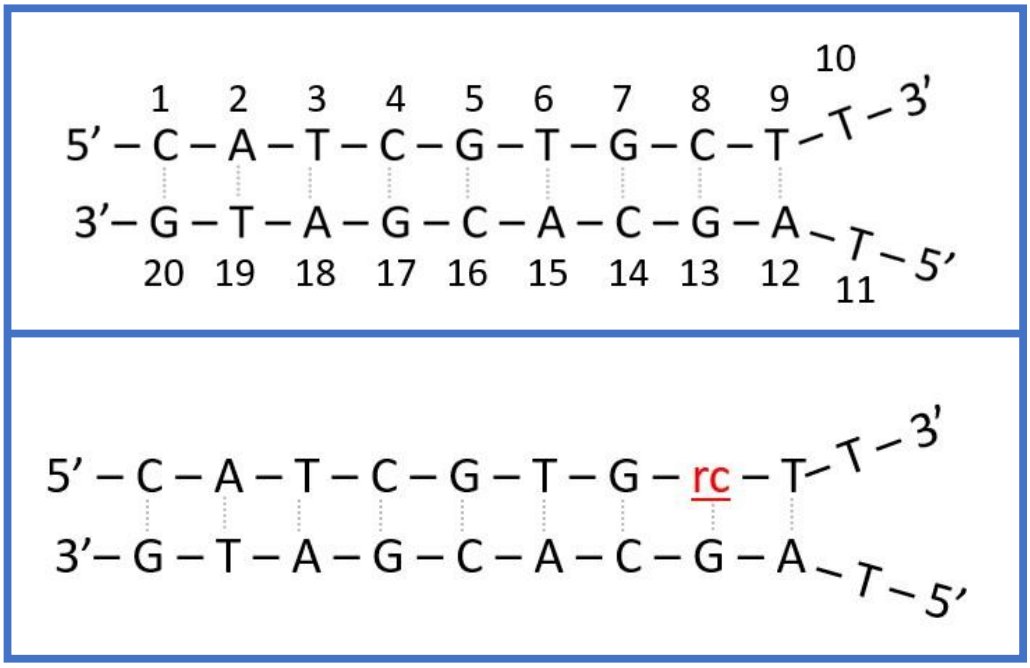


Figure 23: Replication fork analog oligonucleotides with residue numbering.

Figure 24 is a cartoon representation of the numbering convention used to identify the protons present in the sugar ring of our DNA oligonucleotides. It is important during assignment to note that H2″ is not present in the ribocytosine (rC8) containing sample. Distances between H1′ and H2′/H2″ are approximations used for NOESY intensity referencing and assignment.

The sugar ring can have a range of conformations which is defined by the pseudorotation angle. The pseudorotation angle (P) goes from 0 to 360 degrees and is depicted as lying at position on the pseudorotation wheel. Of these conformations 2 dominate: 2′ endo, known as a ‘southern’ conformation (S, P ~170°), and 3′ endo, known as a ‘northern’ conformation (N, P ~10°). Unless the conformation of a ribose or deoxyribose ring is pure N or S then there is an

equilibrium between these two states. For example, if the fraction south (fs) is 0.4 or 40% that signifies an equilibrium of 40% S and 60% N. Figure 25 is the pseudorotation wheel as depicted by Evich, Spring-Connell, and Germann in 2017 and adopted from Altona and Sundaralingam (1972). The wheel highlights the degree regions in which each conformation lies (S, N) with the general shape equilibrium of the sugar ring outside of the wheel in their corresponding regions.

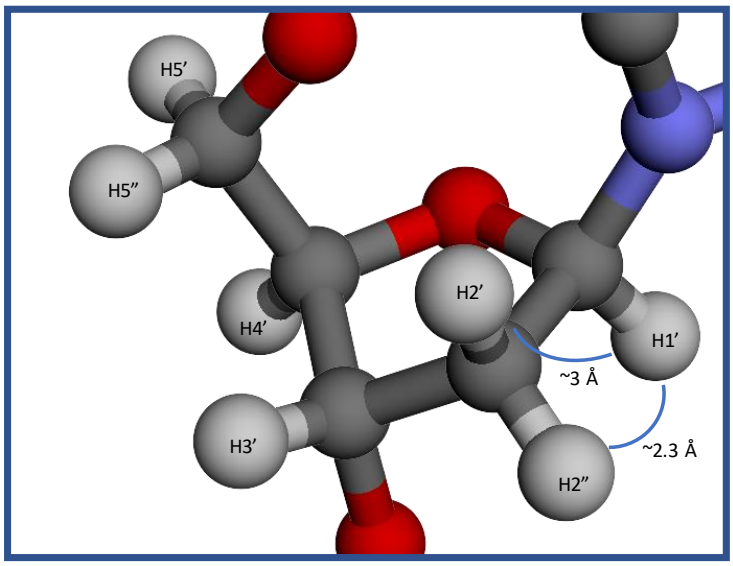


Figure 24: Numbering of deoxyribose sugar ring protons and approximate distances from H1' and H2'/H2'' in a 2' endo (southern) conformation deoxyribose ring.

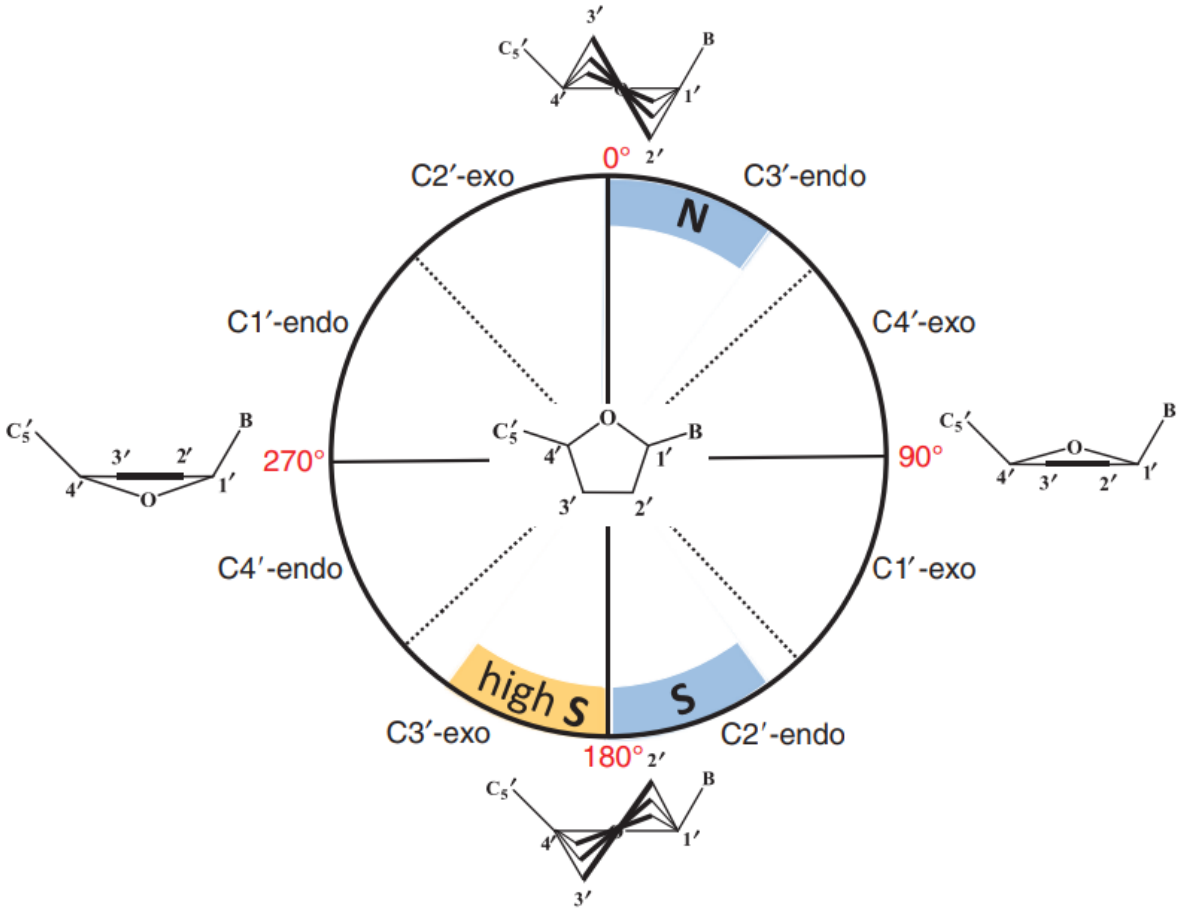


Figure 25: Deoxyribose/Ribose pseudorotation wheel adopted from Altona and Sundaralingam (1972). [26]

4.2 Results

4.2.1 Imino Protons

By identifying the location of each imino proton in the duplex we are able to confirm proper base pairing and a 1:1 molar ratio of our strand and the complement. Nine (9) base pairs are identified though imino proton analysis. Imino protons are especially useful because they are only visible in the spectrum if their respective hydrogen bond is intact. [26] Figure 26 shows a temperature gradient applied to our RF DNA oligonucleotide and its effect on the exchangeable protons. The chemical shifts are more labile for the protons further from the center of our duplex while the center stem protons chemical shifts are less affected by temperature. By 303 K the peaks for the base paired thymine imino protons are too overlapped to be useful in assignment. As the temperature increases exchange causes many of the peaks to become broad.

All five of the guanine imino protons are represented in the lower temperature range spectra. The chemical shift of the G20 imino proton is visible as a shouldered peak at 12.8 ppm on the left side. The transient nature of the end base's imino protons are apparent in how much they shift upon increasing the temperature (G20 and T9 imino protons). By 293 K the overlap is so great it appears as 4 peaks. We determine 288 K to be the 'goldilocks' temperature for imino observation and use that data for assignment. The guanine associated imino protons resonate in the 12.5 to 13.0 ppm range, While the thymine resonance stays in the range of 13.5 to 14.5 ppm. The peak at 14.2 ppm broadens quickly as temperature increases indicating increased exchange.

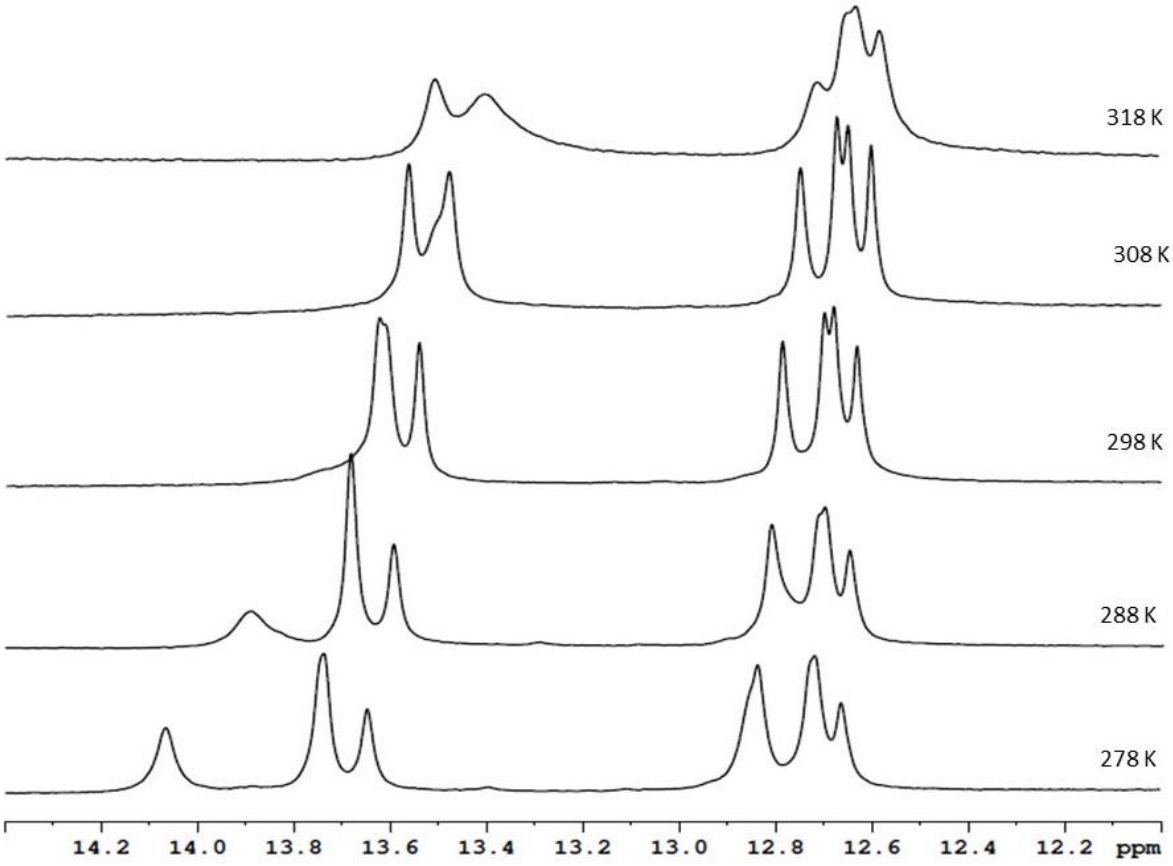


Figure 26: RF DNA imino proton analysis along a temperature gradient on 500 MHz NMR.

Figure 27 is the 1:1 2D NOESY at 283 K ($\boldsymbol{\tau}_M$ = 150 ms, PP: "CHNOESY11" contour levels: 64, contour level increment: 1.2) of the RF DNA oligonucleotide. This spectrum was used for the assignment of imino protons. Three identifiable cross peaks are present. The peak at 14.0 ppm remained broad throughout the temperature gradient. The broadness of this peak indicates it is more likely acceptable to exchange and therefore on the end of the oligonucleotide nearer the mismatch pair. This confirmed our observations in the temperature gradient spectra of figure 26.

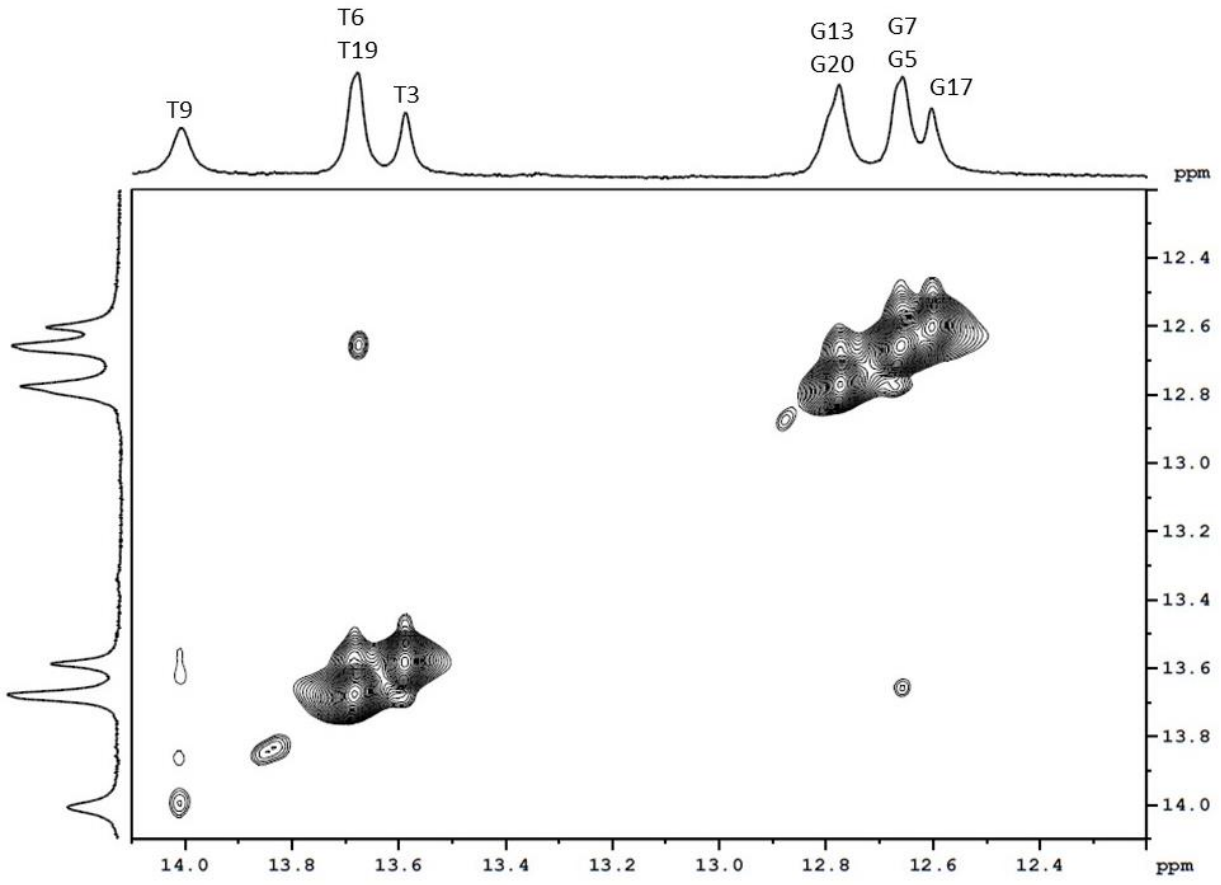


Figure 27: ^{1}H 1:1 NOESY spectrum at 283 K of the RF DNA imino proton region on 500 MHz NMR with 150 ms mixing time.

Figure 28 is AH2 and G-NH_2 region of the above RF DNA NOESY spectrum. NOE development occurs between the thymine imino protons and neighboring guanine amino protons and adenine AH2 protons. These NOE denote contact between the bases of the residues. Each residue is labeled according to its assignment at the top of the figure. The guanine cross peaks that are further upfield appear more weak because they are further from the imino protons than the downfield peaks.

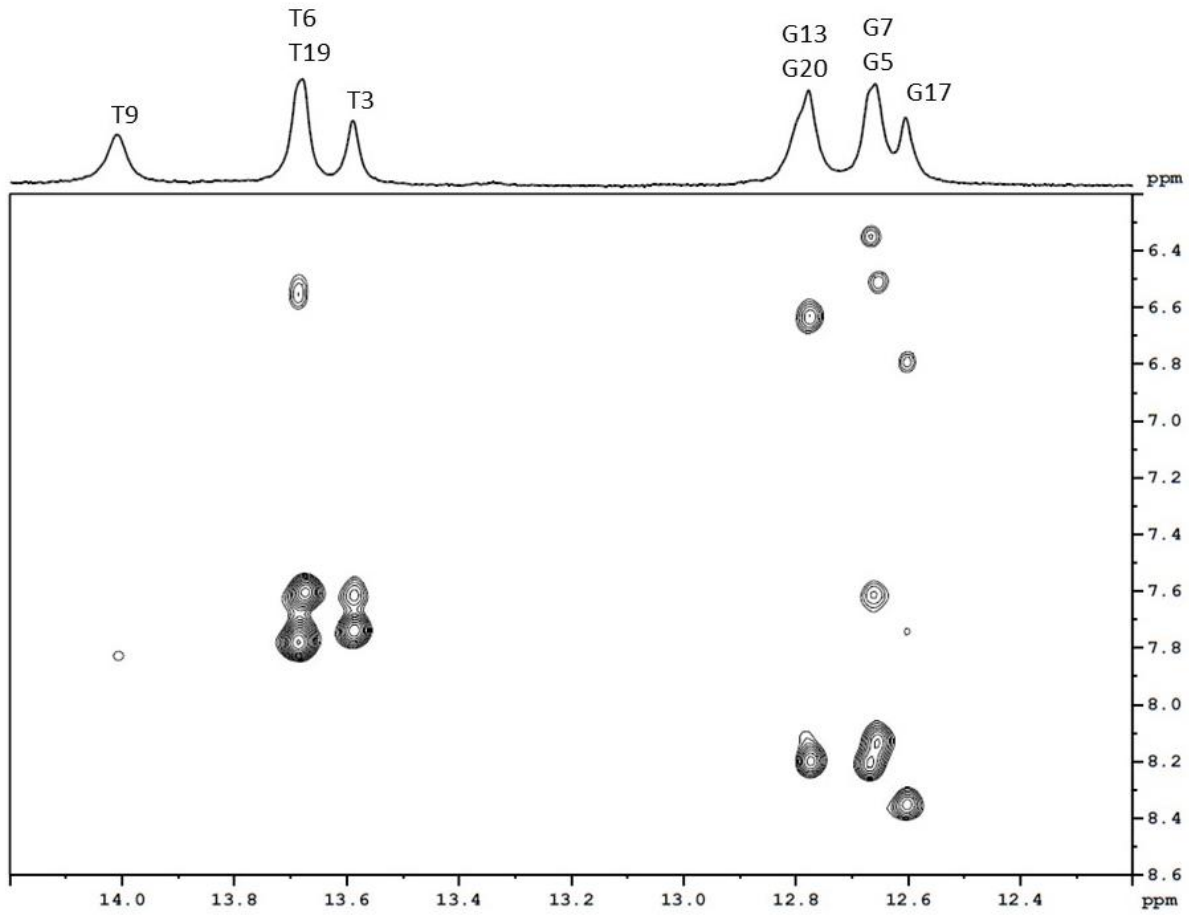


Figure 28: NOESY spectrum at 283 K of the RF DNA imino – AH2/G-NH_2 region.

