Targeting Holliday Junctions

Christopher Hamilton
chamilton6@gsu.edu

Follow this and additional works at: https://scholarworks.gsu.edu/chemistry_theses

Recommended Citation
doi: https://doi.org/10.57709/5813136

This Thesis is brought to you for free and open access by the Department of Chemistry at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Chemistry Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
TARGETING HOLLIDAY JUNCTIONS

by

CHRISTOPHER HAMILTON

Under the Direction of Dr. Markus Germann

ABSTRACT

Holliday junctions are formed as an intermediate during DNA recombination as the two strands come together. Recombination occurs during meiosis, and also during DNA double strand repair. Trapping this branched intermediate could prevent DNA repair from occurring in cells which would prove beneficial during cancer treatment. There are many enzymes that cleave Holliday junctions. One such enzyme, T7 Endonuclease I, was specifically chosen to detect ligand binding at the core of the junction since its binding and cleavage of cruciforms is well documented. Specialized bifunctional ligands were studied in this project that were designed to bind DNA structures that are held in close proximity to one another. These compounds have two identical binding modules that are connected by a linker of various length and rigidity, with each module binding very weakly; however, when both modules bind the binding affinity is greatly enhanced. The interactions of these compounds with cruciforms are currently being studied.

INDEX WORDS: Cruciform binding ligand, Holliday junction, Homologous recombination
TARGETING THE CORE OF A HOLLIDAY JUNCTION

by

CHRISTOPHER HAMILTON

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

2014
TARGETING THE CORE OF A HOLLIDAY JUNCTION

by

CHRISTOPHER HAMILTON

Committee Chair: Markus Germann

Committee: David Wilson
           Kathryn Grant

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
August 2014
ACKNOWLEDGEMENTS

I would like to thank Dr. Germann for his generous support with this project. Without him none of this would be possible. I would like to thank my lab mates for their assistance, especially Alexander Spring who has been of invaluable support throughout every step of this project and throughout my time in the lab. I would also like to thank Marina Evich, Ekaterina Stroeva, and Jin Zhang for their advice and time when assisting me with research issues and in preparation for my various presentations. Lastly I would like to extend my thanks to all my friends, family, and loved ones for their support throughout this project. Thank you.
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS** .................................................................................................................. iv

**LIST OF TABLES** ............................................................................................................................. viii

**LIST OF FIGURES** ............................................................................................................................ ix

**LIST OF ABBREVIATIONS** ............................................................................................................... xi

**LIST OF BUFFERS** ............................................................................................................................. xii

1 **INTRODUCTION** ............................................................................................................................. 1

1.1 Double Strand Break Repair ........................................................................................................ 1

1.2 Cruciform Interactions with Proteins ............................................................................................ 2

   1.2.1 Werner’s Syndrome ................................................................................................................... 2

   1.2.2 Hexameric Peptide Inhibitory Effect on Junction Processing Enzyme ......................... 2

   1.2.3 Alternative Lengthening of Telomeres .................................................................................... 3

1.3 Holliday Junctions ............................................................................................................................ 3

   1.3.1 Holliday Junction Dynamics .................................................................................................... 5

1.4 T7 Endonuclease I ........................................................................................................................... 7

1.5 Ligands of Interest ........................................................................................................................... 7

2 **EXPERIMENTAL** ............................................................................................................................ 9

2.1 Conformational Dynamics ............................................................................................................. 9

   2.1.1 Conformational Effects on Circular Dichroism ................................................................. 9

   2.1.2 Conformational Effects on Gel Mobility ............................................................................. 9

2.2 Ligand Binding to Cruciforms and Branched DNA ..................................................................... 10
2.2.1 T7 Endonuclease I Cleavage of Cruciform DNA ........................................... 10

2.2.2 Ligand Binding Determined by Its Effect on $T_m$ ........................................... 10

2.3 NMR Experiments ..................................................................................................... 11

2.3.1 Temperature Series ............................................................................................... 11

2.3.2 1D NOE ................................................................................................................. 11

2.3.3 2D NOESY ............................................................................................................. 11

2.3.4 T1 Experiments ..................................................................................................... 11

2.3.5 Ligand Titrations ................................................................................................... 12

3 RESULTS ....................................................................................................................... 13

3.1 Branched DNA Substrates ....................................................................................... 13

3.1.1 Cruciform Substrates ............................................................................................ 13

3.1.2 Partial Junctions ........................................................................................................ 16

3.1.3 Y-Structure DNA .................................................................................................... 18

3.2 Conformational Dynamics ....................................................................................... 19

3.2.1 Conformational Effects on Circular Dichroism ....................................................... 19

3.2.2 Conformational Effects on Gel Mobility ................................................................. 24

3.3 Ligand Interaction with Branched DNA ................................................................. 25

3.3.1 Effect of Ligand Binding on the $T_m$ of a Cruciform ............................................. 25

3.3.2 Effect of Ligand Binding by CD Analysis ............................................................... 26

3.3.3 T7 Endonuclease I Cleavage and its Inhibition by Ligands ..................................... 29
LIST OF TABLES

Table 3.1- Ligand Binding Ranking to Cruciform DNA......................................................... 30
Table 3.2- End Modification Effect on Chemical Shift............................................................. 45
Table 3.3- Table of NOE Intensities....................................................................................... 47
Table 3.4- List of T1 values for the Adenine Non-Exchangeable Protons............................... 47
LIST OF FIGURES

Figure 1.1 - X-ray Crystal Structure of LH3.123 Bound to the Core of a HJ.................................................. 4
Figure 1.2 Branch Migration of a Homologous Junction............................................................................. 5
Figure 1.3 Transitions Between Cruciform Conformations........................................................................ 6
Figure 1.4 - Cruciform Binding Ligands...................................................................................................... 8
Figure 3.1 - The Cleavage of Junction3 by T7 Endonuclease........................................................................ 15
Figure 3.2 - Schematic of Cruci4S Folded into a Holliday Junction............................................................. 16
Figure 3.3- Model of the Partial Junctions, JunctionA and JunctionB...................................................... 17
Figure 3.4- Model of the Partial Junctions with Modified Thymidine Tails............................................. 18
Figure 3.5- Model of the Y-Structure DNA............................................................................................... 19
Figure 3.6- Open-X to Stacked-X Conformational Switch......................................................................... 20
Figure 3.7- Effect of MgCl₂ on the CD spectra of Cruci3HL .................................................................. 21
Figure 3.8- Effect of MgCl₂ on the CD spectra of Cruci3HL-Hairpin-Loop .............................................. 22
Figure 3.9- Effect of MgCl₂ on the CD spectra of Junction3..................................................................... 23
Figure 3.10- Conformation Dynamics of Cruci3HL Compared to Single Stranded DNA.................... 25
Figure 3.11- Conformational Dynamics of Cruci3HL compared to Junction3........................................... 25
Figure 3.12- Effect of LH3.123 on the Tₘ of Cruci3HL ............................................................................ 26
Figure 3.13- CD Spectra of LH3.123 Titrated into Cruci3HL no NaCl.................................................... 27
Figure 3.14- CD Spectra of LH3.123 Titrated into Cruci3HL with NaCl.................................................. 28
Figure 3.15- CD Spectra of LH3.123 Titrated into Cruci3HL-Hairpin-Loop no NaCl.......................... 28
Figure 3.16- J3B Inhibition of T7 Endo I Cleavage by Interaction with Ligands on Cruci3HL............. 29
Figure 3.17- Inhibition of T7 Endo I Cleavage by Interaction with Ligands on Junction3 and J3B-Poly-T ........................................................................................................................................... 30
Figure 3.18- T7 Endo I Cleavage of Partial Junctions................................................................................. 32
Figure 3.19- Inhibition of T7 by Ligands Binding to DNA Partial Junctions