Figure 29 is the is the temperature gradient spectra of the riboadulterrated replication fork analog. It is immediately clear that there are significant differences between the two temperature gradient spectra despite the only difference between the two samples being a single oxygen atom. Similar to the previous temperature gradient spectra the peaks become too broad and difficult to distinguish after about 303 K.

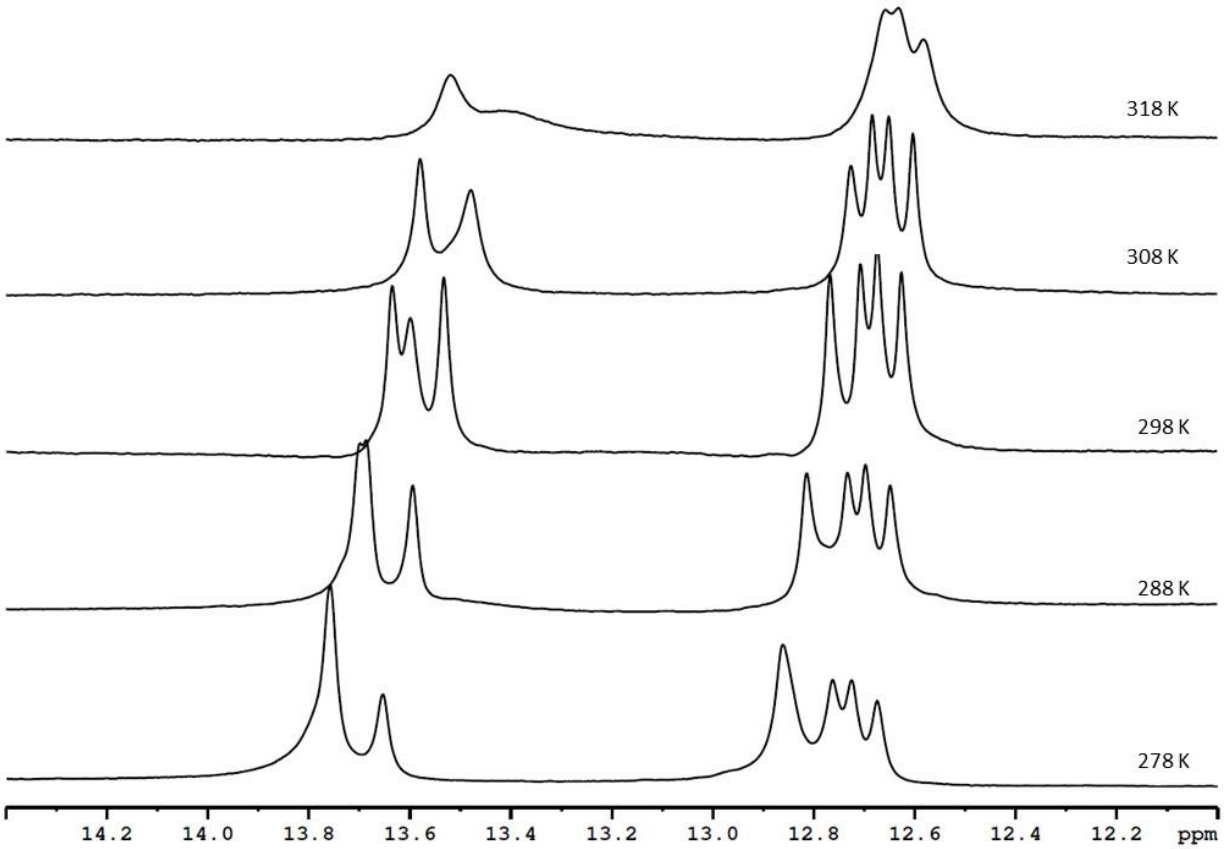


Figure 29: RF RNA imino proton analysis along a temperature gradient on 500 MHz NMR.

Figure 30 is the 2D imino proton NOESY spectrum of the riboadulterrated RF duplex at 283 K ($\boldsymbol{\tau}_M$ = 150 ms, PP: "CHNOESY11" contour levels: 64, contour level increment: 1.2). The spectrum of the 2D NOESY contains crucial differences. Only two cross peaks are visible.

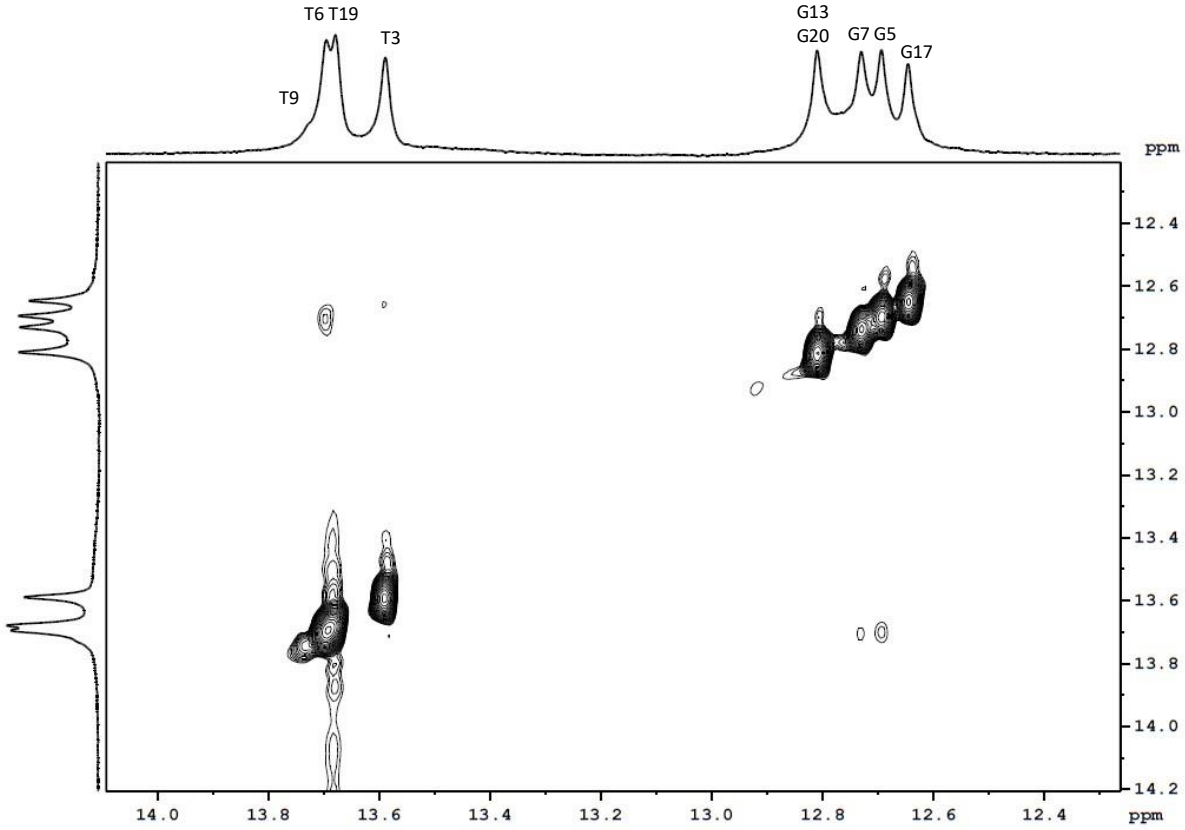


Figure 30: NOESY spectrum at 288 K of the imino proton range in the duplexed RF RNA analog on 500 MHz NMR with 150 ms mixing time.

Figure 31 is the AH2 and G-NH_2 region of the above RF RNA NOESY spectrum. There is an obvious lack of an NOE under the T6 peak while one remains under G7.

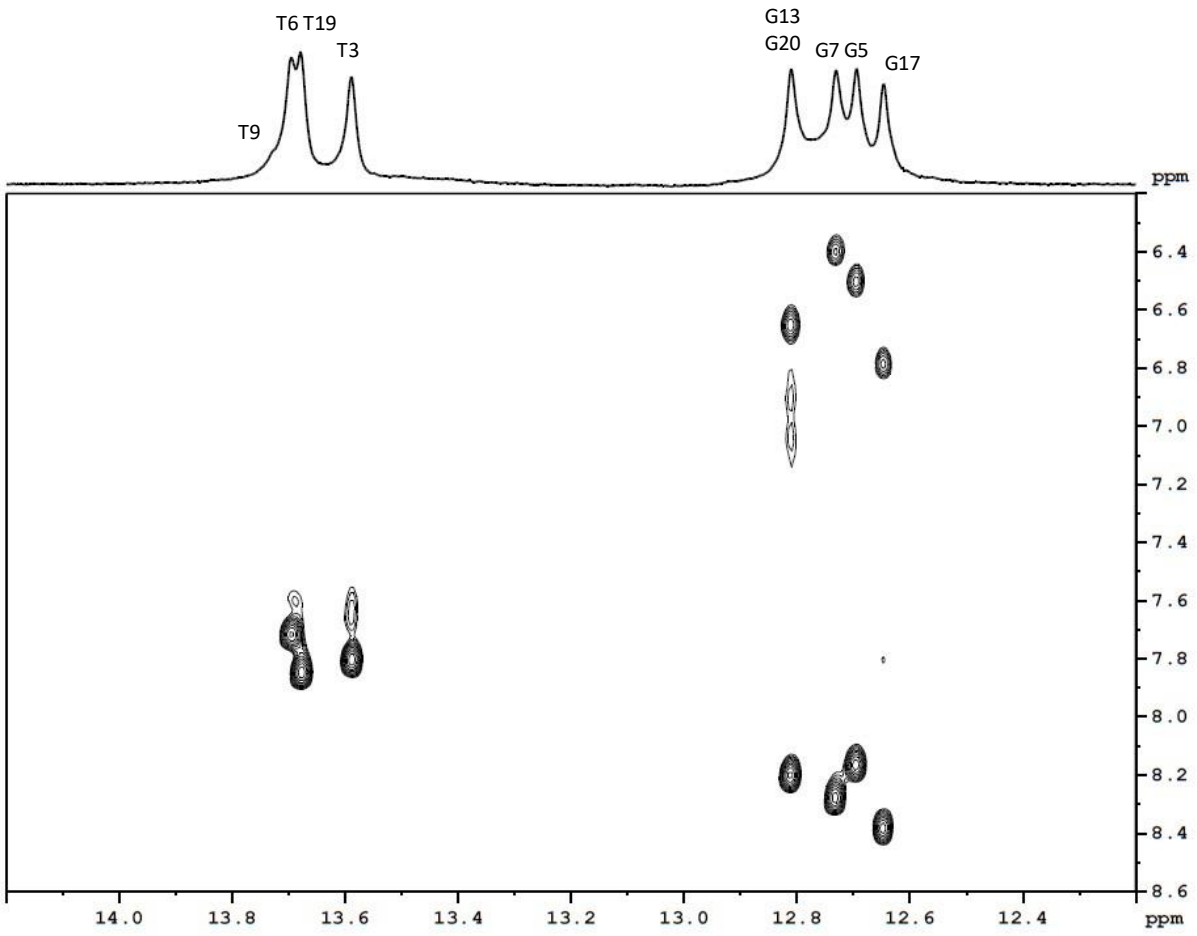


Figure 31: NOESY spectrum at 288 K of the duplex RF RNA analog imino in the AH2/G-NH_2 region.

Figure 32 is an imino proton spectra overlay of the two replication fork analogs. The spectrum labeled 'DNA' is the unadulterated complementary oligonucleotides and the spectrum labeled 'RNA' is the ribonucleotide adulterated sample. Each imino proton has been assigned and labeled on the spectra.

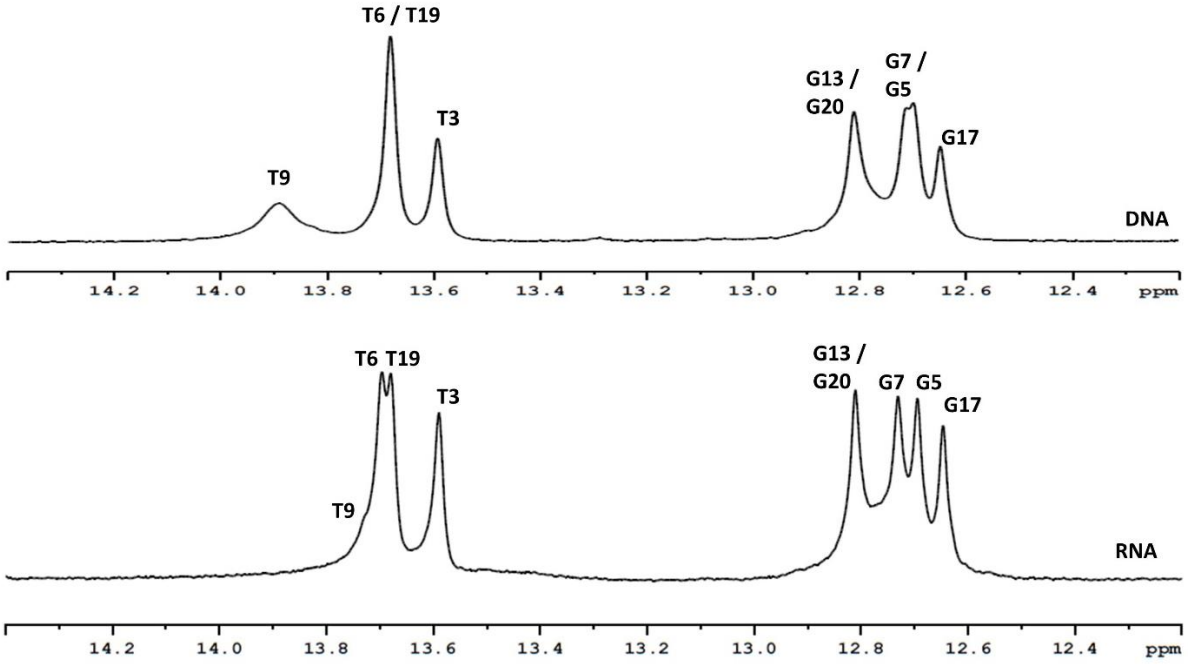


Figure 32: Imino proton spectra at 288 K highlighting disruptions in some hydrogen bonds while stem base pairing remains unchanged.

Figure 33 is a graph of the change in imino proton chemical shift per residue between the DNA control and RNA oligonucleotides. The stem residues of the riboadulterrated replication fork analogue, from C1 to G5, remain consistent at a 0.06 ppm difference from the DNA control. At T6 the chemical shift increases by positive ppm until T9 where the chemical shift changes more up field by -0.1 ppm. T10 has no chemical shift because it does not base pair with its complement, T11. The chemical shift delta for T9 is strongly negative due to its proximity to the non-complementing tail and thus more open to exchange. Table 7 consolidates the data presented in Figure 33, again T10 and T11 imino proton chemical shifts are not present because they do not base pair. The duplex cartoon at the top right of the figure is the residue numbering scheme for reference. Note the ribonucleotide is on the C8 residue.

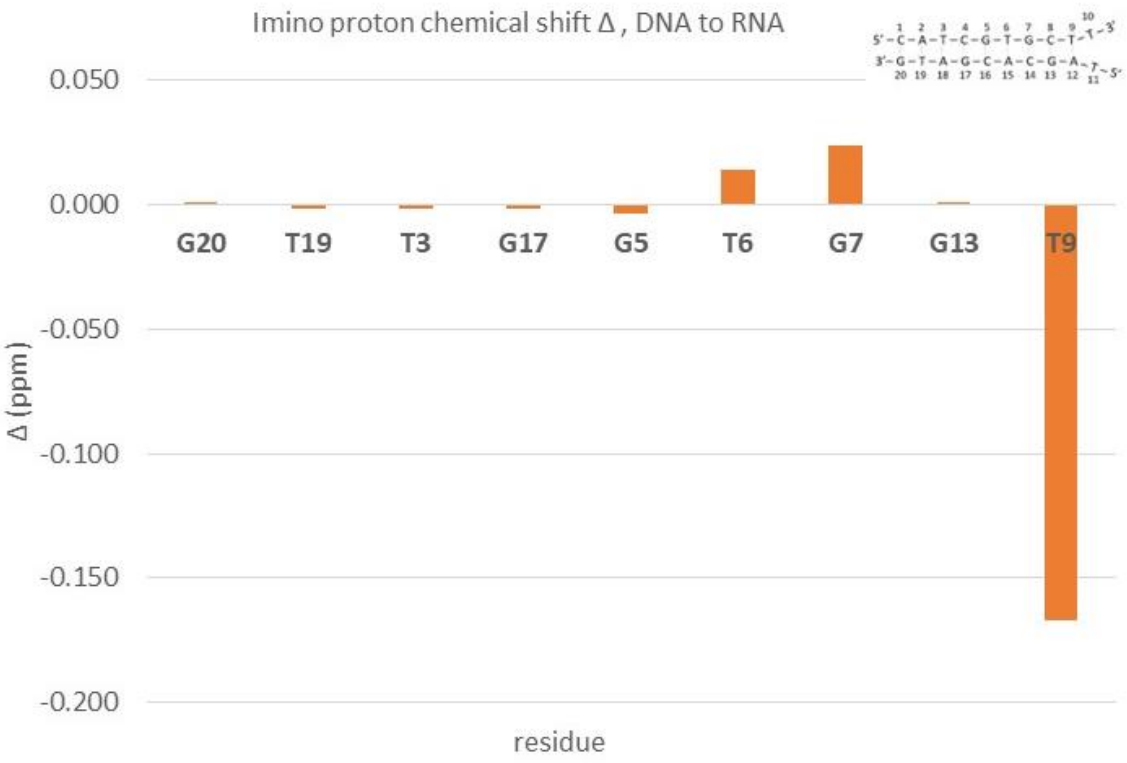


Figure 33: Chemical shift delta of imino protons per residue between RF DNA and RF RNA adulterated oligonucleotides.

Table 7: Comparison of imino proton chemical shifts per residue of RF analog.

residue	DNA imino ^{1}H chemical shift (ppm)	RNA imino ^{1}H chemical shift (ppm)	Δ (ppm)
G20	12.809	12.809	0.000
T19	13.682	13.680	-0.002
T3	13.592	13.590	-0.002
G17	12.649	12.647	-0.002
G5	12.700	12.696	-0.004
T6	13.682	13.696	0.014
G7	12.709	12.733	0.024
G13	12.811	12.812	0.001
T9	13.897	13.730	-0.167

4.2.2 Phosphorous (^{31}P) spectra

Data gathered from running phosphorous NMR experiments helps in identifying backbone integrity. In a similar fashion to the imino proton experiments, a temperature gradient is applied to the first set of phosphorous spectra. Figure 34 is the temperature gradient phosphorous spectra (lb: 4 Hz, ns: 2048) of the RF DNA Duplex. The resolution changes very little with each 10-degree kelvin jump. Figure 35 is the temperature gradient spectra (lb: 4 Hz, ns: 2048) of the RNA containing oligonucleotide. The change of a single peak is evident. The peak at -0.2 ppm is the perturbed backbone of residue C8.

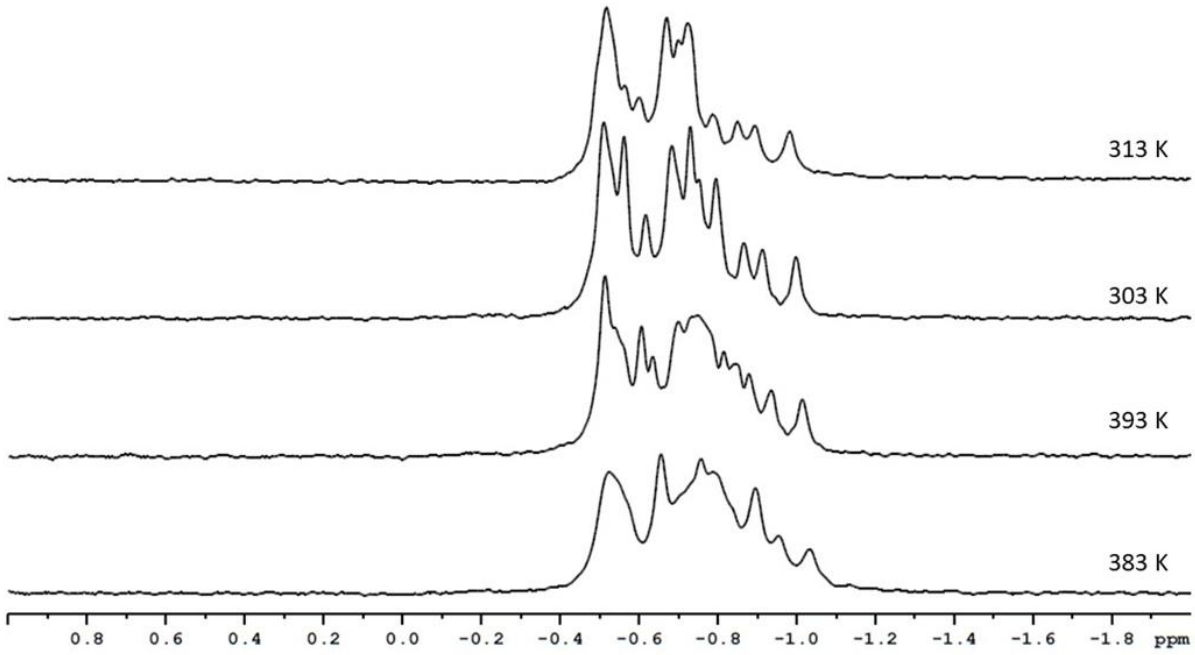


Figure 34: ^{31}P NMR spectra of the duplexed RF analog DNA.