Figure 3.20- Y-Structure Stability Gel

Figure 3.21- NMR Assignment of J3B-HP

Figure 3.22- NMR Assignment of J3B-Poly-T

Figure 3.23- T1 Experiment to Identify the Adenosine H2

Figure 3.24- Watson-Crick Base Pair Compared to a Hoogsteen Base Pair

Figure 3.25- J3B-HP-3’-Overhang Assignment

Figure 3.26- J3B-HP-5’-Overhang Assignment

Figure 3.27- J3B-HP-TT-Mismatch Assignment

Figure 3.28- J3B-Hairpin Temperature Study

Figure 3.29- J3B-Poly-T Temperature Study

Figure 3.30- J3B-HP-5’-Overhang Temperature Study

Figure 3.31- J3B-HP-3’-Overhang Temperature Study

Figure 3.32- J3B-HP-T-T Mismatch Temperature Study

Figure 3.33- Model of a Cruciform Specific Ligand

Figure 3.34- X-ray Crystal Structure of a Stacked-X Holliday Junction

Figure 3.35- X-ray Crystal Structure of an Open-X Holliday Junction
LIST OF ABBREVIATIONS

Holliday junction (HJ)
Double Holliday junction (dHJ)
Homologous recombination (HR)
Nuclear Magnetic Resonance (NMR)
Circular Dichroism (CD)
Melting Temperature (T_m)
Electrospray Ionization Mass Spectrometry (ESI-MS)
Fast Protein Liquid Chromatography (FPLC)
Polyacrylamide Gel Electrophoresis (PAGE)
Ethylenediaminetetraacetic acid (EDTA)
Dimethyl Sulfoxide (DMSO)
Deuterium Oxide (D_2O)
4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS)
Tris(hydroxymethyl)aminomethane (Tris)
Ammonium Persulfate (APS)
Tetramethylethylenediamine (TEMED)
T7 Endonuclease I (T7 Endo I)
Dithiothreitol (DTT)
Sodium Phosphate (NaP)
LIST OF BUFFERS

12% PAGE Native (MgCl₂ or EDTA): 12% acrylamide, 89 mM Tris, 89 mM Boric acid, 2 mM MgCl₂ OR 2 mM EDTA, 0.05% APS, 0.05% TEMED.

12% PAGE Denaturing (MgCl₂ or EDTA): 12% acrylamide, 8 M urea, 89 mM Tris, 89 mM Boric acid, 2 mM MgCl₂ OR 2 mM EDTA, 0.05% APS, 0.05% TEMED.

NEB Buffer 1x: 50 mM NaCl 10 mM Tris-HCl 10 mM MgCl₂ 1 mM DTT pH 7.9 @ 25 °C
1 INTRODUCTION

1.1 Double Strand Break Repair

DNA damage is a common cellular occurrence which can arise from internal factors such as the oxidation of nucleotides by reactive oxygen species (ROS) produced during metabolism as well as from external factors such as irradiation and reactive chemicals. Double strand breaks are a particularly dangerous form of DNA damage that if left unrepaired can lead to cell death or cause chromosomal translocation. Fortunately several mechanisms exist to repair this damage. Two main pathways exist to repair this damage. Non-Homologous End Joining (NHEJ) and Homologous Recombination (HR) can then be broken up into additional subpathways. NHEJ repairs damage by ligating the damaged DNA strands together without the aid of a template; its accuracy is variable, if the 3’ and 5’ overhangs are complementary the repair will proceed error free. If the overhangs require alignment, small deletions or insertions can be introduced into the recombinant DNA due to the lack of a template[1]. Homologous recombination can be achieved by different subpathways depending on the cell type and stage of the cell cycle. HR is the primary pathway for repairing DSBs caused by replication fork collapse[2]. These subpathways all share the same starting steps of forming extended single strand regions which are stabilized by a consortium of enzymes, followed by homology search and strand invasion which forms a displacement loop (D-loop) where DNA synthesis can occur. The D-loop can form a replication fork if it is missing a second end in the process known as Break-induced replication (BIR). This pathway is potentially mutagenic. The DHJ pathway is found most often in meiosis but there is evidence which shows it is formed ~10% of the time in mitotic cells as well[3].
1.2 Cruciform Interactions with Proteins

Cruciforms are targeted by a large range of proteins, from simple peptides to highly regulated enzymes. The dysregulation of these interactions may lead to a variety of diseases and cellular dysfunction. For example, the proto-onco gene DEK, specifically targets cruciform structures over duplex DNA\(^4\); when over-expressed, DEK prevents cell death which can ultimately lead to cancer\(^5\). The mechanism by which DEK bypasses cell death is not well known.

1.2.1 Werner's Syndrome

Werner’s syndrome is an inheritable disease that causes premature aging and is associated with rare types of cancer by mutating the protein WRN which is capable of binding cruciform DNA\(^6\). Werner’s syndrome causes a host of problems ranging from genome instability and reduced repair, which lead to cancer while at the same time causing rapid telomere shortening responsible for premature aging\(^7\). The protein WRN has both a helicase portion and an exonuclease which targets DNA secondary structures such as Holliday junctions or D-loops with a higher affinity than for duplex DNA\(^8\). WRN helicase activity is necessary to unwind the before mentioned secondary structures to perform recombination and DNA repair. The exonuclease also interacts specifically with 4-way junctions and is required to resolve HJ secondary structure. Werner’s syndrome mutates the WRN protein preventing it from performing these tasks yet the protein can still bind HJ’s.

1.2.2 Hexameric Peptide Inhibitory Effect on Junction Processing Enzyme

The hexapeptide WRWYCR is capable of inhibiting junction resolving enzymes and helicases found in bacteria\(^9\). This article claims that this is achieved by binding to the core of an Open-X junction. Their evidence for this is that the addition of spermidine and/or Mg\(^{2+}\) at low concentrations, 2mM and 80µM respectively, causes a marked decrease of 10-100x apparent binding affinity of the hexapeptide. Both of these ions are known to bias the Stacked-X conformation of cruciform DNA. The peptide
interferes with the resolution of HJ by RuvABC, which normally forms an Open-X complex with a HJ, and prevents helicase activity by RecG helicase\textsuperscript{9}. Both of these enzymes are present in bacteria, therefore their inhibition could be of therapeutic interest.

\textbf{1.2.3 Alternative Lengthening of Telomeres}

Telomere lengthening has long been a hallmark of many types of cancer where the enzyme responsible is activated telomerase. Telomerase repairs and lengthens the telomeric repeats which results in the immortalization of a cell; consequently, many cancer treatments target telomerase. An alternate pathway has been discovered that can lengthen telomeres in the absence of telomerase. This pathway is known as the Alternative Lengthening of Telomeres (ALT). The exact mechanism is still unknown, but there are many models that suggest HR is involved in this process\textsuperscript{10}. Therefore, preventing resolution of HJ structures could prevent ALT active telomere lengthening.

\textbf{1.3 Holliday Junctions}

Holliday junctions, proposed by Robin Holliday, are a mobile four-way junction that connects two DNA duplexes together during homologous recombination. Holliday junctions, also known as cruciform DNA, form in the presence of inverted repeats which are prevalent at promoter regions, replication initiation sequences, and at sites of chromosomal DNA damage\textsuperscript{11-13}. The cruciform structure is stabilized by supercoils that are temporarily formed during DNA replication and transcription\textsuperscript{14}. The HJ intermediate is a potential target for ligands which aim to prevent HR. Only a handful of molecules have been described to bind to the core of a Holliday junction (small peptides,\textsuperscript{9} intercalators,\textsuperscript{15} and LH3.123\textsuperscript{16}). Shown below is an X-ray crystal structure of LH3.123 bound to the core of a HJ (Figure 1.1). LH3.123, a ligand developed by Mark Searcey’s group is displayed in figure 1.4 and described in more detail in section 1.5. It is of great interest to find a ligand with a high affinity to the core of the junction while interacting weakly with double stranded DNA outside of the HJ. Preventing resolution of the HJ
could reduce the amount of DSBs that are repaired. A study has shown that by removing the resolving enzymes in yeast cells causes a 40% decrease in the amount of DSBs that are repaired\textsuperscript{[17]}. Apart from DSBR, HJs remain an important target with applications such as gene therapy. Homologous recombination allows for branch migration of the junction (Figure 1.1), which makes structural studies of this intermediate structure difficult to obtain experimentally. To bypass this obstacle, groups have immobilized the junction by creating synthetic constructs that are incapable of branch migration due to their asymmetry.

![Figure 1.1- X-ray Crystal Structure of LH3.123 Bound to the Core of a HJ.](image)

LH3.123 neatly binds to the core of a HJ in this X-ray crystal structure. It is of great interest to see if this ligand will bind to this location in solution. Image taken from Brogden et al. *Angewandte Chemie*, 2007\textsuperscript{[16]}.
Figure 1.2 - Branch Migration of a Homologous Junction. The homologous junction is free to branch migrate. Doing so does not generate any mismatched base pairs.

1.3.1 Holliday Junction Dynamics

HJs even in an artificially constructed immobile form still display an array of conformational changes. One of these conformational changes is the equilibrium between the Open-X form and the Stacked-X form. These conformations create entirely different topologies at the core of the junction. In the Open-X form the core is wide open with ear of the arms held at nearly 90 degree angles to one another. The Stacked-X form is found primarily under conditions where divalent cations are present\textsuperscript{18}. The Stacked-X form is preferred at physiological ionic conditions. This structure has a much narrower junction and is stabilized by coaxial helical stacking. Apart from the different conformations, different isomeric structures can exist as well; this has been shown by FRET and comparative gel electrophoresis\textsuperscript{19}. The cruciform sequence can be engineered in such a way to bias certain isoforms.

Junction3 has a preference to form the isomer IsoII approximately 80% of the time. Additionally the
isomeric forms can change the stereochemistry of the junction by placing the major groove ends on one side of the junction and the minor groove sides on the other\textsuperscript{[20]}. Junction3 was designed in such a way to prevent branch migration in each of its different conformations; branch migration of its junction guarantees a mismatched base pair. Switching between the isomeric states is believed to require forming the Open-X structure as an intermediate\textsuperscript{[19]}. The terms IsoI and IsoII originate from Declais et al.\textsuperscript{[21]}; IsoII refers to the purple strand stacking on the red strand and the green stacking on the black as shown below (Figure 1.1)

**Figure 1.3- Transitions Between Cruciform Conformations.** In the presence of magnesium the Stacked-X conformation is formed. However, there can be more than one configuration of this junction depending on what arms stack with one another. For the sequence Junction3, the Iso II conformation is favored by \(~4:1\). In order for a conformational change to take place the junction goes through the Open-X intermediate form first.

It may also be possible to bias the formation of Open-X over Stacked-X regardless of salt concentration as shown in the gel mobility study we have performed on our synthetic cruciform which shows that it is unaffected by a change in salt concentrations (Figure 3.5). This could mean that the
sequence by default is already in the Stacked-X conformation or switching from Open-X to Stacked-X displays no change in mobility, which is unlikely).

1.4 T7 Endonuclease I

There are several enzymes that are capable of cleaving HJs. T7 Endo I is a 149 amino acid protein encoded by bacteriophage T7 Endo I. We chose to study T7 Endo I due to its availability and selectivity to bind to HJ’s over duplex DNA. T7 Endo I exhibits low nM affinity to HJ’s while binding duplex DNA ~1000 times less selectively\(^\text{[22]}\). T7 Endo I is also capable of cleaving branched DNA structures even when part of the structure is single stranded. The high selectivity of T7 makes it a useful resolvase for determining if a ligand binds to the HJ. The dimeric enzyme T7 Endo I makes several contacts along the arms of the HJ probing for the aberrant backbone angles present in branched DNA. The enzyme introduces two nicks across the junction which resolves the two duplexes. The N-terminal tail of T7 Endo I is believed to be responsible for opening the core of the junction. When it is removed the junction isn’t able to be fully resolved\(^\text{[23]}\). T7 Endo I induces a conformational change that is similar to the Stacked-X conformation with the two stacked helices held almost perpendicular to one another\(^\text{[22]}\).

1.5 Ligands of Interest

In this study we have focused on binding bimodal ligands to cruciform structures developed by Mark Searcey’s group at the University of East Anglia. Each of the ligands shares a common binding motif, the bis-acridine-carboxamide group connected by a linker of variable length and rigidity (Figure 1.4). The linker is of a sufficient size to prevent bis-intercalation along the same duplex, but encourages the ligand to bind across narrow regions such as DNA cross-linking or binding at the core of a HJ. The latter would be more energetically favored. One ligand in particular, LH3.123, has been shown to form a complex with a HJ by binding to the core in an X-ray crystal structure (Figure 1.4). This binding displaces the adenosines at the junction and forms pseudo base pairs with the thymidines\(^\text{[16]}\).
Figure 1.4 - Cruciform Binding Ligands. Each of the ligands, with the exception of LH2.49, has two acridine rings with an attached carboxamide group at varying positions. The ligands also vary in the length and rigidity of their linker chain.
2 EXPERIMENTAL

2.1 Conformational Dynamics

2.1.1 Conformational Effects on Circular Dichroism

CD experiments were performed with GSU’s core facility Jasco J-810 Spectropolarimeter running Spectra Manager version 1.53.00. Raw data was obtained from the instrument and was smoothed using KaleidaGraph (Stineman function with a window size of 15). Cruci3HL (Appendix: Figure SI-1) at a concentration of 1.4µM in a 10 mM Tris-HCl buffer with 10 mM NaCl at pH 7.0 was loaded into a 1 cm cuvette. MgCl₂ was titrated into the cuvette in 4 mM intervals which was then mixed and a new spectrum was run. After the final addition of MgCl₂ (12 mM) the sample was heated to 95 °C for 5 minutes and gradually returned to room temperature. The spectra were recorded over a range of 340 nm to 225 nm and averaged over 4 scans. This was repeated for Cruci3HL-Hairpin-Loop and for Junction3 at a concentration of 1 OD/mL.