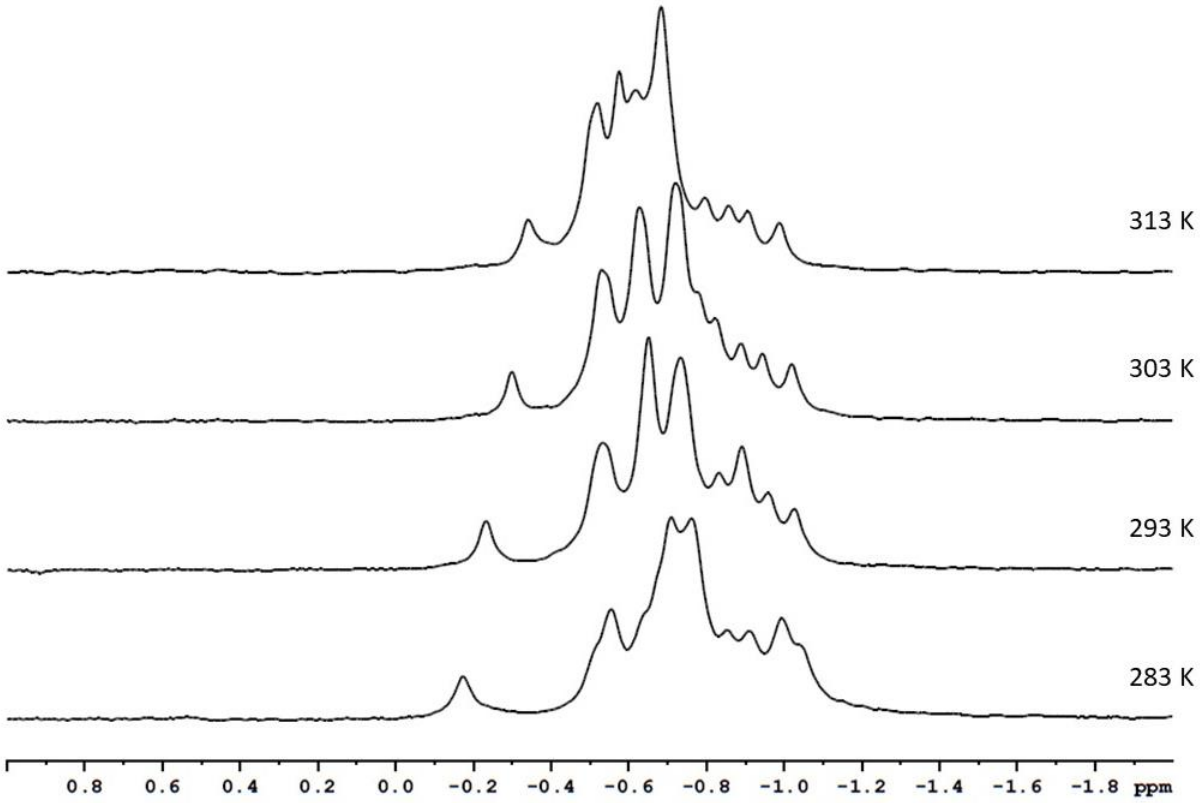


Figure 35: ^{31}P NMR spectra of the RF analog RNA.

Figures 36 and 37 are the residue assignments associated with each peak on the ^{31}P spectrum performed on 600 MHz NMR at 303 K. (ns: 1024, lb: 2.0 Hz)

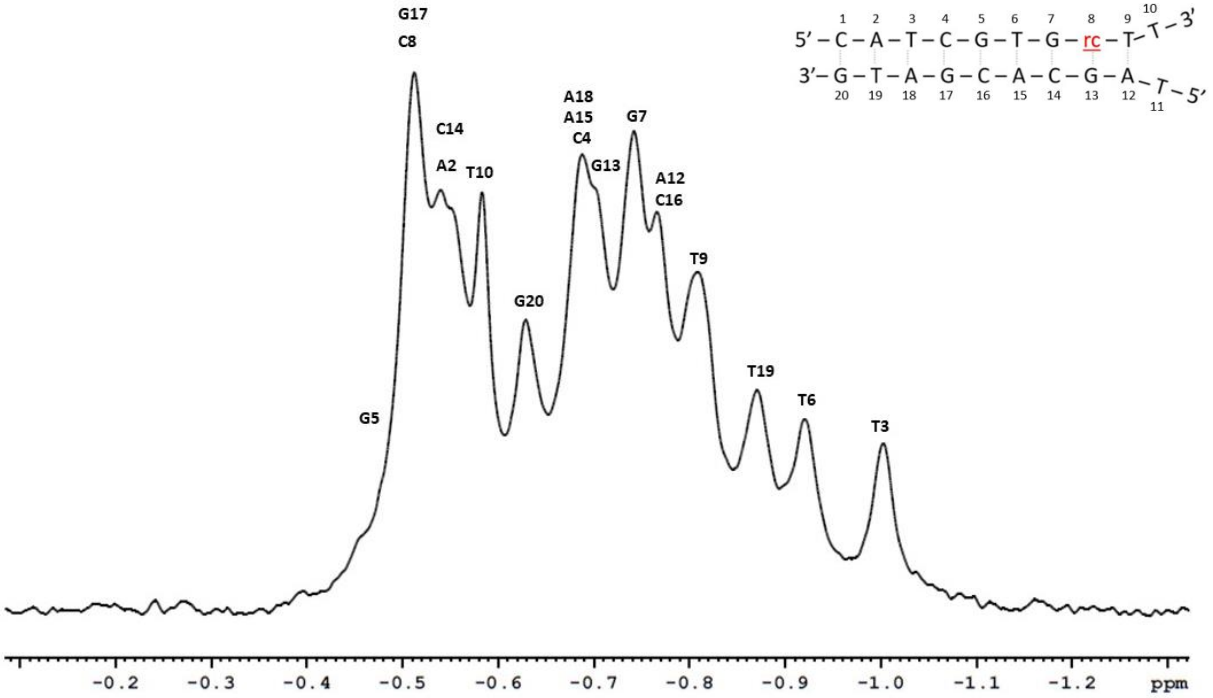


Figure 36: RF DNA ^{31}P spectrum with residue assignment on a 600 MHz NMR at 303 K.

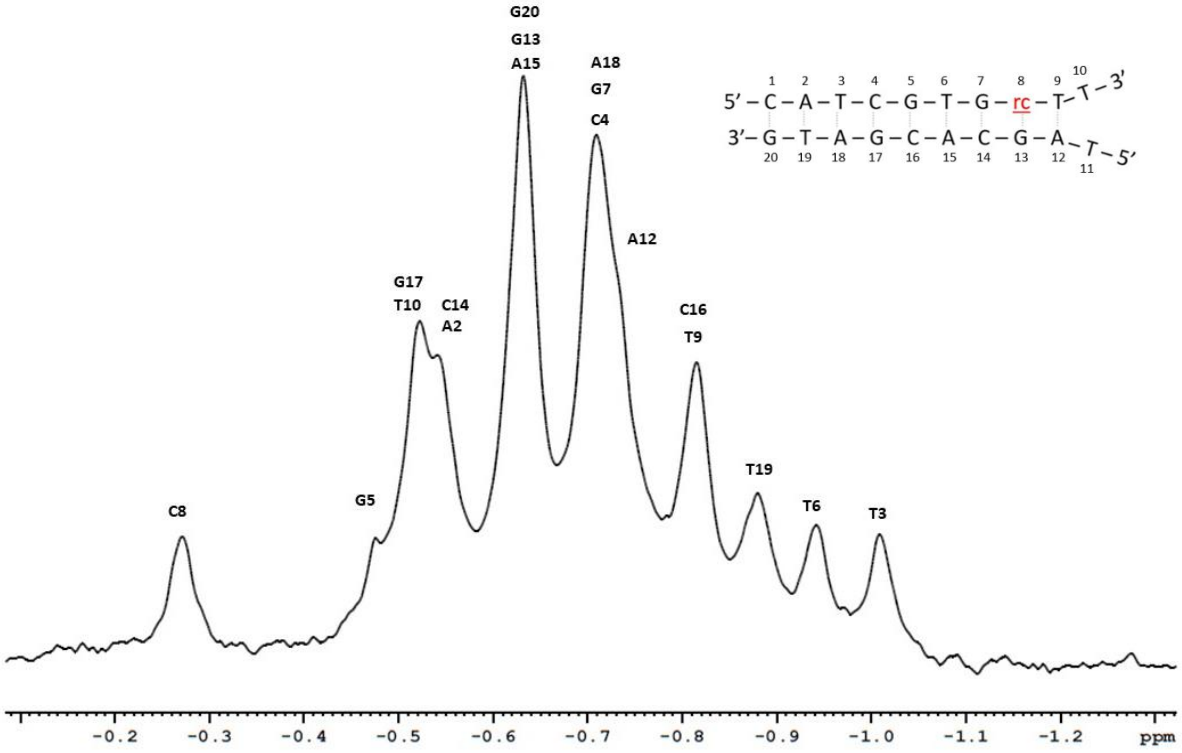


Figure 37: RF RNA ^{31}P spectrum with residue assignment 600 MHz NMR at 303 K.

Table 8 is the chemical shift delta per residue from the RF DNA oligo to the ribo adulterated oligo. The ribo containing residue is denoted with a parenthetical (r). Figure 38 is the graph of ^{31}P chemical shift delta values per residue. The leading strand residues are orange, and the lagging strand is blue. The phosphorus is labeled according to the base for which it is 5′. Therefore, C1 and T11 will not have chemical shifts associated with them.

Table 8: Changes in ^{31}P chemical shifts (ppm) per residue.

	DNA - 31P Chemical Shift	RNA - 31P Chemical Shift	Δ
C1	-	-	-
A2	-0.537	-0.541	-0.004
T3	-0.999	-1.003	-0.004
C4	-0.686	-0.702	-0.016
G5	-0.508	-0.519	-0.011
T6	-0.920	-0.938	-0.018
G7	-0.731	-0.698	0.033
(r) C8	-0.508	-0.268	0.240
T9	-0.814	-0.807	0.007
T10	-0.582	-0.626	-0.045
T11	-	-	-
A12	-0.742	-0.729	0.013
G13	-0.764	-0.612	0.152
C14	-0.556	-0.716	-0.159
A15	-0.680	-0.631	0.049
C16	-0.800	-0.811	-0.011
G17	-0.515	-0.516	0.000
A18	-0.704	-0.706	-0.002
T19	-0.868	-0.879	-0.011
G20	-0.625	-0.633	-0.008

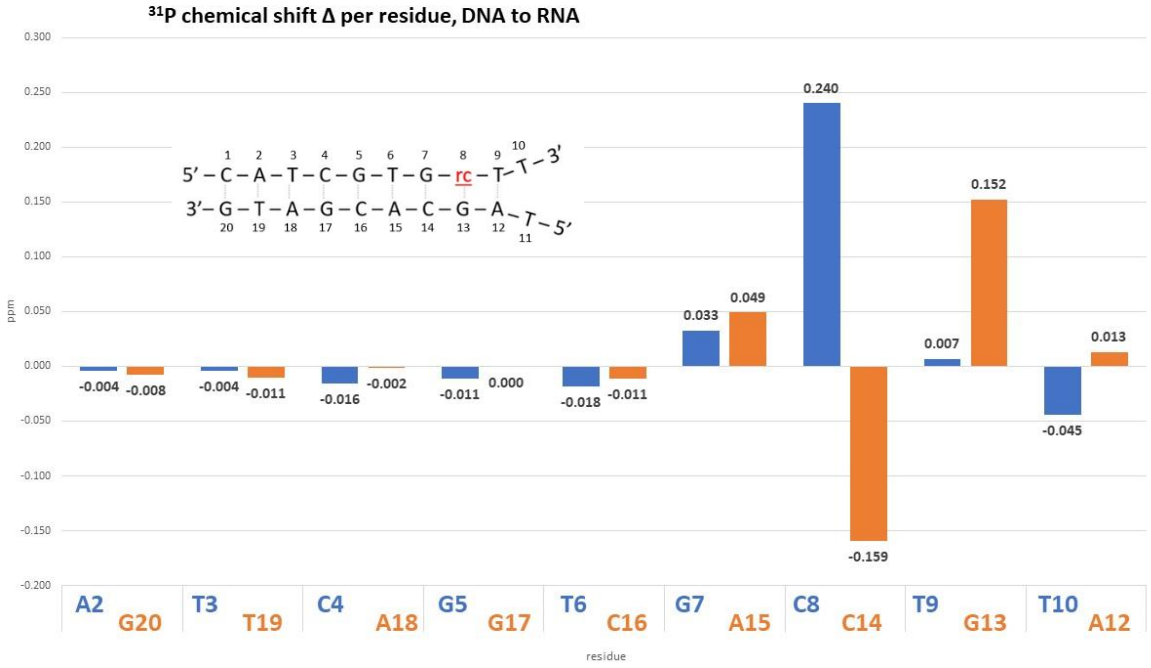


Figure 38: ^{31}P chemical shift delta per residue from DNA to RNA.

4.2.2.1 ^{1}H-^{31}P correlation

^{31}P is correlated with the 3′, 4′, and 5′ protons in a 2D HETCOR experiment. ω2 is the proton frequency and contains water peak, while ω1 is the phosphorus frequency. The 4′ and 5′ resonances are upfield in the ω2 and are crowded. Figure 39 shows the two spectra overlap to highlight change. The DNA spectrum is in black, and the RNA spectrum is in red. Both spectra were referenced to a phosphoric acid insert. Notable changes are highlighted with pointed arrows. C8P to G7H3′ moves downfield in both ω1 and ω2. The most drastic change is the T9P to C8H3′ peak moving far up field in ω2 but does not change in ω1.

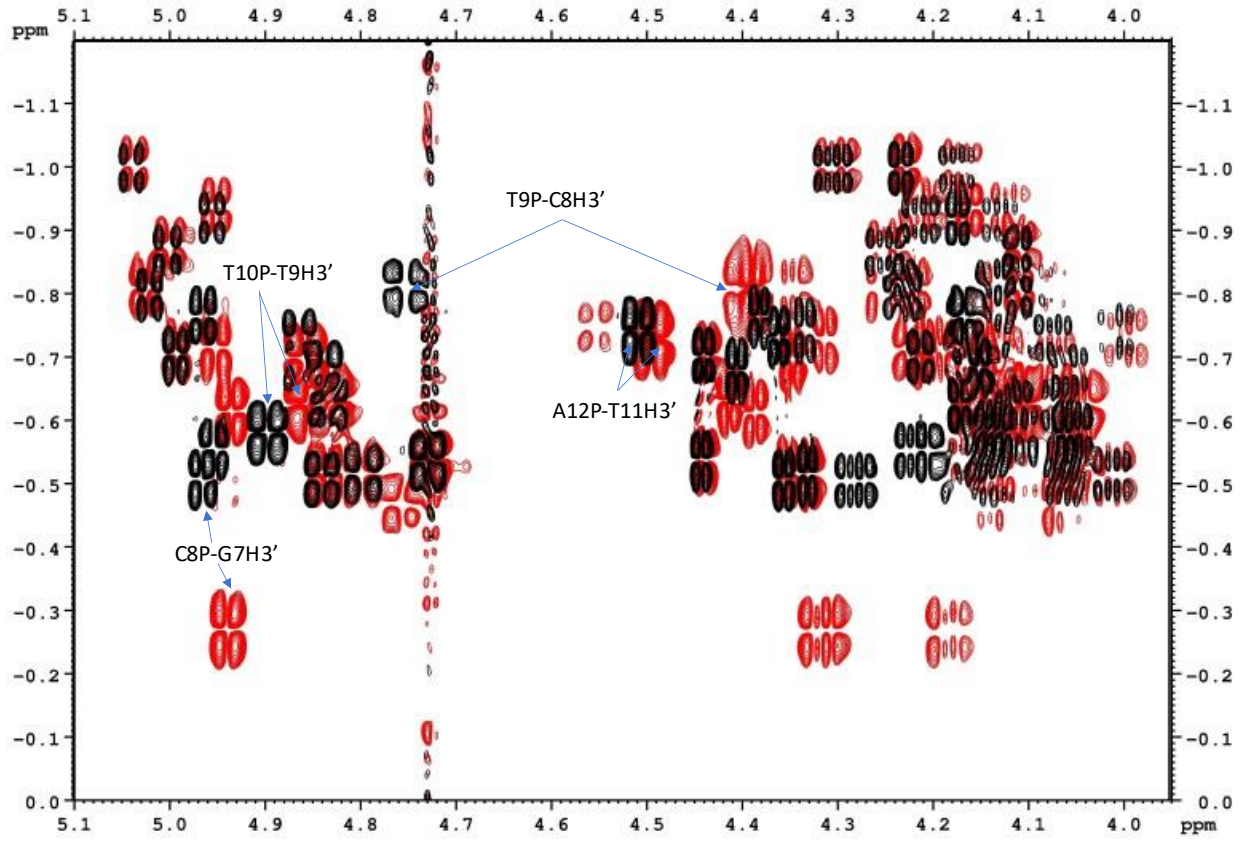


Figure 39: ^{31}P-^{1}H correlation spectra overlay of the RF DNA and RF RNA. Spectra obtained from 600 MHz NMR at 303 K. The DNA cross-peaks are in black and RNA cross-peaks are in red.

4.2.3 Sugar proton assignment

Following ^{31}P temperature gradient experiments a full lyophilization and work up into 99.9998% D_2O was carried out to remove solvent peak at 4.7 ppm. Many of the peaks critical to residue assignment lie under the 4.7 ppm water peak and need to be resolved through solvent suppression. The peaks associated with H4′ and H5′/H5″ resonances overlap and exhibit crowding. Further steps will be taken to resolve these. Figures 40 and 41 are ^{1}H spectra of RF oligonucleotides DNA and RNA in D_2O using a presaturated pulse program with each region of different proton environments labeled. The resolution of these spectra around 4.7 ppm demonstrates the power of presaturation on a completely deuterated solvent. Base protons remain downfield in the region of 7 to 8.4 ppm. Region locations for base and sugar protons are labeled accordingly.