2.1.2 Conformational Effects on Gel Mobility

The gel mobility of the cruciform, Cruci3HL was compared to the mobility of a single stranded 50-mer and a single stranded 90-mer (Appendix: Figure SI-1) separately on a native 12% PAGE gel with and without 5 mM MgCl₂ to observe if there was a difference in the relative mobilities of the species. These gels were cast with the same ionic conditions in the gel as in the loading buffer to ensure that the DNA migrates onto the gel without smearing. Another gel was run in triplicate to ensure accuracy comparing the relative mobility between Cruci3HL and Junction3 with and without 2 mM MgCl₂.
2.2 Ligand Binding to Cruciforms and Branched DNA

2.2.1 T7 Endonuclease I Cleavage of Cruciform DNA

T7 Endo I was obtained from New England Biolabs. All enzymatic cleavages were run under similar conditions to ensure consistent enzymatic activity. The enzyme was brought up in a 1x NEB buffer, 50 mM NaCl 10 mM Tris-HCl 10 mM MgCl₂ 1 mM DTT pH 7.9. The enzymatic cleavage reaction of T7 Endo I on cruciform DNA was performed with a final concentration of 1x NEB in 8µL total volume. T7 Endo I was always added last to ensure that the reaction begins at a consistent time. After addition of T7 Endo I the sample then was mixed in an Eppendorf tube and submerged in a water bath set at 37 °C. Submerging the tube prevents condensation of the sample so the concentrations are constant throughout the reaction. The amount of T7 Endo I used and the reaction time was varied depending on the DNA sequence analyzed. To terminate the reaction the sample was taken out of the water bath and mixed with 2 µL DMSO and immediately frozen at -20 °C. The samples were then mixed with the appropriate loading buffers and run on a corresponding gel to analyze the cleavage. This same process was used, but with the addition of cruciform binding ligands to see if they had any inhibitory effects on the cleavage of cruciforms.

2.2.2 Ligand Binding Determined by Its Effect on Tₘ

LH3.123 is the most studied of the ligands received and it exhibits very tight binding to cruciform DNA. This was shown by measuring the Tₘ of the cruciform at a concentration of 1.4 µM in the presence of 10 mM sodium phosphate buffer with 10 mM NaCl then adding a 5-fold excess of LH3.123 and comparing the Tₘ in a 1 cm cuvette. This was performed on a Cary-UV spectrometer.
2.3 NMR Experiments

2.3.1 Temperature Series

Each of the DNA hairpins corresponding to the arms of Junction3’s J3B were analyzed on a Bruker Avance 600MHz NMR and a Bruker Avance 500 MHz NMR. The temperature was varied from 278K to 318K to observe the effects different temperatures have on the imino proton range of 1D $^1$H spectra. All experiments were performed with a 1-1 jump and return pulse sequence for solvent suppression.

2.3.2 1D NOE

A 1D NOE experiment was carried out on the sample J3B-Hairpin at a concentration of 900 µM. The sample was prepared with 18.3 mM NaCl, 10 mM NaP, 10% D$_2$O, and had a pH of 6.74. The experiment was run at 278 K. Select frequencies were irradiated, at the A-T base pairs and the T-loop regions, individually to measure the NOE effect at a different location.

2.3.3 2D NOESY

2D NOESY were performed on J3B-HP and each of the overhang sequences on a Bruker Avance 600 MHz NMR. The conditions were 20 mM NaCl, 10 mM NaP, 99.8% D$_2$O, and 400-600 µM DNA at a pH of 6.7 and a temperature of 288 K. The mixing time was 200 ms with a relaxation delay of 3 s. The experiment was repeated on the 500 MHz NMR. Each sample was brought up in 99.8% D$_2$O and then lyophilized to replace the exchangeable protons with deuterium. The resultant powdered DNA was then brought up in fresh 99.996% D$_2$O. The buffer and salt conditions remained the same. The mixing time was lowered to 150 ms with the same relaxation delay of 3 s.

2.3.4 T1 Experiments

The same samples used for the 2D NOESY experiment were used to run a T1 experiment at 288K on the Avance 600MHz NMR. A Interscan delay of 30s was used with variable delays of 0.01 s, 0.15 s, 0.3
s, 0.8 s, 1 s, 2 s, 3 s, 5 s, 8 s, 12 s, 18 s, and 30 s for samples containing 99.8% D₂O. This experiment was repeated with small modifications, interscan delay of 25s and variable delays of 0.01s, 0.2s, 0.6s, 1s, 1.5s, 3s, 5s, 9s, 15s, 255s for samples containing 99.996% D₂O. The data was fitted with Bruker Topspin’s 2.1 inversion recovery fitting.

### 2.3.5 Ligand Titrations

Each of the ligands, LH3.123, LH4.14, and LH 4.50 were titrated onto both J3B-HP and J3B-Poly-T. Titrations were all performed under the same conditions of 18.3 mM NaCl, 10 mM NaP, 10% D₂O at a pH of 6.7 and a temperature of 288 K.
3 RESULTS

3.1 Branched DNA Substrates

Branched DNA offers some potential binding sites not typically found on duplex DNA such as the sharp angles of the DNA backbone and the core of a HJ in structures such as cruciform DNA and Y-structures found at replication forks. The scope of this project has led us to create a wide variety of branched DNA sequences. Interestingly, T7 Endonuclease I has activity on more than just Holliday junctions. The enzyme is known to cleave other structures such as branched DNA\(^{24}\). Because of this we can study the effect ligands have on substrates with and without a HJ core. We have investigated the inhibition of T7 Endo I on branched DNA by binding ligands in section 3.3 and we have studied the effect ligands have on the structure of branched DNA via NMR analysis in section 3.4.

3.1.1 Cruciform Substrates

The first sequence studied was the bimolecular cruciform, Cruci2H (Appendix: Figure SI-1). This construct dissociated at low concentrations which made T7 Endo I analysis very problematic (data not shown). To stabilize the junction, the sequence was modified to a single stranded intramolecular cruciform held together by an additional hairpin, denoted as Cruci3HL for its three hairpin loops (Figure 3.1). Despite its stability, Cruci3HL can undergo branch migration in the Stacked-X conformation. Due to this, the sequence at the core of the junction in solution is unknown. A third sequence, Junction3, was created from two strands which contains the well-studied Junction3 core. First published in 1988\(^{25}\), the sequence has arms of a similar length to Cruci3HL. Junction3 cannot undergo branch migration in the Open-X or Stacked-X conformation, and in addition it has a preference to form the Iso II conformation (Figure 1.1). In addition to knowing the dynamics of this junction, we also know the primary cleavage pattern of the enzyme by T7 Endonuclease I; the sequence is cut at two primary locations across the diagonal of the junction\(^{22}\) (Figure 3.1) Another cruciform from the literature was studied, Cruci4S,
which is a 10 nucleotide 4-way self complementary sequence (Appendix: Figure SI-1) (Figure 3.2). Evidence of its formation is shown from its X-ray crystal structure\textsuperscript{26}. Interestingly, Brogden et al. found that adding the ligand LH3.123 spontaneously formed a crystal structure with the ligand bound neatly to the core of this junction\textsuperscript{16} (Figure 1.1). The sequence behaves very differently in solution. On a native gel the sequence runs as a duplex regardless of concentration, salt conditions, or in the presence of LH3.123 (data not shown). Similar tetrameric 10 nucleotide inverted repeat sequences have been studied in solution by measuring their sedimentation coefficient. However, it is unclear if the higher sedimentation values are due to concatemer formation from the high concentration of DNA used (1.02 mM)\textsuperscript{27}.\textsuperscript{26} \textsuperscript{16} \textsuperscript{27}
Figure 3.1- The Cleavage of Junction3 by T7 Endonuclease. This diagram shows the primary cleavage pattern of junction3 denoted by the large arrows. The smaller arrows on the opposite side of the junction show the minor cleavage pattern. This diagram is taken from “The complex between a four-way DNA junction and T7 endonuclease I” [22].
3.1.2 Partial Junctions

The simplest branch structure is just one strand from Junction3, either JunctionA or JunctionB. The structure contains one hairpin loop and two single stranded tails (Figure 3.3). T7 Endonuclease I cleaves these structures at a different rate, despite having the same conditions (Figure 3.9). This discrepancy in enzymatic activity on the two substrates is likely caused by secondary structure formation of the single stranded tails. To test this, the tails were replaced with 13 thymidines on the 5’ side and 11 thymidines on the 3’ side. The new sequences were referred to as J3A-Poly-T and J3B-Poly-T (Figure 3.4). The T-tails are not expected to form secondary structure. These new structures, J3A-Poly-T and J3B-Poly-T are cleaved at the same rate. Partial junctions like these are cleaved primarily one nucleotide away from the branch point on the 5’ side. 

![Figure 3.2- Schematic of Crucifer4S Folded into a Holliday Junction.](image)

\[5’-(TCGTTACCGA)4-3’\]
Figure 3.3- Model of the Partial Junctions, JunctionA and JunctionB. Junction3 is composed of two strands which have been studied independently because their unique single stranded tails form a partial junction when connected to the hairpin stem.
Figure 3.4 - Model of the Partial Junctions with Modified Thymidine Tails. The same hairpin loops from Junction A and Junction B had their single stranded tails replaced with thymidines to reduce secondary structure formation.

3.1.3 Y-Structure DNA

We were also interested in Y-structure DNA. This structure was designed by using the partial junction, Junction B, and adding a complementary sequence to coordinate with the tails. Two sequences were ordered, one which was a perfect complement to the tails (Y-Structure) and one which contained an extra thymidine at the base of the stem loop to give added flexibility (Y-Structure + 1NT) (Figure 3.5). Both structures were not very stable. However, the extra thymidine caused T7 Endo I to cleave this structure to a greater extent than the Y-Structure without it (Figure 3.20). Y-structures, similar to the partial junctions of the last section are also cleaved primarily one nucleotide from the branch point on the 5’ side.
Figure 3.5- Model of the Y-Structure DNA. This diagram shows the two Y-structures we worked with. This model represents the two Y-structures we analyzed, Y-Structure where the red x at the center of the junction is missing and Y-Structure + 1NT, in which the red x is a thymidine.

3.2 Conformational Dynamics

3.2.1 Conformational Effects on Circular Dichroism

It has been shown that Holliday junctions can exist in two different conformations, the Open-X and the Stacked-X forms. It was believed that the difference between the two conformations could be detected with circular dichroism. The Stacked-X conformation has been shown to be the predominant form in the presence of MgCl₂. Several spectra were run with MgCl₂ being titrated into a solution with Cruci3HL.
Figure 3.6 - Open-X to Stacked-X Conformational Switch. This diagram shows the possible conformational swap between the Open-X and Stacked-X form of the intramolecular cruciform Cruci3HL.
Figure 3.7 - Effect of MgCl₂ on the CD spectra of Cruci3HL. CD spectra overlayed of Cruci3HL with MgCl₂ titrated in from 0 mM – 12 mM. There is a significant drop in the molar elipticity at 285 nm when MgCl₂ is added. 1 OD/mL Cruci3HL, 10 mM NaCl, 10 mM Tris-HCl pH 7.0.

The first addition of MgCl₂ drastically changes the CD spectrum by causing a decrease in the molar elipticity at 285 nm and at 250 nm. Further additions of MgCl₂ change the spectra to a lesser extent. Finally, the sample was heated to 95 °C for 5 minutes and returned to room temperature; this had no effect on the spectrum. Under all ionic conditions studied the cruciform exhibited a B-DNA form. To see if the effect of MgCl₂ on the CD spectra was unique to Cruci3HL, a hairpin loop with the same sequence of an arm of Cruci3HL was run with the same conditions (Figure 3.8). The spectra of the hairpin and of Cruci3HL were almost identical. We can conclude that MgCl₂ changes the structure of DNA to some extent (through interaction with the backbone and nucleobases).
Figure 3.8 - Effect of MgCl₂ on the CD spectra of Cruc3HL-Hairpin-Loop. CD spectra overlayed of Cruc3HL-Hairpin-Loop with MgCl₂ titrated in from 0 mM – 12 mM. The spectrum shows the same large dip at 285 nm. 1 OD/mL Cruc3HL-Hairpin-Loop, 10 mM NaCl, 10 mM Tris-HCl, pH 7.0.
Figure 3.9- Effect of MgCl$_2$ on the CD spectra of Junction3. CD spectra overlayed of Junction3 with MgCl$_2$ titrated in from 0 mM – 16 mM. The spectra looks similar to Cruci3HL and Cruci3HL-Hairpin-Loop with the dip at 285 nm being less significant. 1 OD/mL Junction3, 20 mM NaCl, 10 mM Tris-HCl (pH 7.0).

The CD spectra of Junction3 maintain similarities to Cruci3HL and Cruci3HL-Hairpin-Loop such as the dip in molar elipticity at 285 nm and 245 nm as MgCl$_2$ is titrated into the cuvette. However, the change in elipticity is much lower for Junction3 at 285 nm. It remains a challenge to determine exactly why this spectra is different, but a possible explanation is that the cruciform formed from Junction3 (JunctionA + JunctionB) is not entirely stable without the presence of MgCl$_2$ as evidenced on several native gels. With just NaCl present, Junction3 is approximately 90% cruciform and 10% duplex. With the addition of MgCl$_2$ the junction is completely stabilized with no duplex formation (data not shown). This change would impact the CD spectra and hamper the analysis.
3.2.2 Conformational Effects on Gel Mobility

Duckett et al. showed that Junction3 will form the Stacked-X or Open-X based on the ionic conditions present via comparative gel electrophoresis\(^{[25]}\). Adding 1 mM MgCl\(_2\) changed the overall conformation from Open-X to Stacked-X\(^{[25]}\). It is expected that the different conformations would have different electrophoretic mobilities. A Native PAGE was used to test for this, each gel had Cruci3HL and was flanked by a 90-mer control and a 50-mer control. One gel contained MgCl\(_2\) in the running and loading buffer while the other gel had EDTA in each. The gels show that Cruc3HL runs at the same mobility regardless of the presence of MgCl\(_2\). This was determined by measuring the ratios of the distance of Cruci3HL to the 90-mer and to the 50-mer, the ratios are essentially the same. Again, this doesn’t prove or confirm that MgCl\(_2\) causes a conformational change. Continuing on the gel mobility study which has shown us that Cruci3HL is unaffected by salt changes, we can now use this as a control for different cruciform structures such as Junction3. Junction3’s mobility is significantly impacted by the presence of 2 mM MgCl\(_2\) (Figure 3.11). Junction3 travels 16% faster in the presence of MgCl\(_2\) compared to EDTA. It is clear that the structure is undergoing a conformational change.
Figure 3.10- Conformation Dynamics of Cruci3HL Compared to Single Stranded DNA. This gel compares the effect of MgCl₂ on electrophoretic mobility. The lanes are the same in each gel with the exception of the running buffer, on the left each sample is in a 89 mM Tris-borate buffer containing 5mM MgCl₂. The gel on the right has each sample in a 89 mM Tris-borate buffer with 2 mM EDTA. Lanes 1-3 are 90-mer, Cruci3HL, and the 50-mer respectively.