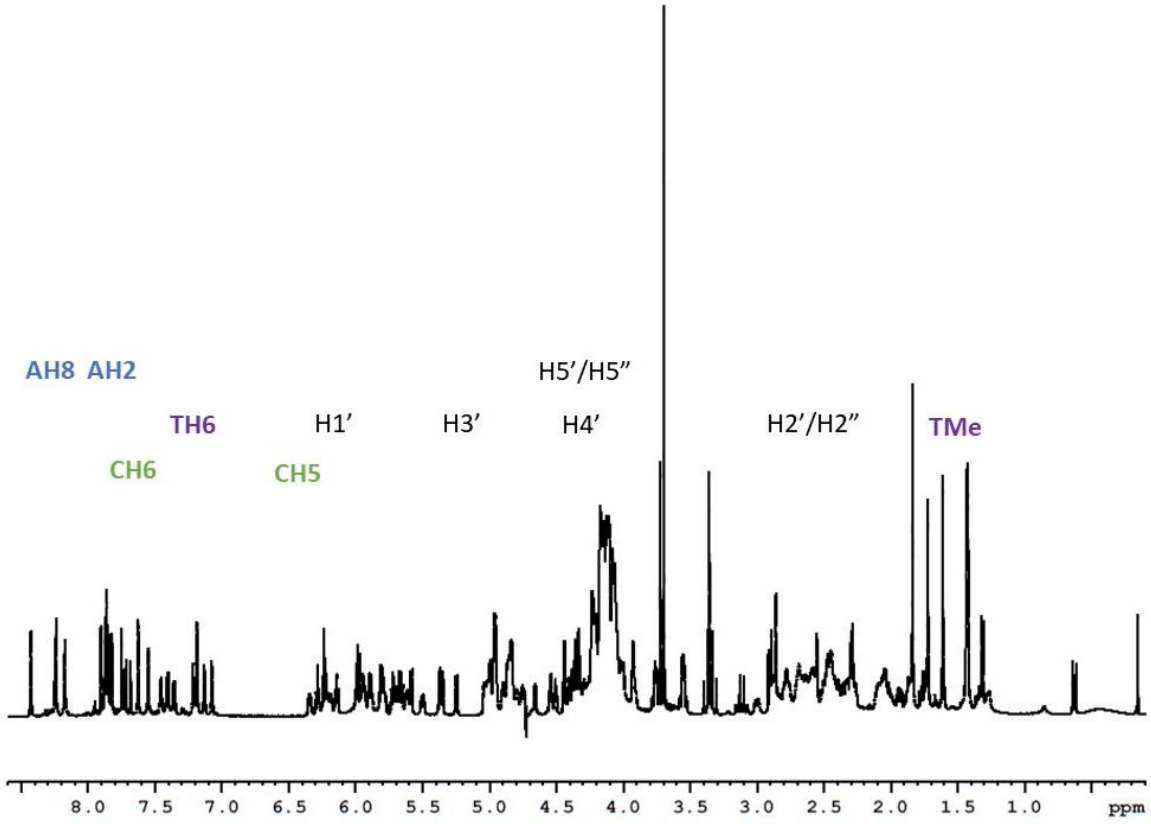


Figure 40: RF DNA ^{1}H spectrum at 303 K with presaturation on a 600 MHz NMR.

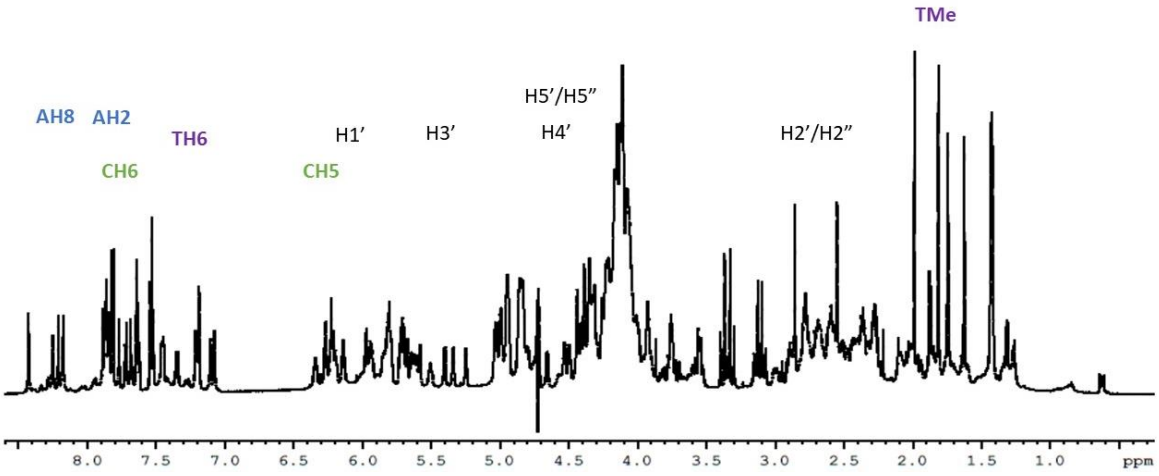


Figure 41: RF RNA ^{1}H spectrum at 303 K with presaturation on a 600 MHz NMR.

4.2.3.1 H1′

Nuclear Overhauser effect spectroscopy (NOESY) yields important information about the spatial relation between neighboring nuclei. The Base to H1′ pathway was the first to be elucidated using 2D NOE experiments. Each actual NOE (not artifact) that appears in the regions of 8.6 – 7.0 ppm (F2) and 6.4 – 5.0 ppm (F1) is an NOE between an anterior base proton (AH8, TH6, etc.) and a sugar proton on the 1′ carbon. With the exception of CH5 – CH6 NOEs which also appear in this region as a doublet cross peak.

Figure 42 is the assignment of the first three residues in the base to H1′ pathway of the RF DNA duplex. 303 K, 250 ms mixing time. Each of the five cytosine H5-H6 NOEs are highlighted magenta. Base proton to sugar H1′ NOEs highlighted blue. As it is a duplex there are two pathways that need to be assigned. The first residue in the oligonucleotide sequence (C1) will have only one base to H1′ cross peak and this is the start of the base to H1′ pathway. However, because the first residue is a cytosine, the first NOE Assignment continues from H1′ to next the next residue base represented by an NOE cross peak. Each residue after the first with have two NOE in the F1 and two NOE in the F2. Vertical progression (ω1) from an NOE represents a connection between a base proton to its own sugar H1′. While a horizontal progression (ω2) represents the sugar H1′ to the base of the next residue. This creates an assignment pattern for each strand (Base to own sugar, to next residue base, etc....)

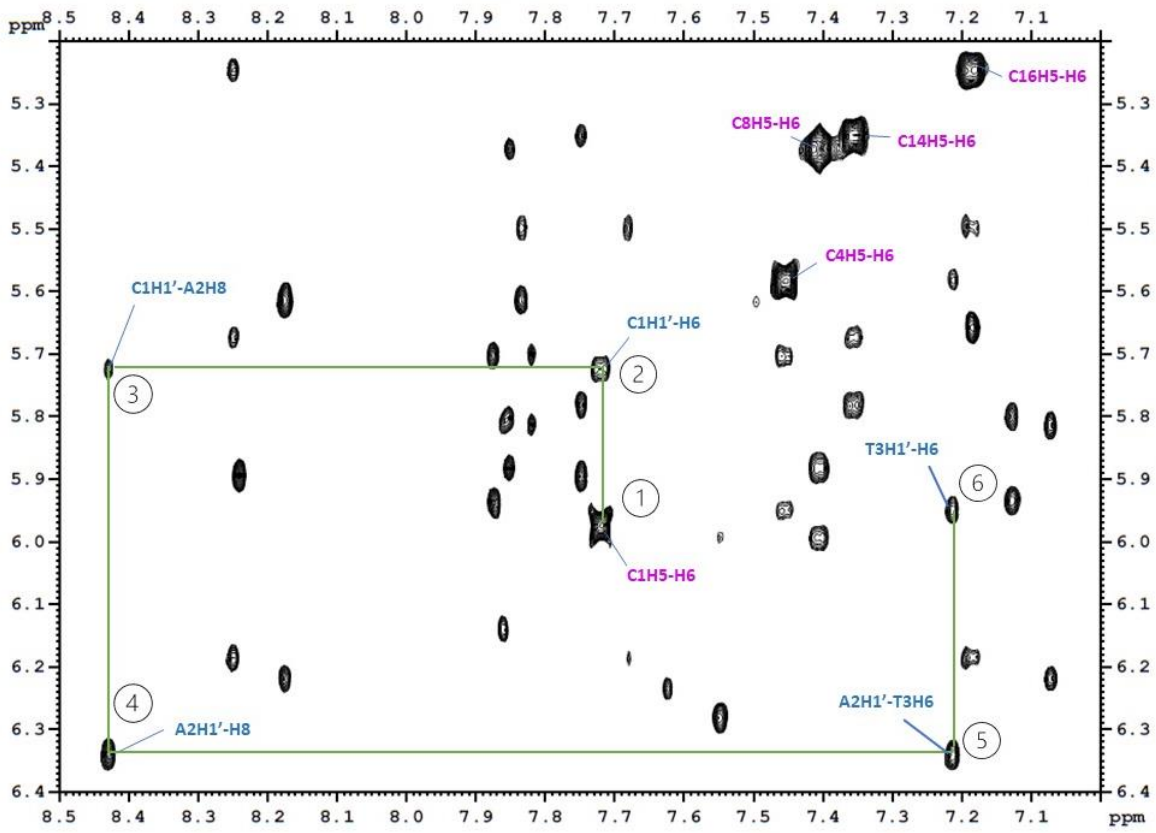


Figure 42: RF DNA ^{1}H NOESY spectrum at 303 K, base to H1′ region (250 ms mixing time).

Figure 43 is the completed assignment of each residue in the base to H1′ region. H1′ resonances are displayed in the vertical axis (ω1) and the base protons in the horizontal (ω2).

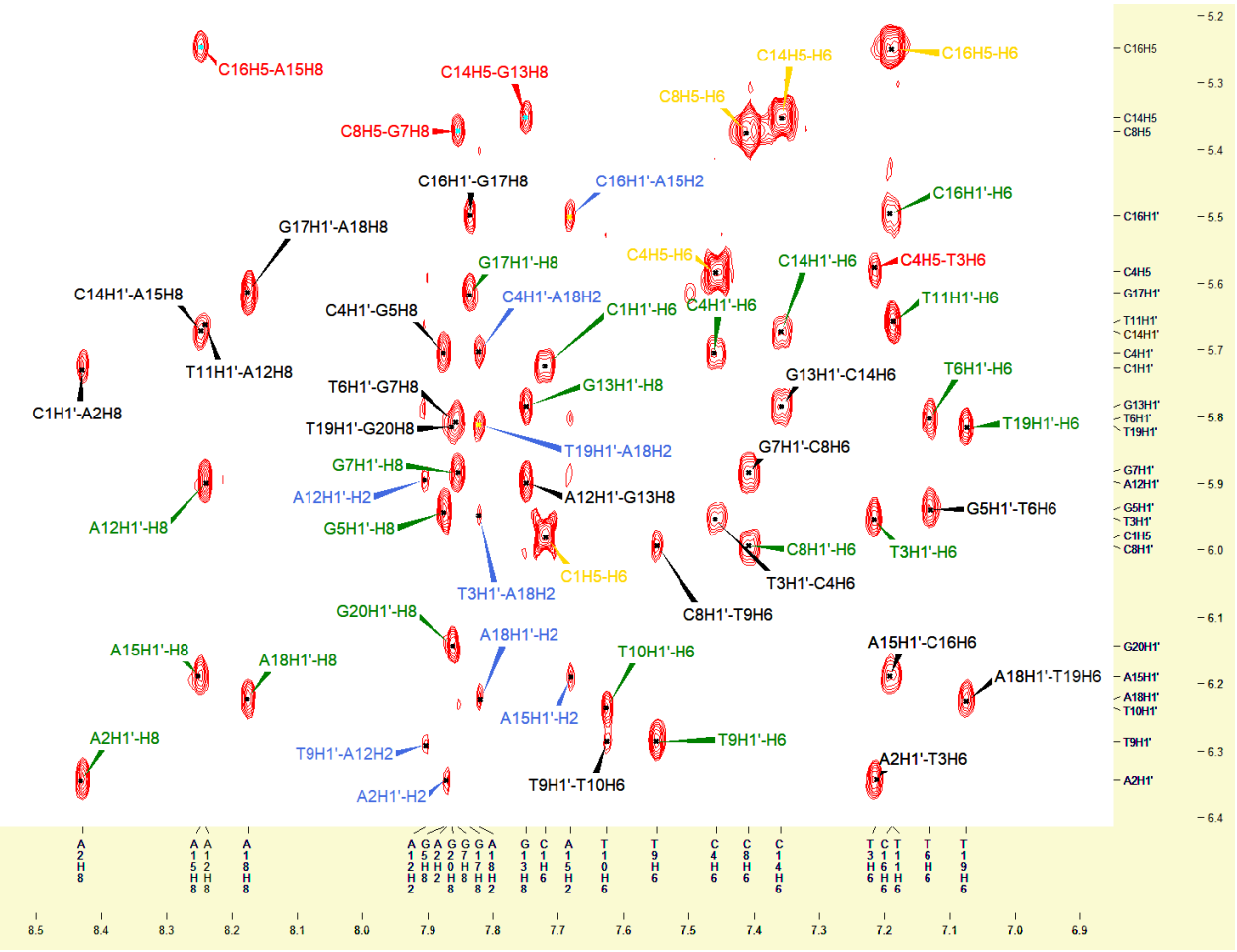


Figure 43: Sparky assignment of each residue in the RF DNA base to H1′ region.

Figure 44 is the base to H1′ region of the RF RNA oligo. The C8 H5-H6 NOE has shifted significantly from Figure 42. Again, showing the base to H1′ pathway for the first three residues. Each of the five cytosine H5-H6 NOEs are highlighted magenta. Base proton to sugar H1′ NOEs highlighted blue. Only the first three residues are shown.

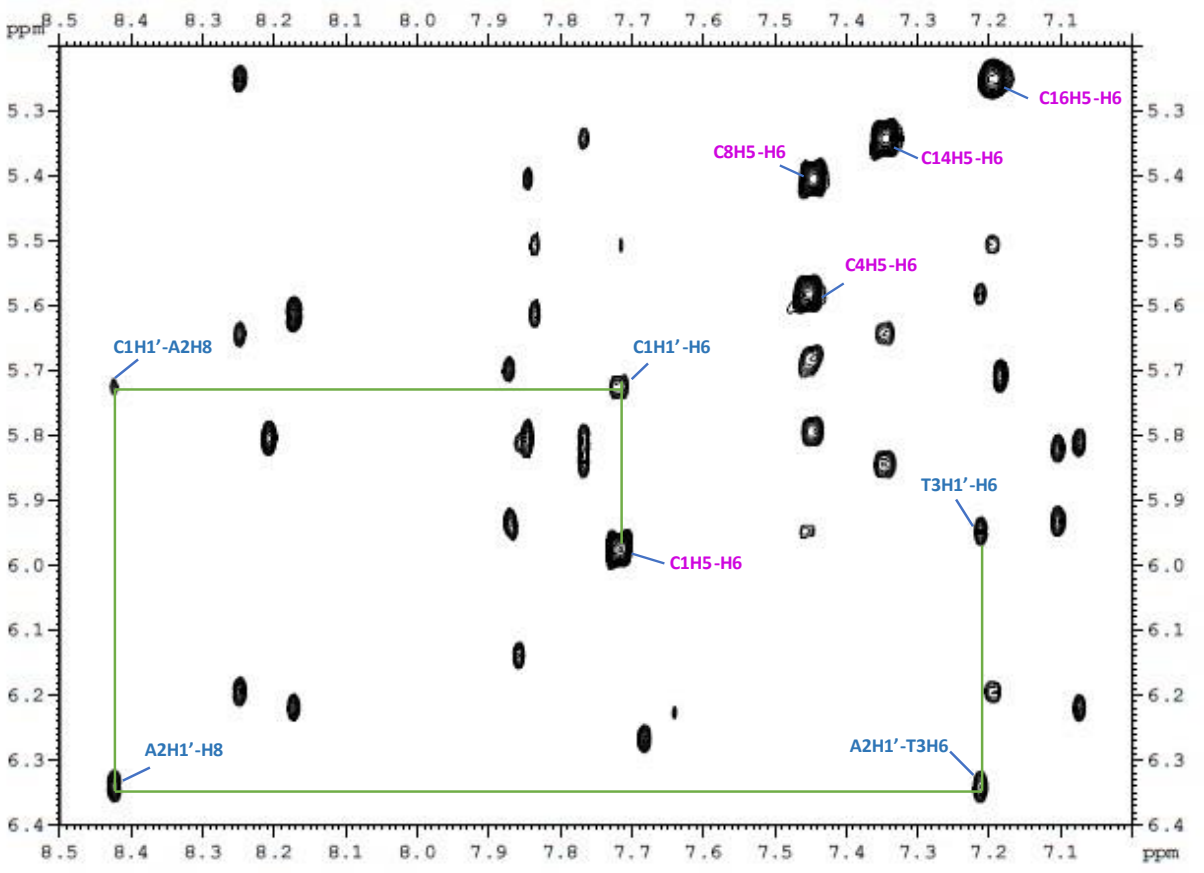


Figure 44: RF RNA ^{1}H NOESY spectrum at 303 K, base to H1′ region (250 ms mixing time).

Figure 45 is the assignment of each residue in the base to H1′ region of the RF RNA NOESY spectrum. Table 9 is the H1′ chemical shift delta per residue from DNA to RNA.

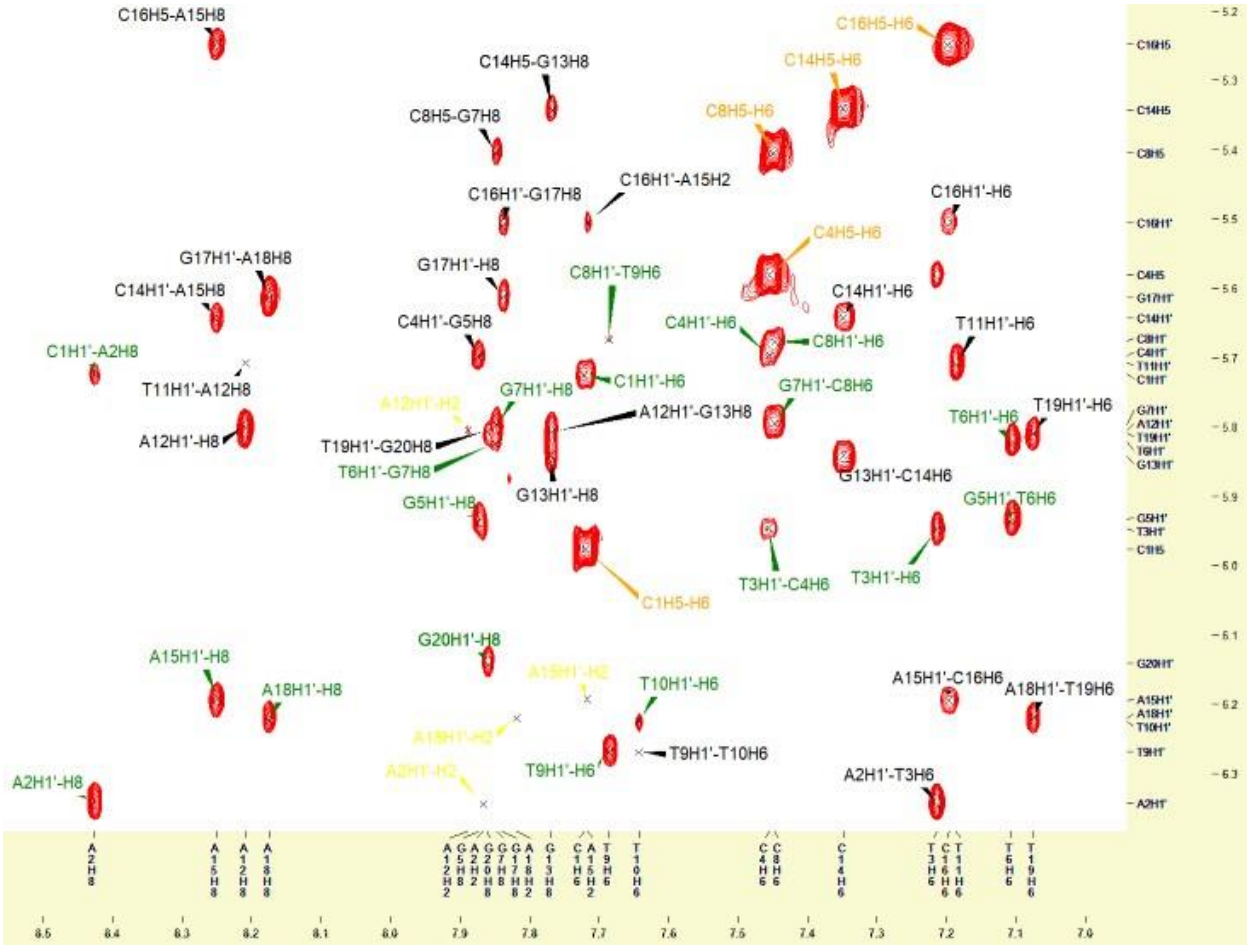


Figure 45: Sparky assignment of each residue NOE in the RF RNA base to H1' region.

Table 9: H1' Chemical shift delta (ppm) per residue from DNA to RNA

	DNA H1’ chemical shift (ppm)	RNA H1’ chemical shift (ppm)	Δ (ppm)
C1	5.726	5.724	0.002
A2	6.345	6.344	0.001
T3	5.951	5.947	0.004
C4	5.704	5.698	0.006
G5	5.938	5.935	0.003
T6	5.803	5.822	-0.019
G7	5.882	5.796	0.086
(r)C8	5.994	5.676	0.318
T9	6.286	6.269	0.017
T10	6.237	6.226	0.011
T11	5.66	5.708	-0.048
A12	5.896	5.804	0.092
G13	5.784	5.842	-0.058
C14	5.674	5.643	0.031
A15	6.189	6.194	-0.005
C16	5.499	5.506	-0.007
G17	5.614	5.614	0.000
A18	6.224	6.221	0.003
T19	5.815	5.809	0.006
G20	6.142	6.141	0.001

Figure 46 is the graph of the H1′ chemical shift changes per residue from the DNA to RNA containing oligonucleotide. As with the imino proton graph previously, the stem protons remain unperturbed, while the mass of the change occurs around the C8 residue.

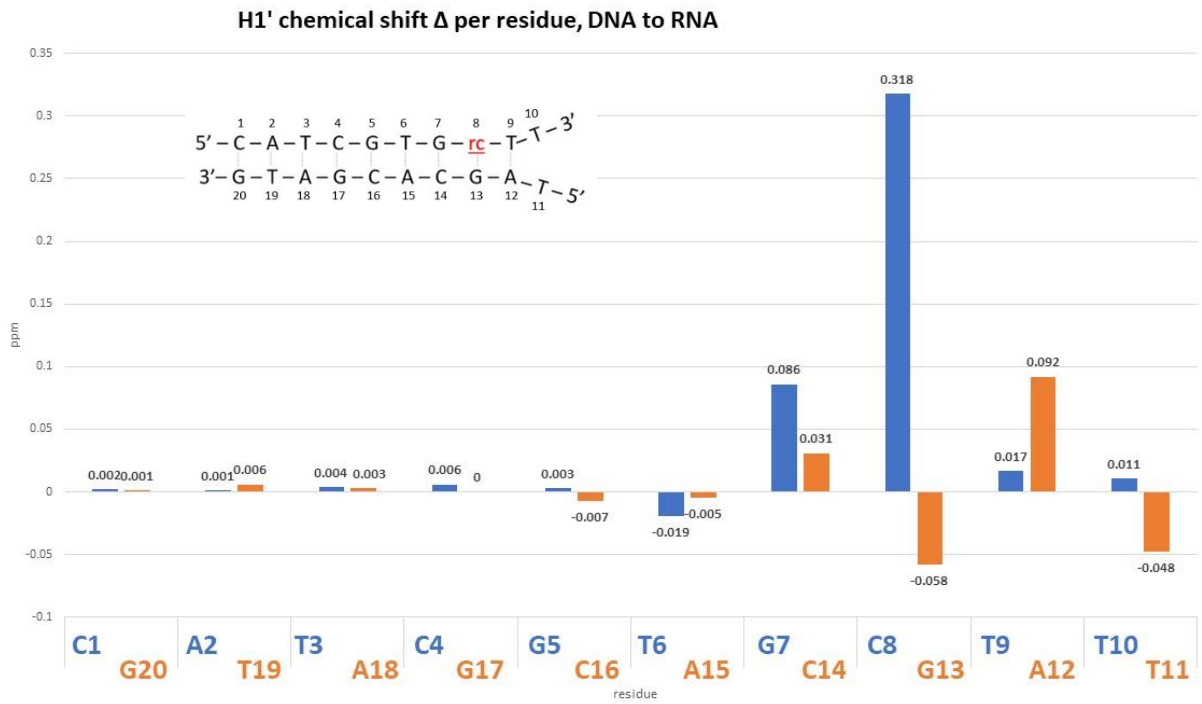


Figure 46: Chemical shift delta of H1' protons per residue between RF DNA and RF RNA oligonucleotides.

Figure 47 is an overlay of the RF DNA and RF RNA spectra in the H1′ region with a few changes highlighted. The NOE of C1, A2, and T3 are perfectly on top of each other exhibiting the localization of the ribo lesion damage. While NOE associated with the ribo lesion has shifted quite a bit. C8 H5-H6 has moved downfield in both directions while the neighboring NOE for C14 H5-H6 remains in the same resonance. Other changes include the NOE for G13H1′ to

G13H8, this NOE has moved downfield in both directions as well. The NOE for A12H1′ to G13H8 has moved downfield in the ω1 but moved up field in the ω2.

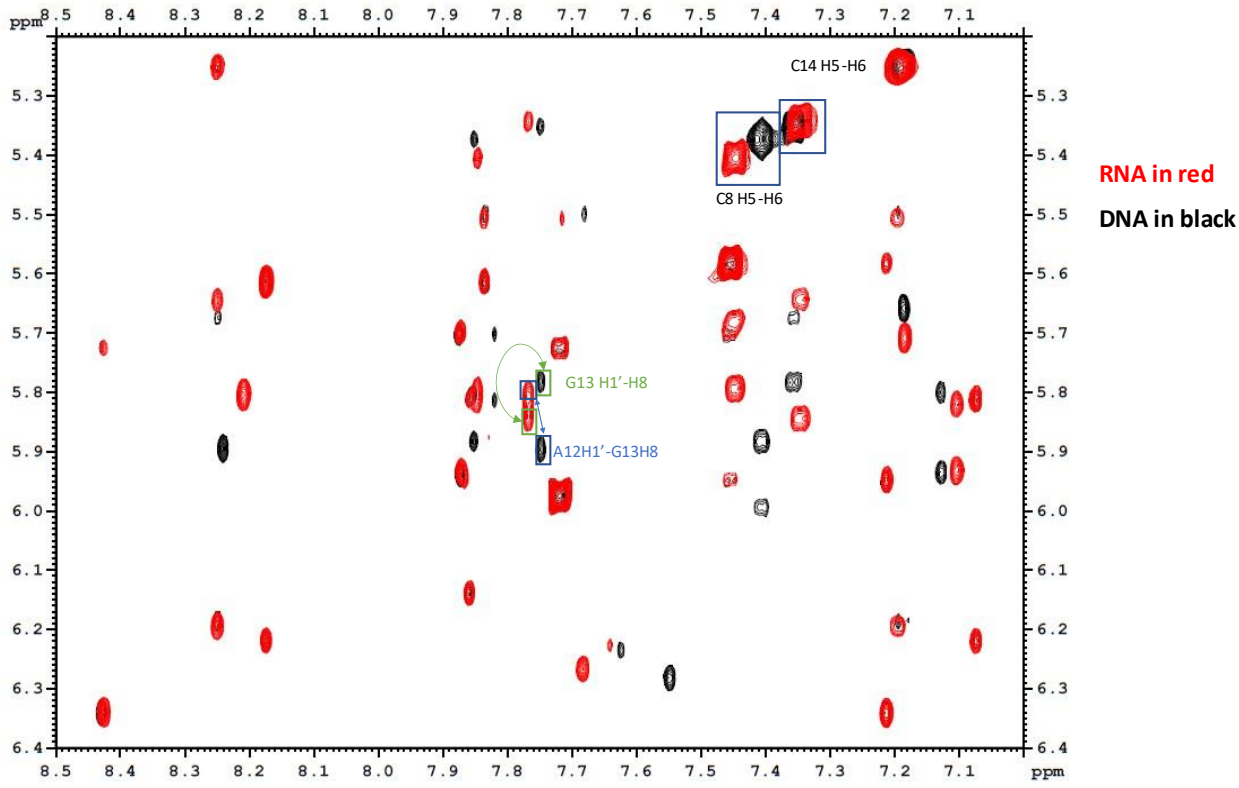


Figure 47: Overlay of the ^{1}H NOESY spectra DNA and RNA RF analogs, base to H1' region.

4.2.3.2 H2′/H2″

The 2′ protons are assigned in a pathway differing from the one followed H1′. Figure 48 is an example of the H2′/H2″ pathway and the expected NOE generation. Only the first three residues are shown as an example of assignment. The residues are labeled, and arrows color coded accordingly. The weight intensity of each arrow is adjusted to increase with the greater strength of each NOE that is to be expected based on special relation. The RF RNA oligonucleotide is lacking an H2″ proton on the C8 residue so there will not be an H2″ resonance.

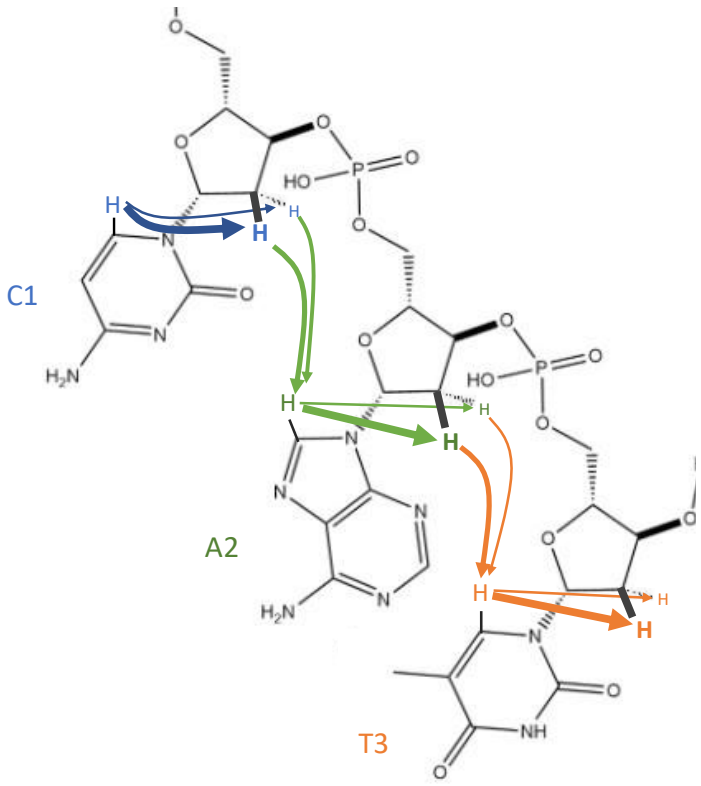


Figure 48: Method of proton assignment in the NOESY base to H2′ / H2″ assignment pathway.

Figure 49 shows the pathway for assigning H2′ and H2″ NOE on the RF RNA oligonucleotide as they interact with base protons. Each vertical progression of the H2′/H2″ pathway is a connection from a base proton to its own sugar. While each horizontal progression along the H2′/H2″ pathway is a connection to the next residues base proton. For simplicity, the assignments of the first three residues are labeled. The labels are color coordinated to that of Figure 48.

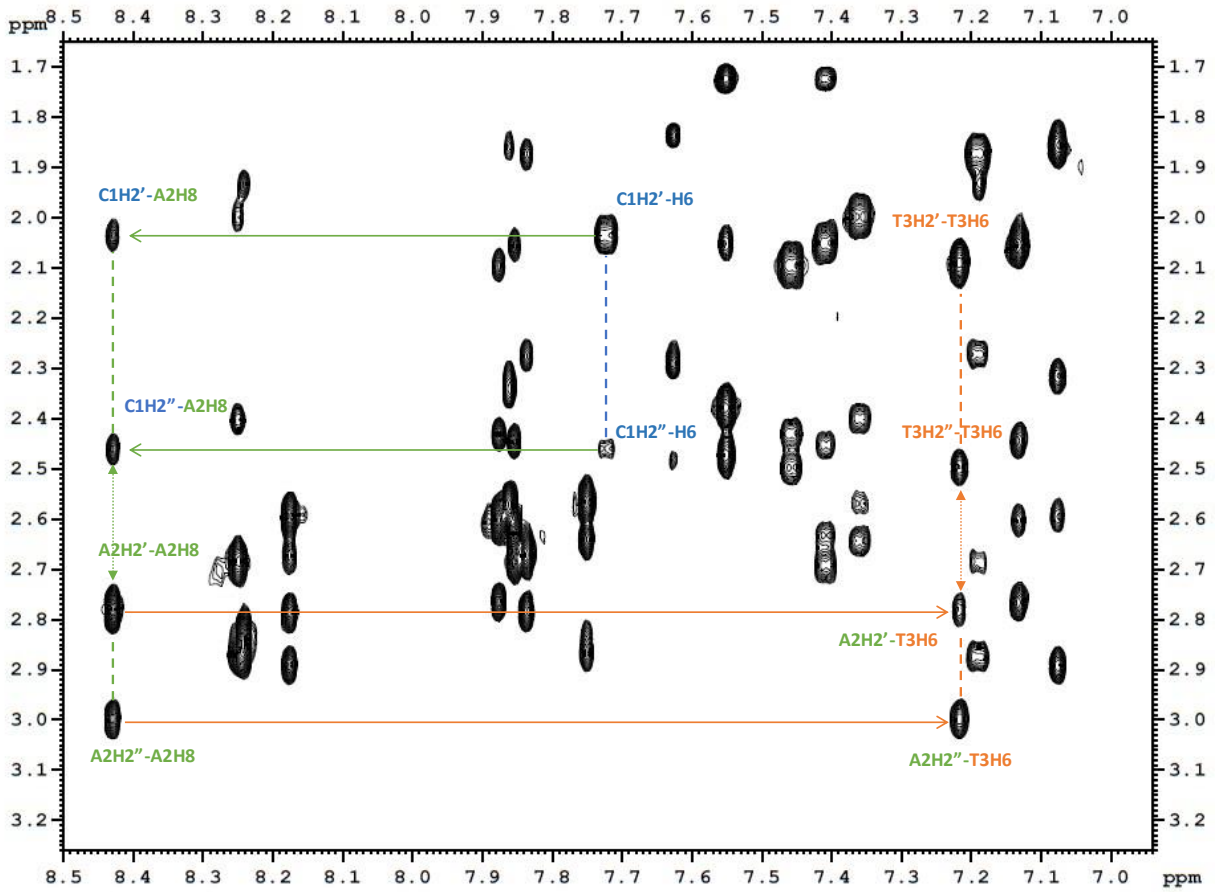


Figure 49: NOESY Base to H2'/H2" assignment pathway for RF DNA oligonucleotide.

Figure 50 below shows the final assignment for all RF DNA H2′/H2″ resonances.

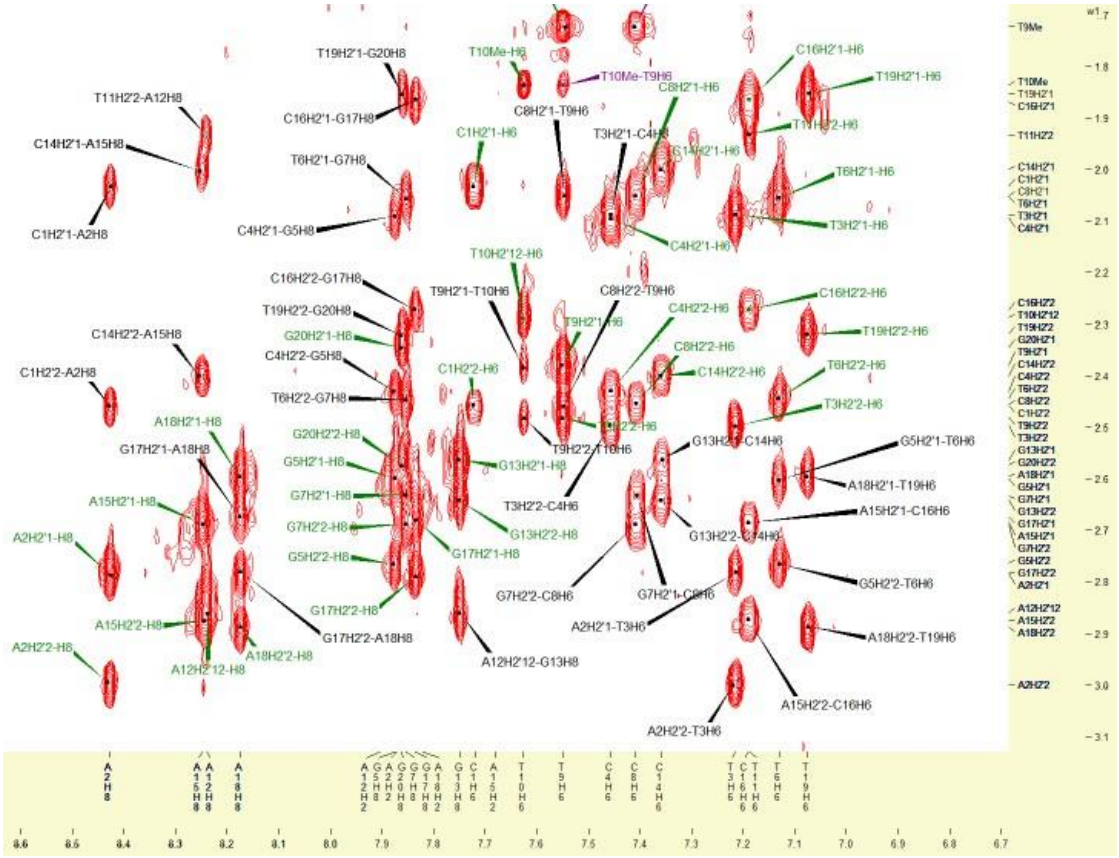


Figure 50: RF DNA H2'/H2" resonance assignments.

Figure 51 is the RF RNA NOESY spectrum in the H2′/H2″ region with the first three residue assignment resonances displayed on top of the pathway. These resonances remain the same as the first three residues of the RF DNA oligonucleotide. No change in resonance for H2′/H2″ resonances is observed for the first 5 residues.

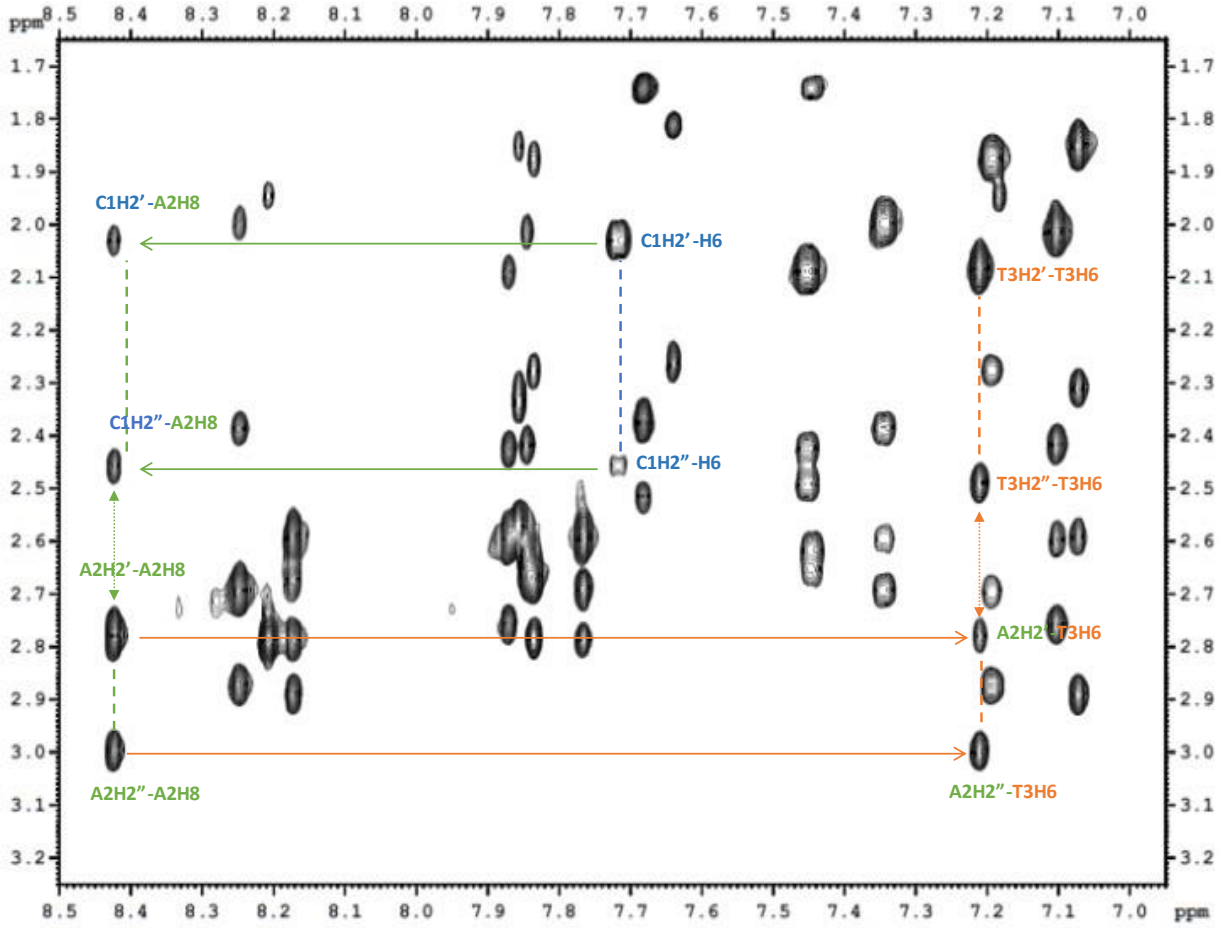


Figure 51: Base to H2'/H2" pathway in the RF RNA NOESY spectrum.

Figure 52 below shows the final assignment for all RF RNA H2′/H2″ resonances.