Figure 3.11- Conformational Dynamics of Cruci3HL compared to Junction3. The gel on the left shows the mobility of Cruci3HL compared to the mobility of Junction3 in the presence of 2 mM MgCl₂. Since Cruci3HL is not affected by salt changes it is used as a reference for the relative mobility of Junction3. The gel on the right has the same conditions but without MgCl₂ and in the presence of 2 mM MgCl₂.

3.3 Ligand Interaction with Branched DNA

3.3.1 Effect of Ligand Binding on the $T_m$ of a Cruciform

Measuring the $T_m$ of a DNA structure is a widely used technique that can be used to determine a variety of thermodynamic properties. The $T_m$ value indicates the temperature when half of the DNA is
melted. As the DNA melts the bases become unstacked and exhibit higher absorbance. The melting curve is typically sigmoidal representative of the cooperative melting. In the case of Cruci3HL there are two transitions instead of one, this is due to the intramolecular design only allowing melting to occur at the open stem; once the first stem melts the remaining arms can melt simultaneously. The melting profiles show an increase of 10 degrees in the Tm of the first transition after the addition of LH3.123, this is representative of strong binding that greatly stabilizes the cruciform structure (Figure 3.12).

**Figure 3.12 - Effect of LH3.123 on the Tm of Cruci3HL.** (A) The Tm of 1.4µM Cruci3HL is shown by itself in 10 mM NaP buffer with 10 mM NaCl; two Tm’s can be easily determined at 260 nm wavelength. (B) A 5-fold excess of LH3.123 was added to the sample under the same conditions, only the first Tm is obtainable. A 240 nm excitation wavelength was used since LH3.123 has a λmax at 280 nm. Adding a 20-fold excess of LH3.123 to the cruciform prevents the DNA from fully melting, even at 95 °C (Data not shown).

### 3.3.2 Effect of Ligand Binding by CD Analysis

We used CD as a means to test if ligands were binding to cruciform DNA and to monitor what structural changes they induce. LH3.123 causes a positive change in the CD spectra at 330 nm and a negative change at 248 nm regardless of salt conditions and on both Cruci3HL and the duplex sequence Cruci3HL-Hairpin-Loop.
Figure 3.13 - CD Spectra of LH3.123 Titrated into Cruci3HL no NaCl. The CD Spectra of Cruci3HL experiences two main changes as LH3.123 is titrated in. There is a positive shift at 330nm and a negative shift at 248nm. 1 OD/mL (1.1 µM) Cruci3HL, 10 mM Tris-HCl (pH 7.0), 10 mM NaCl.
Figure 3.14- CD Spectra of LH3.123 Titrated into Cruci3HL with NaCl. The spectra exhibit changes at the same positions as the spectra without NaCl. 1 OD/mL (1.1 µM) Cruci3HL, 10 mM Tris-HCl (pH 7.0), 10 mM NaCl.

Figure 3.15- CD Spectra of LH3.123 Titrated into Cruci3HL-Hairpin-Loop no NaCl. The spectra is affected in the same way as the full cruciform, Cruci3HL. 1 OD/mL (4.5 µM) Cruci3HL-Hairpin-Loop, 10 mM Tris-HCl (pH 7.0), no NaCl.
3.3.3  

**T7 Endonuclease I Cleavage and its Inhibition by Ligands**

![Figure 3.16](image.png)

Lane 1: Cruci3HL + T7 (Control)
Lane 2: Cruci3HL + T7 + 20x LH2.49
Lane 3: Cruci3HL + T7 + 20x LH3.123
Lane 4: Cruci3HL + T7 + 5x LH3.123
Lane 5: Cruci3HL + T7 + 20x LH3.124
Lane 6: Cruci3HL + T7 + 5x LH3.124
Lane 7: Cruci3HL + T7 + 20x LH4.14
Lane 8: Cruci3HL + T7 + 5x LH4.14
Lane 9: Cruci3HL + T7 + 20x LH4.49
Lane 10: Cruci3HL + T7 + 5x LH4.49
Lane 11: Cruci3HL + T7 + 20x LH4.50
Lane 12: Cruci3HL + T7 + 5x LH4.50
Lane 13: Cruci3HL + T7 (Control)

**Figure 3.16- Inhibition of T7 Endo I Cleavage by Interaction with Ligands on Cruci3HL.** 12% Native PAGE showing how the various ligands inhibited the cleavage of the cruciform at differing concentrations. All T7 Endo I enzymatic reactions were run for 1 hour with 1 unit of T7 Endo I with 445 nM DNA. Lanes 3, 7, and 8 have nearly 100% inhibition which points to strong interaction.

The above gel shows that all of the LH ligands, with the exception of LH2.49, show some degree of inhibition (Figure 3.16). It is clear that these ligands can bind to cruciform DNA, but determining if the ligands are binding to the core of a Holliday junction requires a different screen.
Running the best binding ligands with Junction3 yielded similar results. J3B-Poly-T was run on the same gel to compare binding of the ligands to a branched DNA structure containing the same hairpin loop, but lacking a junction core (Figure 3.17). LH4.50 inhibits the partial junction slightly more than the full Junction3.

Figure 3.17- Inhibition of T7 Endo I Cleavage by Interaction with Ligands on Junction3 and J3B-Poly-T. 12% Native PAGE. All T7 Endo I enzymatic reactions were run for 40 minutes with 0.9 units of T7 Endo I with 445 nM DNA.
3.3.4  T7 Endo I Cleavage of Partial Junctions

As a control we examined what effect T7 Endo I would have on half of our first bimolecular cruciform, Cruci2H. A gel was run where T7 Endo I was mixed with Cruci B and Cruci3HL to compare cleavage. Surprisingly the enzyme cleaved both DNA structures with similar rates (data not shown). This isn’t necessarily bad news, as T7 Endo I will not cleave duplex DNA, but it will cleave hairpin loops with single strands attached. We can now use this as a test to compare inhibition of T7 Endo I when ligands are bound to either half the junction or to the whole junction. If the ligand only interacts with normal duplex regions and not the junction itself, then the inhibition should be the same, while if the ligand has specificity to the core of the junction it will inhibit the partial junctions significantly less. As a preliminary step before we begin studying the inhibition of the partial junctions we need to see how they are cleaved in comparison to a whole junction. Figure 3.18 shows us that Junction A is cleaved to a lesser extent than the other structures. It was believed this was caused by secondary structure formation in the single stranded region. To test for this we removed that portion of the partial junctions and replaced them with thymidines so they can no longer form secondary structures (J3A-Poly-T and J3B-Poly-T). Replacing these tails with T’s remediates the discrepancy in cleavage (Figure 3.18).
3.3.5 \textit{T7 Endo I Inhibition of Partial Junctions by Ligand Interaction}

After determining that secondary structure can cause different cleavage rates, we chose to use Junction B Poly-T (referred to as J3B-Poly-T above) for our inhibition study. Figure 3.19 compares the ligands which bind to Crucii3HLL with the highest affinity as they bind to a partial cruciform. LH3.123 completely inhibits the cleavage which means that it’s binding is non-specific to Holliday junctions. LH4.14 and LH4.50 are more promising as they show a diminished degree of inhibition compared to a whole junction.
Figure 3.19- Inhibition of T7 Endo I by Ligands Binding to DNA Partial Junctions. This gel shows how the most potent ligands effect T7 Endo I cleavage on a partial cruciform. T7 Endo I reactions were run for 1 hour with 1 unit of T7 Endo I with a DNA concentration of 445 nM. LH3.123 inhibits almost 100% while LH4.14 and LH4.50 still inhibit but to a much less extent than they do with Junction3 and Cruc3HL.