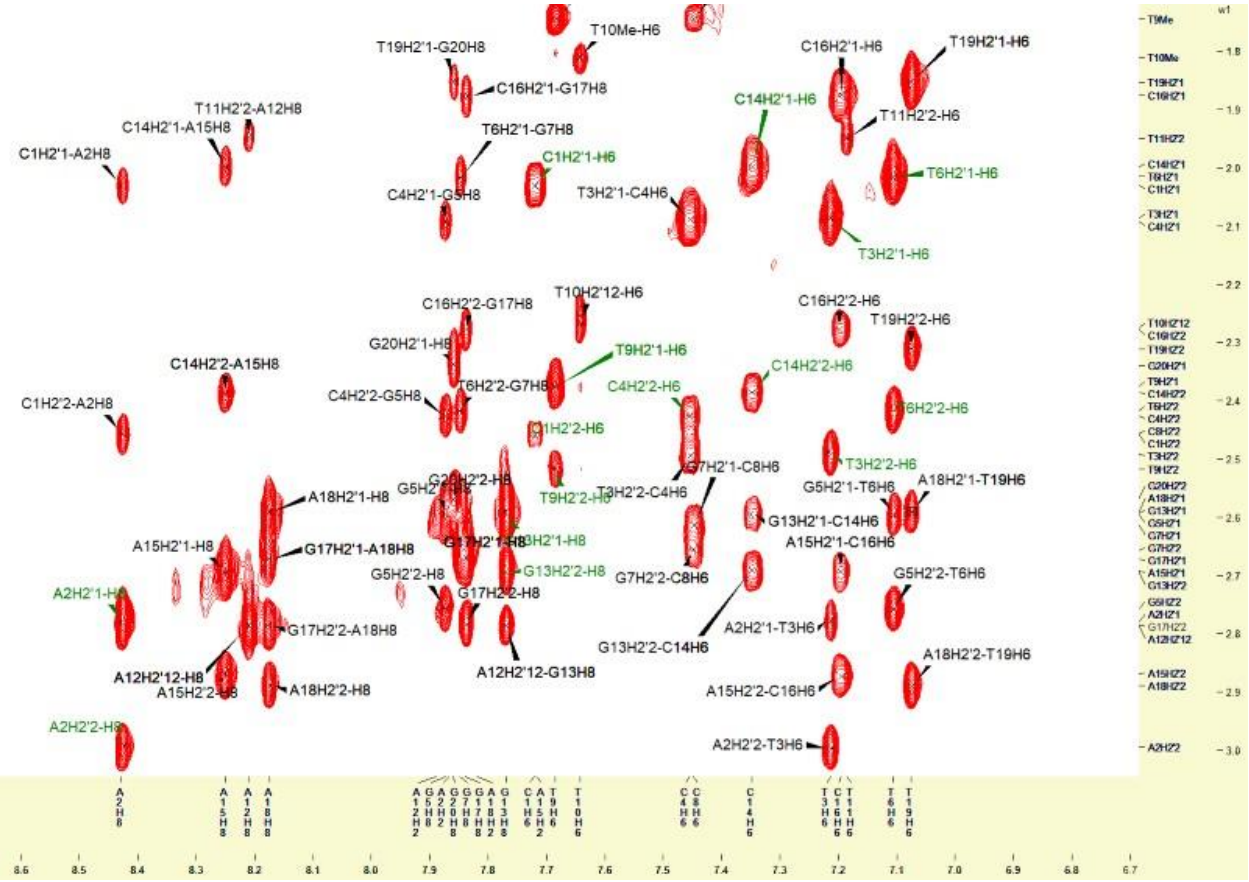


Figure 52: RF RNA H2'/H2'' resonance assignments in the base to H2'/H2'' region.

Table 10 below consolidates the H2′ chemical shifts data and changes for each residue across DNA to RNA oligonucleotides.

Table 10: Change in H2' chemical shift (ppm) in the RF analogs.

	DNA H2' Chemical shift (ppm)	RNA H2' Chemical shift (ppm)	Δ(ppm)
C1	2.032	2.031	0.001
A2	2.784	2.780	0.004
T3	2.087	2.087	0.000
C4	2.093	2.092	0.001
G5	2.6	2.598	0.002
T6	2.054	2.015	0.039
G7	2.633	2.614	0.019
C8	2.052	4.128	-2.076
T9	2.379	2.375	0.004
T10	-	-	
T11	1.335	1.336	-0.001
A12	-	-	
G13	2.564	2.594	-0.030
C14	2.00	1.998	0.002
A15	2.685	2.692	-0.007
C16	1.867	1.877	-0.010
G17	2.674	2.674	0.000
A18	2.595	2.593	0.002
T19	1.852	1.856	-0.004
G20	2.345	2.341	0.004

Graphing the H2′ chemical shift changes, figure 53, shows a drastic change in the C8 residue while all the other residue H2′ resonances are relatively unaffected.

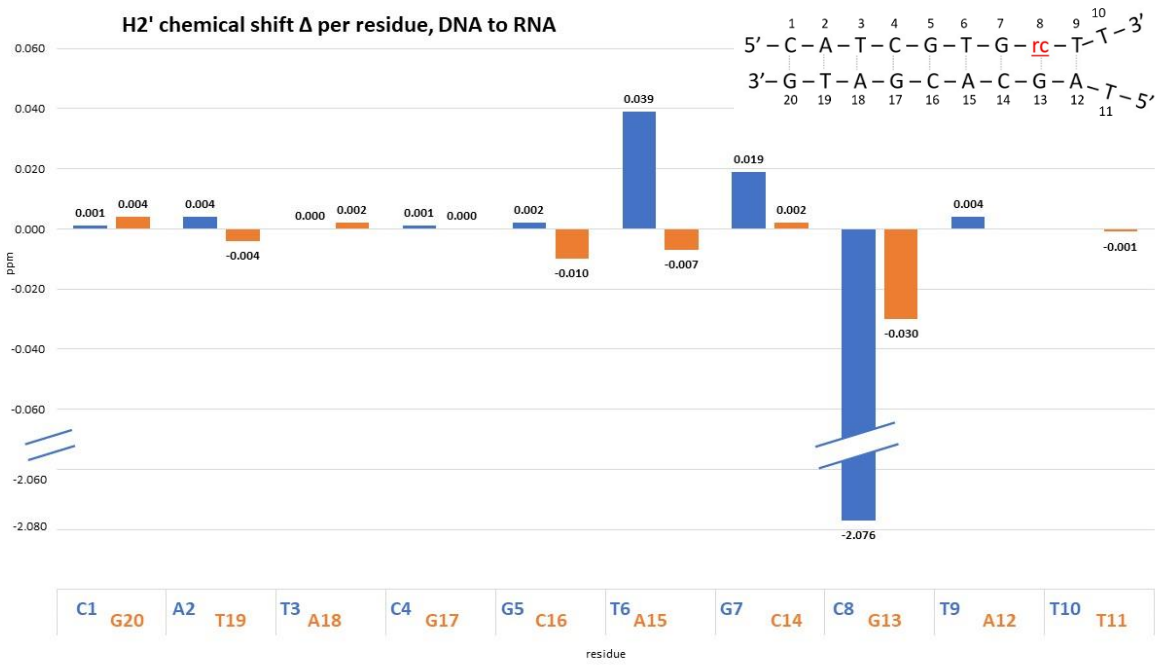


Figure 53: Change in chemical shifts of H2' protons between the DNA and RNA containing RF analogs.

Table 11 displays the H2′ chemical shifts data and delta for each DNA and RNA residue.

Table 11: H2" chemical shift delta (ppm) per residue from DNA to RNA.

	DNA H2” Chemical shift (ppm)	RNA H2” Chemical shift (ppm)	Δ (ppm)
C1	2.458	2.457	0.001
A2	2.999	2.995	0.004
T3	2.497	2.492	0.005
C4	2.43	2.427	0.003
G5	2.764	2.756	0.008
T6	2.444	2.417	0.027
G7	2.69	2.656	0.034
C8	2.456	-	
T9	2.481	2.518	-0.037
T10	-	-	
T11	1.932	1.950	-0.018
A12	-	-	
G13	2.641	2.694	-0.053
C14	2.401	2.386	0.015
A15	2.874	2.867	0.007
C16	2.272	2.279	-0.007
G17	2.784	2.786	-0.002
A18	2.887	2.887	0.000
T19	2.319	2.313	0.006
G20	2.572	2.567	0.005

Figure 54 shows the relative change in resonances about the H2″ chemical shifts, of which C8 is missing.

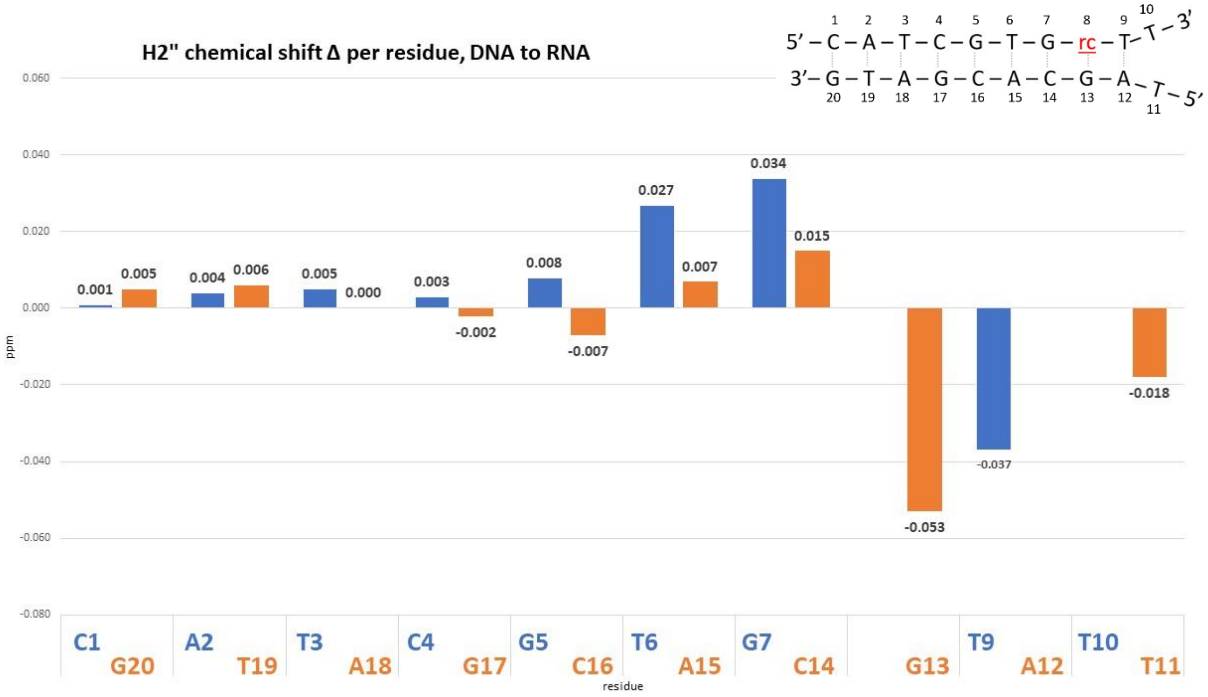


Figure 54: Change in chemical shifts of H2" protons per residue between our DNA and RNA containing RF analogs.

Figure 55 is the overlay of the two NOESY spectrum in the H2′/H2″ region. The DNA oligonucleotide resonances are in black, and the RNA resonances are in red. Similar to the H1′ region the NOE for the stem residues overlap perfectly. Notable changes in the spectra are highlighted with labels. The RNA oligonucleotide does not possess an H2″ so there is no resonance for C8H2″ in the RNA spectrum. We still observe the effect from the perturbed residue. G7, C8, T9, and A12 NOE's are labeled and highlighted with arrows. The C8H2′ resonances disappear from the region and move far downfield out of the range of the plot. The neighboring nucleotides, G7 and T9, H2′/H2″ resonances all move downfield in the ω2 frequency. The G7 NOE moves downfield in the ω2 but slightly up field in the ω1.

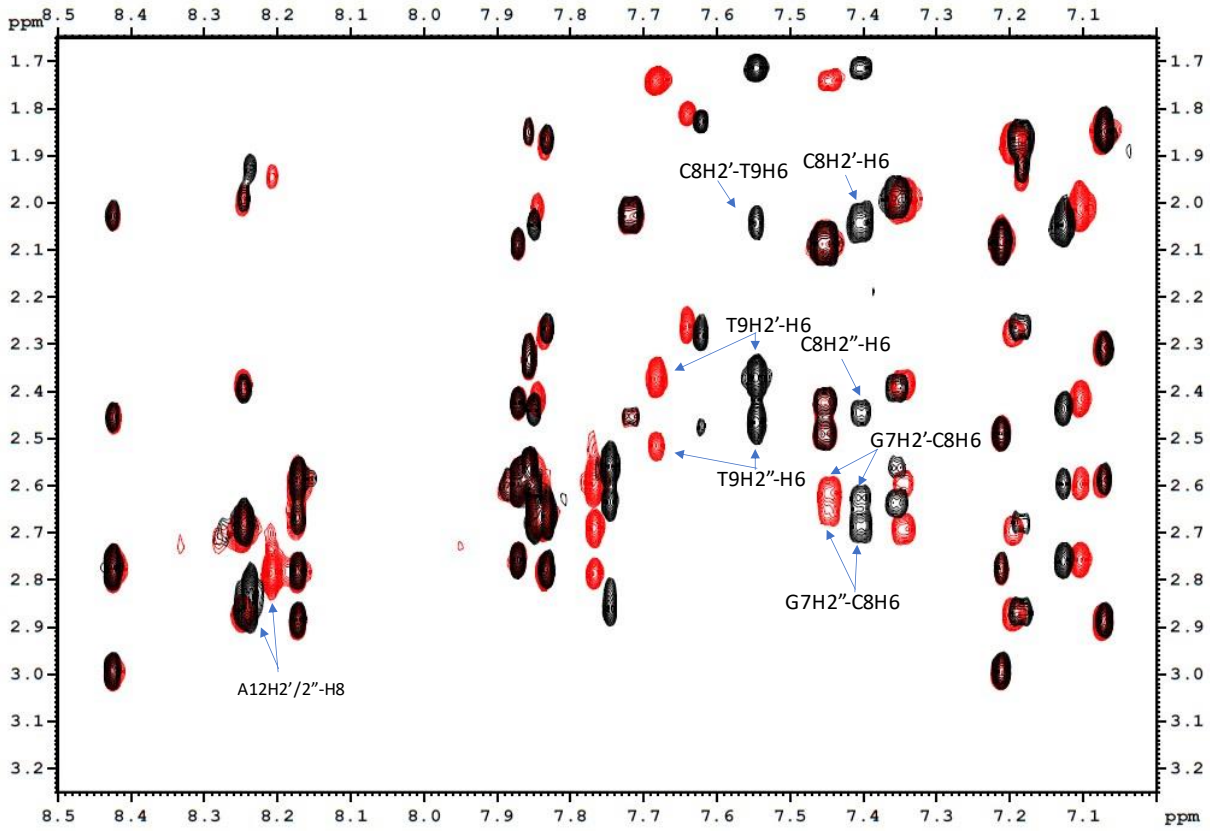


Figure 55: Overlay NOESY spectra of the RF DNA and RF RNA oligonucleotides in the base to H2'/H2'' region. DNA is in black, and RNA is in red.

Assignments for the rest of the residue protons continued in a similar manner to the H1′ and H2′/H2″ pathways. Tables 12 and 13 list all the proton chemical shift assignments for each residue. Table 14 is the change in chemical shift with the degree of change emphasized by gradient shading. The shading shows the degree of the damage with a darker red color indicating greater positive chemical shift change and a darker green indicating a greater negative change.

Table 12: Final residue chemical shift assignments, RF DNA.

DNA assignments (ppm)											
	H1'	**H2'**	**H2"**	**H2'/"**	**H3'**	**H4'**	**H2**	**H5**	**H6**	**H8**	**Me**
C1	5.726	2.032	2.458	-	4.728	4.078	-	5.981	7.720	-	-
A2	6.345	2.784	2.999	-	5.039	4.442	7.871	-	-	8.428	-
T3	5.951	2.087	2.497	-	4.860	4.231	-	-	7.215	-	1.421
C4	5.704	2.093	2.430	-	4.838	4.151	-	5.582	7.458	-	-
G5	5.938	2.600	2.764	-	4.954	4.354	-	-	-	7.875	-
T6	5.803	2.054	2.444	-	4.862	4.172	-	-	7.130	-	1.428
G7	5.882	2.633	2.690	-	4.964	4.360	-	-	-	7.854	-
C8	5.994	2.052	2.456	-	4.753	4.176	-	5.373	7.408	-	-
T9	6.286	2.379	2.481	-	4.898	4.229	-	-	7.549	-	1.723
T10	6.237	-	-	2.287	4.545	4.103	-	-	7.626	-	1.836
T11	5.660	1.335	1.932	-	4.507	3.923	-	-	7.187	-	1.610
A12	5.896	-	-	2.861	4.963	4.333	7.905	-	-	8.240	-
G13	5.784	2.564	2.641	-	4.949	4.380	-	-	-	7.750	-
C14	5.674	2.000	2.401	-	4.836	4.201	-	5.352	7.359	-	-
A15	6.189	2.685	2.874	-	5.020	4.404	7.681	-	-	8.249	-
C16	5.499	1.867	2.272	-	4.797	4.107	-	5.247	7.190	-	-
G17	5.614	2.674	2.784	-	4.992	4.330	-	-	-	7.835	-
A18	6.224	2.595	2.887	-	5.000	4.438	7.820	-	-	8.175	-
T19	5.815	1.852	2.319	-	4.832	4.112	-	-	7.075	-	1.433
G20	6.142	2.345	2.572	-	4.658	4.168	-	-	-	7.862	-

Table 13: Final residue chemical shift assignments, RF RNA.

RNA assignments (ppm)											
	H1'	**H2'**	**H2”**	**H2’/”**	**H3'**	**H4'**	**H2**	**H5**	**H6**	**H8**	**Me**
C1	5.724	2.031	2.457	-	4.732	4.082		5.977	7.721	-	-
A2	6.344	2.780	2.995	-	5.037	4.439	7.867	-	-	8.428	-
T3	5.947	2.087	2.492	-	4.859	4.229	-	-	7.214	-	1.418
C4	5.698	2.092	2.427	-	4.837	4.150	-	5.581	7.455	-	-
G5	5.935	2.598	2.756	-	4.951	4.353	-	-	-	7.874	-
T6	5.822	2.015	2.417	-	4.858	4.163	-	-	7.106	-	1.426
G7	5.796	2.614	2.656	-	4.942	4.336	-	-	-	7.848	-
rC8	5.676	4.128	-	-	4.390	-	-	5.405	7.448	-	-
T9	6.269	2.375	2.518	-	4.858	4.258	-	-	7.685	-	1.745
T10	6.226	-	-	2.285	4.530	4.108	-	-	7.643	-	1.810
T11	5.708	1.336	1.950	-	4.495	3.929	-	-	7.186	-	1.626
A12	5.804	-	-	2.786	4.939	4.314	7.889	-	-	8.209	-
G13	5.842	2.594	2.694	-	4.953	4.387	-	-	-	7.769	-
C14	5.643	1.998	2.386	-	4.835	4.189	-	5.342	7.348	-	-
A15	6.194	2.692	2.867	-	5.027	4.415	7.716	-	-	8.251	-
C16	5.506	1.877	2.279	-	4.801	4.111	-	5.248	7.197	-	-
G17	5.614	2.674	2.786	-	4.993	4.333	-	-	-	7.839	-
A18	6.221	2.593	2.887	-	5.001	4.438	7.819	-	-	8.175	-
T19	5.809	1.856	2.313	-	4.833	4.111	-	-	7.075	-	1.433
G20	6.141	2.341	2.567	-	4.658	4.170	-	-	-	7.861	-

Table 14: Proton chemical shift delta for all residues assigned. Conditional color gradient format is applied. The darker the shade of red equals the greater positive values. While the darker the shade of green indicates the greater negative values.

	Assignment Δ chemical shifts (ppm)										
	H1'	H2'1	H2'2	H2'12	H3'	H4'	H2	H5	H6	H8	Me
C1	0.002	0.001	0.001		-0.004	-0.004		0.004	-0.001		
A2	0.001	0.004	0.004		0.002	0.003	0.004			0.000	
T3	0.004	0.000	0.005		0.001	0.002			0.001		0.003
C4	0.006	0.001	0.003		0.001	0.001		0.001	0.003		
G5	0.003	0.002	0.008		0.003	0.001				0.001	
T6	-0.019	0.039	0.027		0.004	0.009			0.024		0.002
G7	0.086	0.019	0.034		0.022	0.024				0.006	
C8	0.318	-2.076			0.363			-0.032	-0.040		
T9	0.017	0.004	-0.037		0.040	-0.029			-0.136		-0.022
T10	0.011			0.002	0.015	-0.005			-0.017		0.026
T11	-0.048	-0.001	-0.018		0.012	-0.006			0.001		-0.016
A12	0.092			0.075	0.024	0.019	0.016			0.031	
G13	-0.058	-0.030	-0.053		-0.004	-0.007				-0.019	
C14	0.031	0.002	0.015		0.001	0.012		0.010	0.011		
A15	-0.005	-0.007	0.007		-0.007	-0.011	-0.035			-0.002	
C16	-0.007	-0.010	-0.007		-0.004	-0.004		-0.001	-0.007		
G17	0.000	0.000	-0.002		-0.001	-0.003				-0.004	
A18	0.003	0.002	0.000		-0.001	0.000	0.001			0.000	
T19	0.006	-0.004	0.006		-0.001	0.001			0.000		0.000
G20	0.001	0.004	0.005		0.000	-0.002				0.001	

4.2.3.3 Fraction south

The dominant sugar conformation and the extent of the N/S equilibrium can be determined from coupling constants. These 3-bond J coupling (^{3}JH1′-H2′, etc.…) can be measured from the low flip-angle COSY spectra. The ^{3}JH1′-H2′ coupling constant is measured from the passive coupling from the H1′-H2″ COSY cross peak while the ^{3}JH1′-H2″ coupling constant is measured from the passive coupling in the H1′-H2′ cross peak. These are measured in the spectrum from the cross peaks above the diagonal to take advantage of the higher resolution in the F2 dimension. The sum of 2′ and 2″ are measured from the cross peaks below the diagonal for the same reason.

The fraction south of each sugar ring is calculated graphically from J coupling measurements taken from the correlation spectroscopy experiments. Figures 56 and 57 below display the low flip angle COSY spectra and focus on H2′ to H1′ region of the RF DNA and RF RNA, respectively. In figure 55 highlighted residues include C8 while figure 56 omits the C8 highlight because C8H1′ jumps far downfield and out of the region.

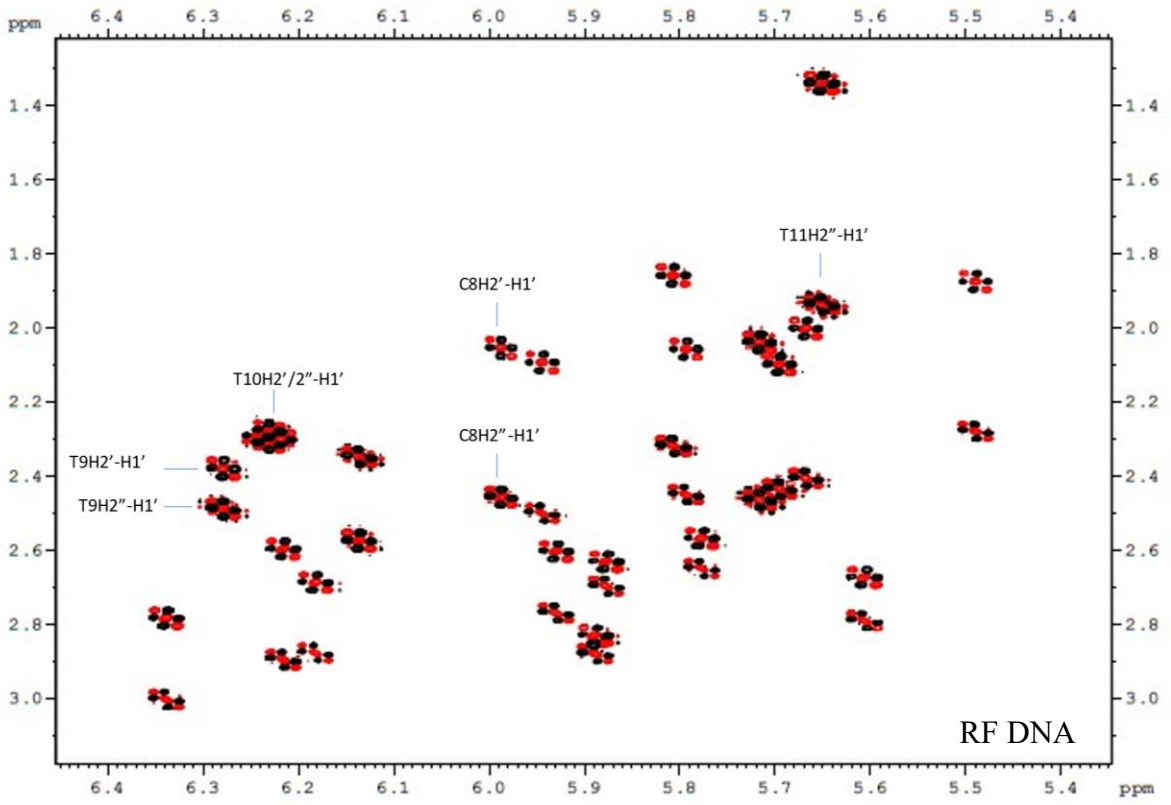


Figure 56: RF DNA low-flip angle COSY spectrum, H2'/ H2'' to H1' region.

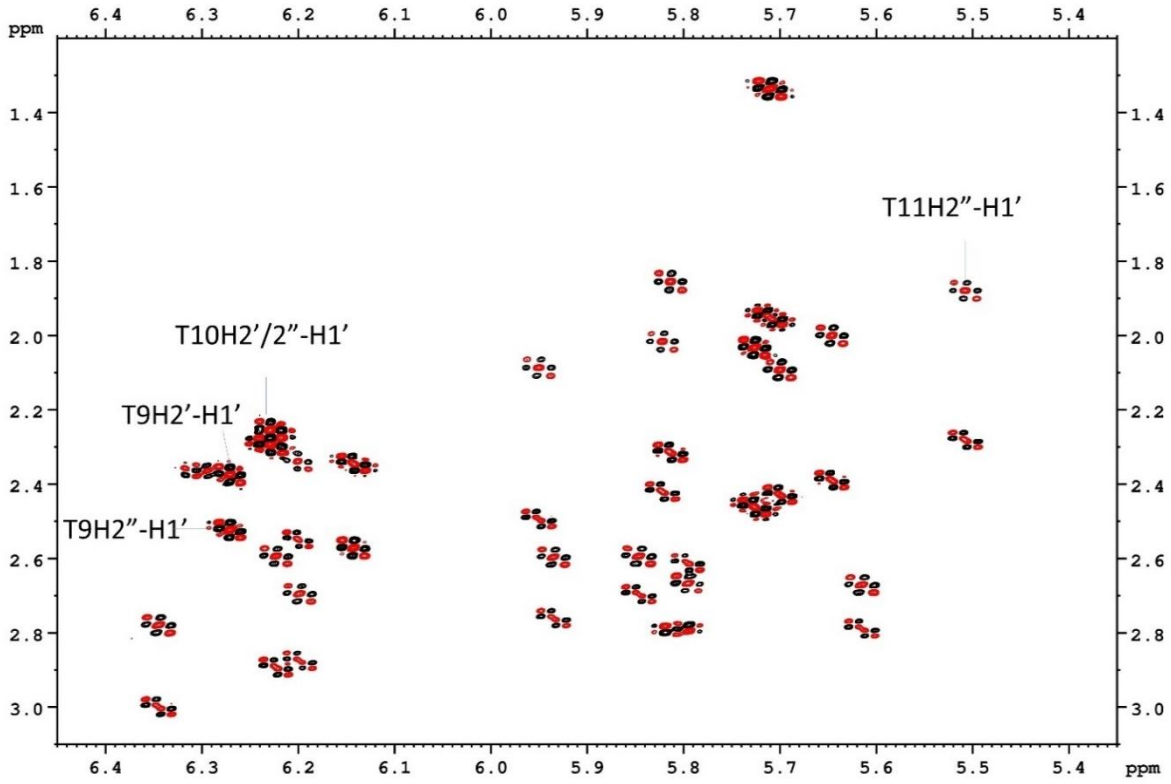


Figure 57: RF RNA low-flip angle COSY spectrum, H2'/ H2'' to H1' region.

After strip transformation, the coupling constants can be measured from the COSY peaks. The fraction south is calculated using the graphical method of correlating the sum of the coupling constants to the percentage of southern confirmation. Figure 58 is an example J coupling sum measurement. The peak highlighted in figure 54 is the DNA C8H2′-H1′. The axis are in Hz to show relevance in J coupling measurement. Since this measurement is for the H2′ peak the measurement is made in frequency 1. Therefore, in the example ∑ J coupling is 27.28 Hz.

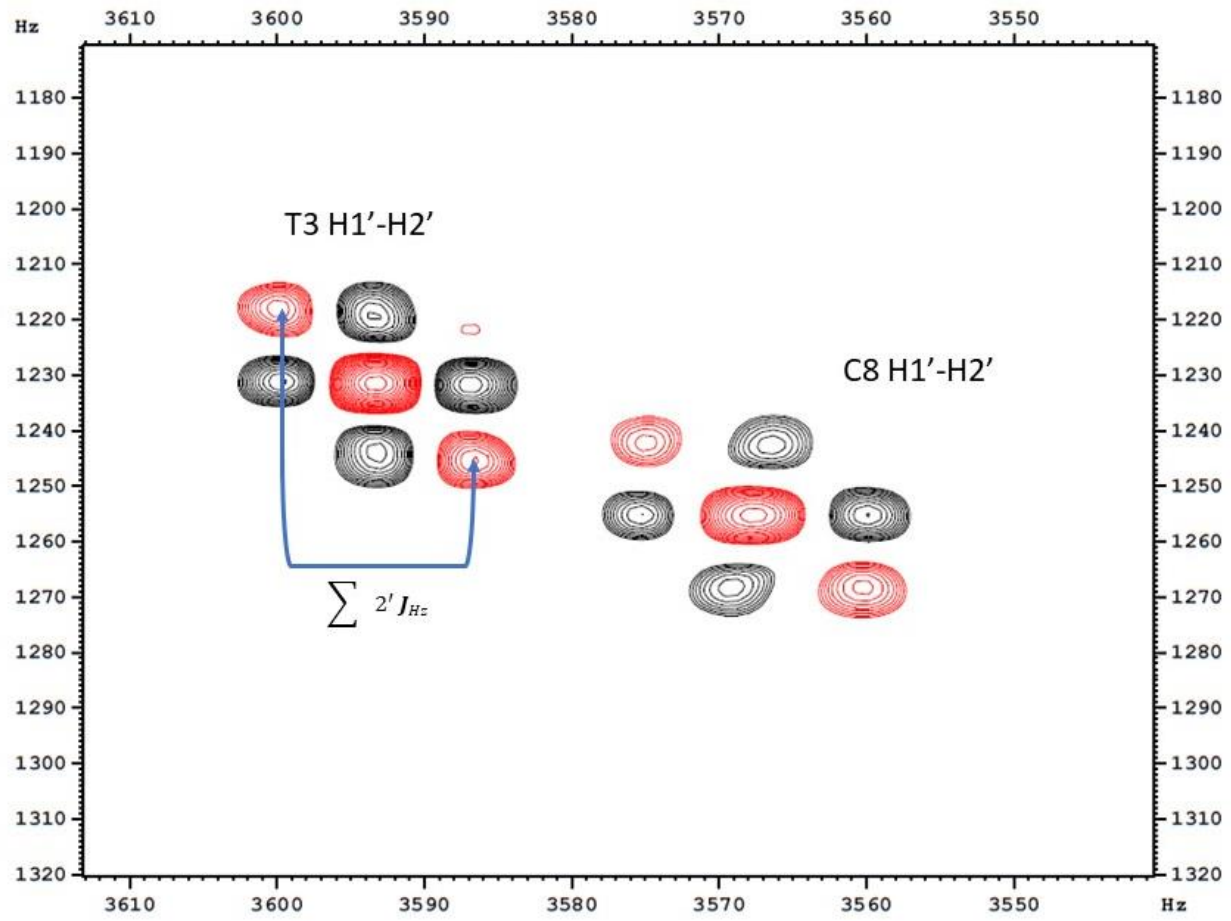


Figure 58: Example measurement of the T3 H1'-H2' ^{3}J coupling constant after strip transformation.

All of the J coupling measurements for H2′/H2″ are collected in tables 15 and 16 below. Table 15 is the collection of J coupling measurements for the RF DNA and table 16 is the collection of measurements for the RF RNA. The measurements are made above and below the diagonal to ensure the best possible accuracy. For instance, the measurement for the ∑ of H1′-H2′ and H1′-H2″ is reliable in the indirect frequency below the diagonal.

Table 15: RF DNA H1'-H2'/H2'' ^{3}J coupling measurements. Values highlighted green are used in fraction south calculations due to their highest inherent accuracy.

	Above Diagonal						Below Diagonal					
	H2″			H2′			H2″			H2′		
	ω2	ω2	ω1	ω2	ω2	ω1	ω2	ω1	ω2	ω2	ω1	ω2
	JH1'-H2"	JH1'-H2'	∑	JH1'-H2"	JH1'-H2'	∑	JH1'-H2"	JH1'-H2'	∑	JH1'-H2"	JH1'-H2'	∑
C1	5.98	**8.01**	22.94	**5.98**	6.80	24.37	6.89	7.43	**20.76**	6.5	11.8	**25.3**
A2	6.07	**9.47**	23.92	**5.92**	8.32	24.95	6.57	9.16	**21.07**	6.61	10.01	**23.2**
T3	6.26	**9.63**	23.40	**6.09**	8.47	26.38	6.515	9.567	**20.73**	7.26	15.54	**29.75**
C4	6.09	**8.78**	22.72	**6.01**	8	24.04	8.24	8.71	**20.93**	6.98	12.98	**24.78**
G5	6.42	**9.26**	23.66	**6.11**	8.31	24.38	6.81	9.26	**21.13**	6.16	10.43	**23.07**
T6	6.19	**9.25**	23.70	**5.99**	8.02	25.84	6.88	8.88	**21.02**	7.12	13.82	**25.37**
G7	6.30	**9.44**	23.65	**5.72**	8.18	24.32	6.62	9.08	**20.57**	5.53	10.43	**22.95**
C8	6.67	**6.53**	24.22	**6.90**	6.52	27.28	9.08	6.28	**23.3**	6.94	14.12	**28.25**
T9	6.47	**7.96**	24.17	**6.60**	7.21	25.92	7.748	7.456	**22.11**	7.19	13.2	**27.57**
T10												
T11	6.17	**9.29**	23.08	**6.13**	8.24	24.02	6.41	9.177	**20.90**	5.71	10.21	**23.55**
A12												
G13	6.40	**10.04**	22.95	**5.28**	9.31	23.49	6.42	9.59	**20.42**	6.38	10.34	**23.15**
C14	6.22	**8.66**	22.91	**6.10**	7.48	24.92	6.67	8.93	**20.77**	6.6	12.59	**25.34**
A15	6.47	**9.45**	24.41	**5.98**	8.44	23.78	7.27	9.40	**21.89**	6.34	10.25	**23.13**
C16	6.33	**8.77**	22.57	**6.37**	7.76	26.35	7.32	8.69	**20.75**	6.84	14.82	**28.31**
G17	6.22	**9.92**	23.22	**5.64**	8.96	24.22	6.4	9.82	**21.2**	6.06	10.41	**22.97**
A18	6.17	**8.82**	23.77	**6.25**	7.87	24.79	6.65	8.61	**21.39**	6.3	11.09	**23.37**
T19	6.26	**8.82**	23.69	**6.70**	7.60	27.07	7	8.6	**21.6**	8.05	15.39	**30.89**
G20	7.06	**6.46**	24.60	**8.42**	6.17	23.30	12.1	7.28	**14.9**	8.28	7.27	**20.85**

Table 16: RF RNA H1'-H2'/H2'' ^{3}J coupling measurements. Values highlighted green are used in fraction south calculations due to their highest inherent accuracy. C8 H2'' values are absent.

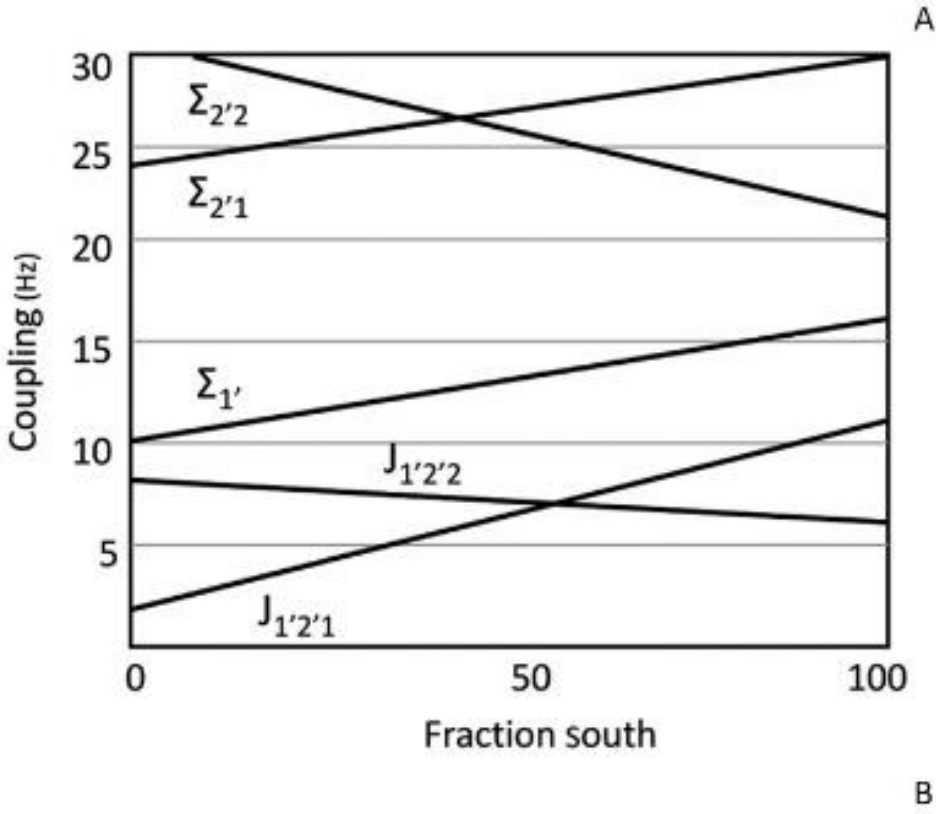
	Above Diagonal						Below Diagonal					
	H2″			H2′			H2″			H2′		
	ω2	ω2	ω1	ω2	ω2	ω1	ω2	ω1	ω2	ω2	ω1	ω2
	JH1'-H2''	JH1'-H2'	Σ	JH1'-H2''	JH1'-H2'	Σ	JH1'-H2''	JH1'-H2'	Σ	JH1'-H2''	JH1'-H2'	Σ
C1	5.88	7.98	23.42	7.00	6.07	24.55	6.96	7.41	20.91	6.30	12.56	25.44
A2	6.24	9.24	24.08	6.08	8.11	24.29	6.65	9.32	21.08	6.66	12.06	23.34
T3	6.20	9.63	24.23	6.02	8.68	26.67	6.91	9.15	21.37	7.23	15.90	27.45
C4	5.98	8.66	23.05	5.77	6.67	24.52	7.01	8.62	20.77	5.89	12.55	24.75
G5	6.24	9.07	23.39	6.46	8.03	24.05	6.60	9.44	21.00	6.20	11.11	22.74
T6	6.22	9.69	23.85	5.91	8.16	25.58	7.09	8.95	21.39	6.73	14.32	26.20
G7	8.29	5.50	23.23	9.03	9.06	23.01	11.33	6.53	22.52	8.76	7.11	20.53
(r)C8				5.59	5.66	16.92				8.20	6.49	13.08
T9	6.41	6.48	23.47	6.42	6.47	24.87	13.28	5.77	22.87	7.65	12.95	25.86
T10												
T11	5.85	9.20	22.78	5.88	7.61	24.61	6.31	9.09	21.01	6.40	12.37	23.77
A12												
G13	6.14	9.82	23.14	5.36	9.06	24.78	6.31	10.17	20.58	6.31	10.99	23.33
C14	5.07	9.16	22.98	6.26	7.48	25.45	6.84	6.77	21.03	6.31	12.64	25.04
A15	6.42	9.28	24.61	6.14	8.39	24.61	7.04	10.27	21.19	6.71	11.26	23.27
C16	6.06	9.01	23.41	6.49	7.88	26.29	6.72	8.40	20.77	6.93	13.27	27.02
G17	6.26	9.96	23.27	5.70	8.87	24.30	6.65	9.99	20.64	6.49	10.98	22.72
A18	6.20	9.10	23.87	6.02	7.94	24.54	6.97	8.46	21.15	6.56	12.25	23.67
T19	6.37	8.72	23.26	6.72	7.51	26.90	7.20	8.77	21.47	6.22	13.88	24.50
G20	7.08	6.37	25.00	8.21	6.23	22.60	13.35	6.80	25.18	7.90	7.30	20.80

Table 17 lists the distances in Angstroms between each sugar proton pair as modeled using Spartan. The rightmost column described the change is protons distances from southern to northern conformation. With a pure northern species, we expect to see very strong cross peaks between the H2″-H4′. While pure northern sugars will not exhibit as much of a change in H2′-H4′ NOE intensity. The delta values are shaded red for the most drastic change. Unfortunately, the resonance overlap is also a problem in the NOESY spectrum even though the cross-peaks are better defined. We therefore visually compare the available H1′-H2′/H2″ COSY cross-peaks. This showed that for residue T10 the cross-peaks are very similar which suggests that the sugar puckering for T10 in either the DNA or RNA sample is also similar.