3.3.6 T7 Endonuclease I Cleavage of Y-Structures

Additional sequences were examined to test the enzymatic cleavage on branched structures such as a Y-structure junction. Two structures were designed based on the Junction B stem loop that had a complimentary strand to coordinate with the tail region (Appendix: Figure SI-1). Two complimentary strands were designed, one which contained an extra thymidine to help stabilize the structure and one without. The Y-Junctions are less stable than cruciform as a whole, both the Y-structure with and without the extra nucleotide show a band corresponding to the unformed junction. However, enzymatic activity is higher with the extra nucleotide at the core (Figure 3.20). This difference in enzymatic activity is believed to be due to the extra flexibility that comes from adding an additional nucleotide at the branch point.
12% Native Gel TBM

Lane 1: Y-Structure
Lane 2: Y-Structure + 0.9units T7 40min
Lane 3: Y-Structure + 1NT
Lane 4: Y-Structure + 1NT + 0.9units T7 40 min
Lane 5: Junction 3 + 0.9units T7 40min

Figure 3.20- Y-Structure Stability Gel. Native gel showing the different enzymatic activity on the same structure with and without an extra nucleotide at the core. All enzymatic reactions used 445nM DNA in a 8µL reaction. Notice how the Y-Structure with the extra nucleotide is cleaved more similarly to Junction3

3.4 NMR Analysis of Branched DNA and Ligand Binding

To determine if any of the ligands target the core of branched DNA we used NMR spectra to identify binding sites.

3.4.1 NMR Assignment of a Branched DNA Structure and its Hairpin

Before observing the effect the ligand has on the DNA spectra the imino proton resonances of the hairpin (J3B-HP) have to be assigned. The spectrum of the hairpin gave 9 distinct peaks, however, the imino proton of thymine at position 1 exhibited an unusual chemical shift (12.15 ppm). This shifted resonance has been associated with Hoogsteen base pairing where the purine is flipped into the syn orientation\(^{[29]}\). Typically, Watson-Crick A-T base pairs are seen farther downfield in the range of 13-14 ppm. Additionally, the spectra of the Poly-T sample showed the terminal A-T base pair in its expected location (Figure 3.22). It follows that the poly-T tail at the end of the hairpin is responsible for returning the terminal base to its expected chemical shift.
Figure 3.21- NMR Assignment of J3B-HP. The NMR spectrum has been assigned to indicate which imino proton peaks correspond to what base pair. The T-loop can be seen even though they do not form a basepair in the region from 11.4-10.6 ppm. 900 μM J3B-HP 18.3 mM NaCl, 10 mM NaP, pH 6.79, 500 MHz NMR.
The NMR spectrum of the hairpin loop with a poly-T tail shows imino proton peaks at the same position as the hairpin with the major exception of the terminal A-T base pair which has shifted from 12.2 ppm to 13.8 ppm indicative of a shift of adenine from syn to an anti glycosidic torsion. 181 μM Poly-T, 18.3 mM NaCl, 10 mM NaP, pH 6.74, 500 MHz NMR.

3.4.1.1 1D NOE to Confirm Terminal Base Assignment

Confirmation was needed to determine that the lone peak at 12.2 ppm was in fact from the terminal base pair. This was accomplished by using a 1D NOE experiment where each imino proton from an A-T base pair in J3B-HP was irradiated; any peak that shows a drop in intensity corresponds to a proton that is within 5 Å through space from the originally irradiated proton (Appendix Figure SI-3). Irradiation of T1 caused a drop at 7.82 ppm, T20 at 7.78 ppm, and T8 at 8.00 ppm. These resonances are in the adenosine-H2 range. We then ran a simple T1 experiment to confirm the adenosine-H2’s
resonances by varying the relaxation delay. AH2 protons have long T1 values, so as the relaxation delay is increased the signal from the slowly-exchangeable AH2 increases dramatically. This experiment showed an increase in signal at 7.82 ppm, 7.78 ppm, and 8.00 ppm (Figure 3.23) confirming their assignment.

Figure 3.23- T1 Experiment to Identify the Adenosine H2. The spectra above show that as the relaxation delay (D1) is increased, the signal from the slow exchanging adenosine-H2s is dramatically increased. The numbering above the peaks correspond to the H2s of adenosine at position 3, 15, and 22. The red, blue, and black lines correspond to a relaxation time of 1 s, 4 s, and 8 s respectively.

3.4.2 NMR Analysis of Ligand Binding

With the DNA sequences assigned we can now explore the effect ligands have on the NMR spectra of J3B-HP compared to J3B-Poly-T. We investigated the effects of the strongest cruciform binding ligand, LH4.14, the well-studied LH3.123, and LH4.50, which has been shown to inhibit T7 Endo I to a greater extent in cruciform DNA compared to partial junctions (Figure 3.19). LH4.14 doesn’t affect the chemical shift of the terminal A-T base pair for either J3B-HP or J3B-Poly-T. LH4.14 also causes little change to the T-loop. LH4.14 mainly affects the base pairs on the stem of the hairpin. LH3.123 also has
no effect on the terminal A-T base pair or on the T-loop. It affects G-C base pairs more than A-T, which remain relatively unperturbed through the titration. Addition of LH4.50 causes a drop in the signal intensity, this could be due to precipitation or non-specific binding. One similarity between all the spectra is that titrating ligands has no effect on the terminal A-T base pair. This test implies that the ligands interact preferentially with canonical duplex DNA (Ligand Titrations can be found in Appendix Figure SI-5).

### 3.4.3 NMR Structural Studies of a Hairpin Sequence with a T-Overhang

DNA duplexes experience weak base pairing at their stems due to a less constrained environment. This freedom can give rise to the non-canonical Hoogsteen base pair where the purine base flips into the syn conformation\(^{[29]}\). We are interested in investigating non-intrusive end modifications that can suppress the conformational heterogeneity.

![Watson-Crick vs Hoogsteen](image)

**Figure 3.24- Watson-Crick Base Pair Compared to a Hoogsteen Base Pair.** The WC base pair has both bases in the anti configuration while the Hoogsteen base pair has the adenosine flipped around into the syn configuration.

Since adding a T tail to the end of a hairpin evidently shifts the terminal A-T basepair to its typical resonance position, it would be of interest to determine if simple end modifications could elicit a similar response. Three additional sequences were studied, J3B-HP-3'-Overhang, J3B-HP-5'-Overhang,
and J3B-T-T Mismatch, each of these hairpins have the same sequence as J3B-HP with the only
difference being a thymidine added to the 3’ or 5’ end and a thymidine added to both respectively. The
sequences were analyzed by 1D NMR with a 1-1 jump and return water suppression pulse and also by a
2D NOESY in 99.8% D₂O to measure the distance from the H8-H1’ of the terminal adenosine.

Figure 3.25- J3B-HP-3’-Overhang Assignment. Adding an additional thymidine to the 3’ end (nearest the
terminal adenosine) shifts the terminal A-T peak slightly downfield compared to the unaltered J3B-
Hairpin. The adenosine is still predominantly in the syn conformation here. 20 mM NaCl, 10 mM NaP,
500 µM hairpin, pH 6.67, 600 MHz NMR.
Figure 3.26: J3B-HP-5’-Overhang Assignment. Adding a thymidine to the 5’ end shifts the terminal A-T base pair almost all the way to the A-T base pair region. 20 mM NaCl, 10 mM NaP, 582 µM hairpin, pH 6.68, 600 MHz NMR.