Table 17: Sugar proton distance changes from pure 2 ' endo (southern) to pure 3 ' endo (northern) conformations as modeled in Discovery studio suite. Delta values highlighted red indicate the greatest change, and values highlighted green indicate the neighboring proton with the least change used as built-in control.

2' endo Distance		3' endo Distance		Δ
1'-2"	**2.372**	1'-2"	**2.377**	**0.00**
1'-2'	**3.032**	1'-2'	**2.788**	**0.24**
1'-3'	**3.933**	1'-3'	**3.846**	**0.09**
1'-4'	**3.646**	1'-4'	**3.381**	**0.27**
2"-3'	**2.703**	2"-3'	**3.061**	**0.36**
2'-3'	**2.401**	2'-3'	**2.421**	**0.02**
2"-4'	**4.095**	2"-4'	**2.782**	**1.31**
2'-4'	**3.863**	2'-4'	**3.812**	**0.05**
3'-4'	**2.648**	3'-4'	**3.017**	**0.37**

The fractions south relate linearly to the respective J coupling for each proton. Therefore, each fraction south can be calculated by deriving a linear equation graphically. The Y intercept is measured and used in the linear equation for fractions south. Figure 59 (A) is the graph of J coupling trends in relation to their southern characteristics. Figure 59 (B) is the linear equation used in southern calculations.



$$\Sigma 1' fs = \frac{JH1'H2' + JH1'H2'' - 9.8}{0.059} \qquad fs = \frac{JH1'H2' - 1.8}{0.084}$$

$$x\,fs = \frac{fsH1'H2' + fs\Sigma H1'}{2}$$

Figure 59: Derived fraction south values and linearly related coupling. Graphical method adopted from 'NMR of Macromolecules'.[26][28]

Table 18 contains the calculated fraction south for each residue using the above graphical method. The three residues with the greatest delta, indicating the most northern characteristics, are highlighted. Unfortunately, resonance overlaps precluded us from measuring the coupling constant for T10 and A12. We, therefore, attempted to get some insight from comparing NOE

intensities for H2″-H4′ with that of a 100% fraction south residue G13. H2″-H4′ was selected because it exhibits the greatest change in proton distance from southern to northern conformations. Northern characteristics for residues T10 and A12 were obtained by comparing NOE intensities for H2″-H4′ with that of a 100% fraction south residue G13. H2″-H4′ was observed because it exhibits the greatest change in proton distance during its transition from southern to northern.

Table 18: Calculated fraction south values and the delta from DNA to RNA.

	fs		
	DNA	RNA	Δ
C1	78	77	1
A2	92	89	4
T3	93	93	1
C4	85	85	0
G5	89	84	5
T6	90	94	5
G7	94	60	34
C8	56	44	12
T9	72	59	13
T10			
T11	89	89	0
A12			
G13	100	100	0
C14	83	87	3
A15	92	89	3
C16	82	83	1
G17	99	99	0
A18	83	88	5
T19	81	80	2
G20	45	41	3

Figure 60 is a cartoon of the RF analog and comprises all the accumulated chemical shift and fractions south delta values. Normal B conformation DNA has a high fraction south characteristic, therefore a high $f_s \Delta$ in the figure indicates a greater northern character. The fraction south delta is present in the figure 60 sugar ring interiors. Evident is the great change in fraction south character among residue G7 at a delta of 34. T9 has a fraction south of 59 in the riboadulterrated RF analog but only a delta of 13 from the DNA control sample. Most evident is the change in fraction south character for residue G7 at a delta of 34%.

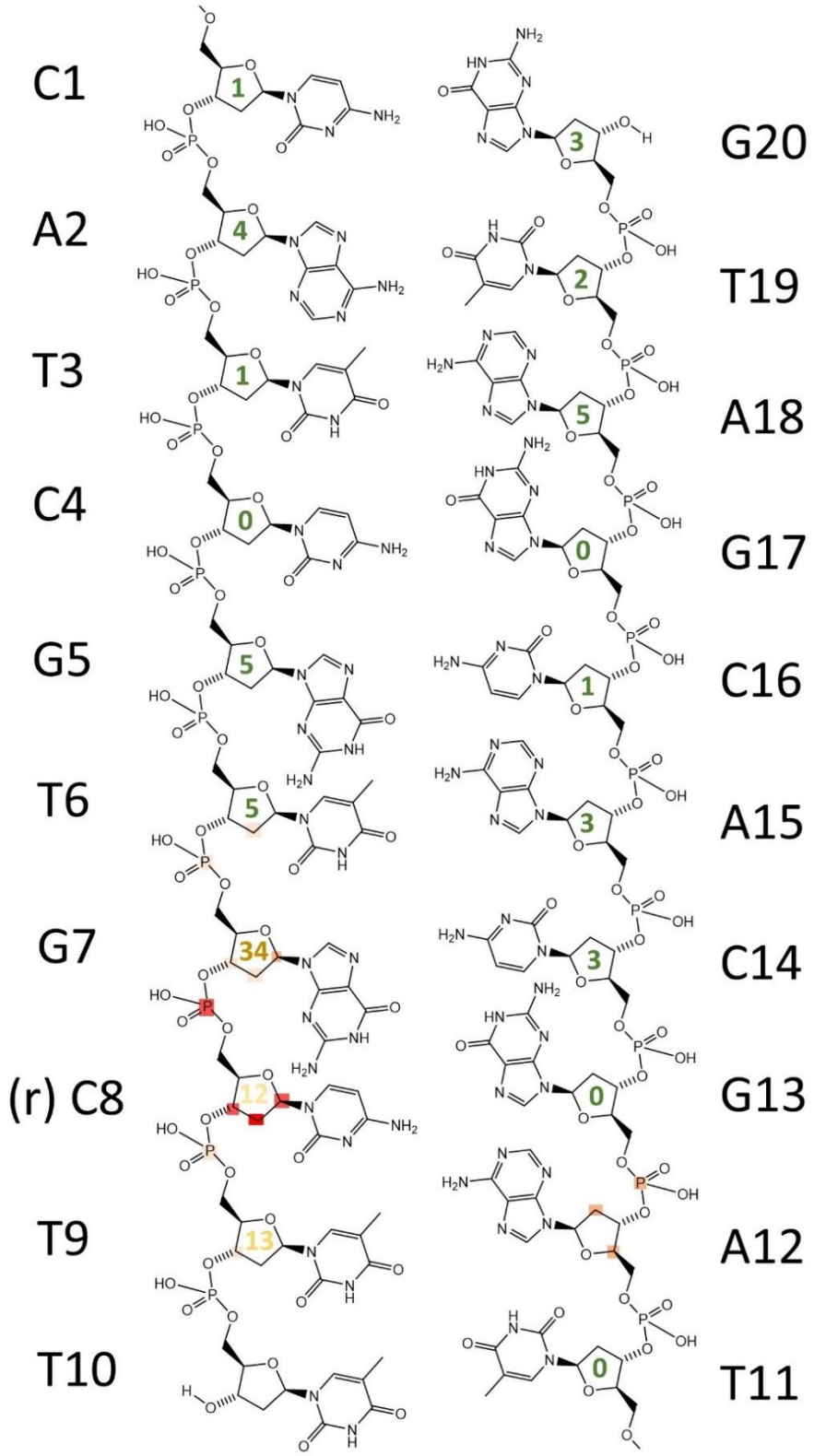


Figure 60: Composite figure containing all perturbation detected in the RF analog oligonucleotide substrates. Perturbation includes chemical shifts and sugar fraction south deltas.

4.2.4 Discussion: High S conformation is disrupted

Our RF analog is properly base paired with all residues excluding thymine mismatched pair. It is also strongly exhibiting overall B structure with the exception of the residues neighboring (r)C8. Structural changes of the ribo damaged DNA are localized to the oxygen containing sugar and nearby, 5′ phosphate group. We have previously shown the phosphate damage can be localized to either the phosphate group 5′ or 3′ of the ribo damage. [7]

In the RF analog changes are observed 5′ of the ribo C, while the 3′ side is still affected to a lesser degree. In fact, the major chemical shift differences are located on the ribo C sugar, the immediate phosphates. Minor chemical shift changes were observed on the same strand as far back as T6 and across strand in residue A12. Sugar puckering changes are only observed in residues G7, (r)C8, and T9. While the G7 exhibits the greatest fraction south delta, as expected (r)C8 exhibits the most northern characteristics. The neighboring nucleotides of our riboadulterrated residue impact the localization of the damage such that a different nucleotide sequence will impact the resonances and structural perturbation in a different manner. As was observed by our RNase HII experiments in chapter 1, a 5′ purine and 3′ pyrimidine reacted much more quickly with the enzyme.

Future research includes determining the structure from distance information acquired by integrating all assigned NOESY cross-peaks. The Mardigras and Corma suite, which takes spin diffusion into account, will be used to calculate the inter-proton distances from the NOESY volume data if we use longer mixing times. Alternatively, a simple distance binning approach can be employed to generate distance restraints as strong (1.8 to 3.6 Å), medium (2.6 to 5.0) and weak (3.5 to 6.5 Å) [27]. In that case, NOESY's with mixing times of 50-75 ms will be used as

spin diffusion is minimal compared to longer mixing times. This method is less labor-intensive than a full relaxation matrix method while providing good-quality structures.

The distance restraints along with base pair-, sugar puckering-, and backbone restraints will be used to determine the solution NMR structure by restrained MD simulations using Assisted Model Building with Energy Refinement (AMBER). The rMD simulations of the DNA duplexes will be run in an octahedral box, solvated, and neutralized with Na^+ and additional ions.

These restrained MD calculations, including sugar puckering and backbone restraints, can be repeated for every iteration of ribonucleotide inclusion (XxX). This will expound on our understanding of the neighboring nucleotide environmental effects on a single ribonucleotide incorporation at a replication fork.

REFERENCES

1- Marteijn, J. A., Lans, H., Vermeulen, W., & Hoeijmakers, J. H. (2014). Understanding nucleotide excision repair and its roles in cancer and ageing. *Nature Reviews Molecular Cell Biology*, *15*(7), 465–481. https://doi.org/10.1038/nrm3822

2- Jackson, S. P., & Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature*, *461*(7267), 1071–1078. https://doi.org/10.1038/nature08467

3- De Bont, R. (2004). Endogenous DNA damage in humans: A review of Quantitative Data. *Mutagenesis*, *19*(3), 169–185. https://doi.org/10.1093/mutage/geh025

4- Chiu, H.-C., Koh, K. D., Evich, M., Lesiak, A. L., Germann, M. W., Bongiorno, A., Riedo, E., & Storici, F. (2014). RNA intrusions change DNA elastic properties and structure. *Nanoscale*, *6*(17), 10009–10017. https://doi.org/10.1039/c4nr01794c

5- Cerritelli, S. M., & Crouch, R. J. (2016). The Balancing Act of ribonucleotides in DNA. *Trends in Biochemical Sciences*, *41*(5), 434–445. https://doi.org/10.1016/j.tibs.2016.02.005

6- Cronan, G. E., Kouzminova, E. A., & Kuzminov, A. (2019). Near-continuously synthesized leading strands in escherichia coli are broken by ribonucleotide excision. Proceedings of the National Academy of Sciences, 116(4), 1251–1260. https://doi.org/10.1073/pnas.1814512116

7- Evich, M., Spring-Connell, A. M., Storici, F., & Germann, M. W. (2016). Structural impact of single ribonucleotide residues in DNA. *ChemBioChem*, *17*(20), 1968–1977. https://doi.org/10.1002/cbic.201600385

8- Chakarov, S., Petkova, R., Russev, G., & Zhelev, N. (2014). DNA damage and mutation. types of DNA damage. *BioDiscovery*, (11), 1. https://doi.org/10.7750/biodiscovery.2014.11.1

9- James, K., Gamba, P., Cockell, S. J., & Zenkin, N. (2016). Misincorporation by RNA polymerase is a major source of transcription pausing *in vivo*. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkw969

10- Chang A., Jeske L., Ulbrich S., Hofmann J., Koblitz J., Schomburg I., Neumann-Schaal M., Jahn D., Schomburg D. BRENDA, the ELIXIR core data resource in 2021: new developments and updates. *Nucleic Acids Research*., 49: D498-D508. DOI: 10.1093/nar/gkaa1025 PubMed: 33211880

11- Chon, H., Sparks, J. L., Rychlik, M., Nowotny, M., Burgers, P. M., Crouch, R. J., & Cerritelli, S. M. (2013). RNase H2 roles in genome integrity revealed by unlinking its activities. *Nucleic Acids Research*, *41*(5), 3130–3143. https://doi.org/10.1093/nar/gkt027

12- Zuker, M. (2003). mFold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, *31*(13), 3406–3415. https://doi.org/10.1093/nar/gkg595

13- Krężel, A., & Bal, W. (2004). A formula for correlating P K A values determined in D_2O and H_2O. *Journal of Inorganic Biochemistry*, *98*(1), 161–166. https://doi.org/10.1016/j.jinorgbio.2003.10.001

14- Lee, W., Tonelli, M., & Markley, J. L. (2014). NMRFAM-Sparky: Enhanced Software for biomolecular NMR spectroscopy. *Bioinformatics*, *31*(8), 1325–1327. https://doi.org/10.1093/bioinformatics/btu830

15- Kojima, K., Baba, M., Tsukiashi, M., Nishimura, T., & Yasukawa, K. (2018). RNA/DNA structures recognized by RNase H2. *Briefings in Functional Genomics*, *18*(3), 169–173. https://doi.org/10.1093/bfgp/ely024

16- Cerritelli, S. M., & Crouch, R. J. (2009). Ribonuclease H: The enzymes in Eukaryotes. *FEBS Journal*, *276*(6), 1494–1505. https://doi.org/10.1111/j.1742-4658.2009.06908.x

17- Rychlik, M. P., Chon, H., Cerritelli, S. M., Klimek, P., Crouch, R. J., & Nowotny, M. (2010). Crystal structures of RNase H2 in complex with nucleic acid reveal the mechanism of RNA-DNA junction recognition and cleavage. *Molecular Cell*, *40*(4), 658–670. https://doi.org/10.1016/j.molcel.2010.11.001

18- Yang, W., Lee, J. Y., & Nowotny, M. (2006). Making and breaking nucleic acids: Two-Mg^{2+}-ion catalysis and substrate specificity. *Molecular Cell*, *22*(1), 5–13. https://doi.org/10.1016/j.molcel.2006.03.013

19- Rychlik, M. P., & Nowotny, M. (2013). *T. Maritima* RNase H2 G21s in complex with nucleic acid substrate and calcium ions. https://doi.org/10.2210/pdb4hht/pdb

20- Dobosy, J. R., Rose, S. D., Beltz, K. R., Rupp, S. M., Powers, K. M., Behlke, M. A., & Walder, J. A. (2011). RNase H-dependent PCR (RHPCR): Improved specificity and single nucleotide polymorphism detection using blocked cleavable primers. *BMC Biotechnology*, *11*(1). https://doi.org/10.1186/1472-6750-11-80

21- Huttner, D., & Hickson, I. D. (2013). Helicases. *Brenner's Encyclopedia of Genetics*, 406–408. https://doi.org/10.1016/b978-0-12-374984-0.00687-2

22- Pike A. C., Gomathinayagam S., Swuec P., Berti M., Zhang Y., Schnecke C., Marino F., von Delft F., Renault L., Costa A., Gileadi O., Vindigni A. (2015). Human RECQ1 helicase-driven DNA unwinding, annealing, and branch migration: Insights from DNA complex structures. *Proceedings of the National Academy of Sciences of the United States of America*. https://pubmed.ncbi.nlm.nih.gov/25831490/

23- Brosh, R. M. (2013). DNA helicases involved in DNA repair and their roles in cancer. *Nature Reviews Cancer*, 13(8), 542–558. https://doi.org/10.1038/nrc3560

24- Mentegari, E., Crespan, E., Bavagnoli, L., Kissova, M., Bertoletti, F., Sabbioneda, S., Imhof, R., Sturla, S. J., Nilforoushan, A., Hübscher, U., van Loon, B., & Maga, G. (2016). Ribonucleotide incorporation by human DNA polymerase η impacts translesion synthesis and RNase H2 activity. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkw1275

25- Harika, N. K., Paul, A., Stroeva, E., Chai, Y., Boykin, D. W., Germann, M. W., & Wilson, W. D. (2016). Imino proton NMR guides the reprogramming of a•t specific minor groove binders for mixed base pair recognition. *Nucleic Acids Research*, 44(10), 4519–4527. https://doi.org/10.1093/nar/gkw353

26- Evich, M., Spring-Connell, A. M., & Germann, M. W. (2017). Impact of modified ribose sugars on nucleic acid conformation and function. *Heterocyclic Communications*, 23(3), 155–165. https://doi.org/10.1515/hc-2017-0056

27- Spring-Connell, A. M., Evich, M., & Germann, M. W. (2018). NMR structure determination for oligonucleotides. *Current Protocols in Nucleic Acid Chemistry*, *72*(1). https://doi.org/10.1002/cpnc.48

28- Wijmenga, S. S., & Mooren, M. M. W. (1993). In G. C. K. Roberts (Ed.), *NMR OF MACROMOLECULES: A practical approach* (pp. 217–288). essay, IRL Press at Oxford University Press.

APPENDICES

Appendix A

Appendix A.1

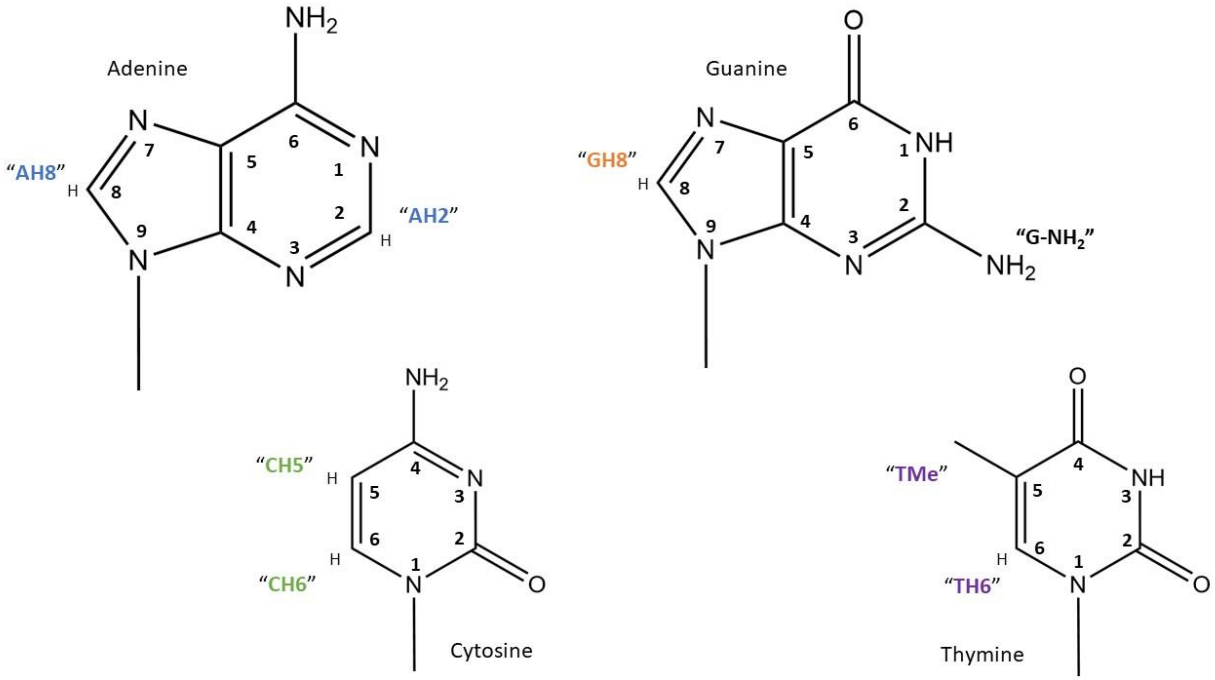


Figure 61: Numbering and labeling convention of DNA base protons