Figure 3.27: J3B-HP-TT-Mismatch Assignment. Placing thymidines on both sides of the hairpin fully shifts the terminal A-T base pair to a typical resonance. 20 mM NaCl, 10 mM NaP, 585 µM Hairpin, pH 6.70, 600 MHz NMR.

In the following figures, we can see how varying temperature effects the terminal A-T base pairs resonance. The terminal base pair is marked by a red arrow (Figures 3.28-3.32).
Figure 3.28- J3B-Hairpin Temperature Study. NMR Spectra of J3B-Hairpin from 278 K-308 K. At low temperatures the terminal base pair can be seen between 12.5 ppm and 12.0 ppm. As the sample is heated the proton exchanges more rapidly with water and the signal drifts towards the water peak. 18.3 mM NaCl, 10 mM NaP, 900 uM DNA, pH 6.79, 500 MHz NMR.
Figure 3.29- J3B-Poly-T Temperature Study. NMR Spectra of J3B-Poly-T from 278 K-308 K. As temperature is increased the peak shifts out towards the water peak. 18.3 mM NaCl, 10mM NaP, 181 uM J3B-Poly-T, pH 6.74, 500 MHz NMR.
Figure 3.30 - J3B-HP-5'-Overhang Temperature Study. NMR spectra of J3B-HP-5'-Overhang from 278 K to 308 K. The terminal AT base pair is significantly shifted towards the typical Watson-Crick A-T base pair region at low temperatures. 20 mM NaCl, 10 mM NaP, 582 µM hairpin, pH 6.68 600 MHz NMR.
Figure 3.31- J3B-HP-3'-Overhang Temperature Study. NMR spectra of J3B-HP-3'-Overhang from 278 K to 298 K. These spectra shows that the terminal A-T base pair is actually buried with the G-C base pairs at 278 K as temperature is increased it once again shifts out to water. 20 mM NaCl, 10 mM NaP, 500 µM hairpin, pH 6.67, 600 MHz NMR.
Figure 3.32- J3B-HP-T-T Mismatch Temperature Study. NMR spectra of J3B-T-T Mismatch from 278 K to 308 K. These spectra shows that the terminal A-T base pair is stabilized similarly to J3B-Poly-T. As temperature increases the peak shifts towards water much more drastically than the other base pairs indicating that it is not entirely kinetically stable. 20 mM NaCl, 10 mM NaP, 585 µM Hairpin, pH 6.70, 600MHz NMR.

Table 3.2- End Modification Effect on Chemical Shift

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Chemical Shift of Terminal A-T Base Pair at 288 K (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J3B-Hairpin</td>
<td>12.50</td>
</tr>
<tr>
<td>J3B-HP-3’-Overhang</td>
<td>12.58</td>
</tr>
<tr>
<td>J3B-HP-5’-Overhang</td>
<td>13.52</td>
</tr>
<tr>
<td>J3B-HP-TT-Mismatch</td>
<td>13.82</td>
</tr>
<tr>
<td>J3B-HP-Poly-T</td>
<td>13.75</td>
</tr>
</tbody>
</table>

The NOESY spectra were assigned and the A22-H8-H1’ cross peak was integrated via the program Sparky. The A22-H8-H1’ volume was divided by a reference distance of the cytosine H5-H6 cross peak. This value is listed in Table 3.3 under the NOE Ratio column; this ratio represents the distance of the A22 H8 proton to the proton on its sugar H1’. If the ratio is 1, the distance between H8
and H1′ is 2.46 Å (the distance of the reference cytosine H5-H6), as the ratio drops, the distance increases.
Table 3.3 - Table of NOE Intensities

<table>
<thead>
<tr>
<th>Sequence</th>
<th>NOE Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>J3B-T-T Mismatch</td>
<td>0.39±0.02</td>
</tr>
<tr>
<td>J3B-HP-5'-Overhang</td>
<td>0.33±0.01</td>
</tr>
<tr>
<td>J3B-HP-3'-Overhang</td>
<td>0.52±0.06</td>
</tr>
<tr>
<td>J3B-HP</td>
<td>0.25±0.01</td>
</tr>
</tbody>
</table>

The NOESY integrations imply that J3B-HP has the least amount of syn conformation compared to all the other hairpin structures. This is inconsistent with the chemical shift data. We believe this NOE value is affected by unusual relaxation effects of the H8 proton. We have therefore recorded a T1 experiment to support this notion. The data shows that J3B-HP has the shortest T1 value for the terminal AH8 which may explain this aberrant NOE ratio (Table 3.4).

Table 3.4 - List of T1 values for the Adenine Non-Exchangeable Protons

<table>
<thead>
<tr>
<th>Sequence</th>
<th>A22/A23 H8</th>
<th>A22/A23 H2</th>
<th>A15/A16 H8</th>
<th>A15/A16 H2</th>
<th>A3/A4 H8</th>
<th>A3/A4 H2</th>
</tr>
</thead>
<tbody>
<tr>
<td>J3B-TT-Overhang</td>
<td>2.36</td>
<td>4.15</td>
<td>2.90</td>
<td>5.22</td>
<td>2.69</td>
<td>5.94</td>
</tr>
<tr>
<td>J3B-HP-5' Overhang</td>
<td>1.78</td>
<td>4.20</td>
<td>2.84</td>
<td>N/A</td>
<td>2.84</td>
<td>6.11</td>
</tr>
<tr>
<td>J3B-HP-3' Overhang</td>
<td>2.01</td>
<td>N/A</td>
<td>2.73</td>
<td>N/A</td>
<td>2.78</td>
<td>5.87</td>
</tr>
<tr>
<td>J3B HP</td>
<td>1.24</td>
<td>4.2</td>
<td>2.17</td>
<td>N/A</td>
<td>1.77</td>
<td>6.29</td>
</tr>
</tbody>
</table>

4 CONCLUSIONS

During our NMR studies on branched DNA, we observed an unusual conformational behavior on the terminal base pair that may be a target for ligands and T7 Endo I. In an attempt to return the base pair to its normal anti conformation, we added terminal thymidines on the ends of J3B-HP. These modifications had varying effects on the chemical shift in the imino proton spectra; J3B-HP-TT-Mismatch and J3B-HP-5'-Overhang shifted the terminal base pair to the normal anti conformation while J3B-HP-3'-Overhang and J3B-HP remained in the syn conformation. The NOESY ratios exhibit a different trend which can be explained by an unusual relaxation that could be resultant from its neighboring bases.
Finding a ligand that binds to the core of a Holliday junction is not a simple task as Holliday junctions are dynamic and are able to have very different local environments. Small changes in the sequence as far as 8 base pairs away from the point of strand exchange can cause a change in the isomeric structure, which in turn can cause the core of the junction to have either an exposed minor or major groove\cite{20}. Synthetic HJ constructs are capable of conformational changes from the Open-X to the Stacked-X structure. Our gel electrophoresis experiments indicate that the Junction3 sequence changes from the Open-X to the Stacked-X conformation by addition of 2 mM MgCl$_2$. The core of the junction is much narrower in the Stacked-X form and it is likely that ligands will interact differently depending on the conformation of the HJ. In cells Holliday junctions are even more mobile, due to branch migration, which is enabled by sequence homology. Thus, an effective drug that binds to the core of a HJ must be able to interact favorably to a wide variety of sequences and structures while maintaining selectivity to HJs over duplex DNA. In addition, cruciform contain a wide array of binding targets that are not unique to branched DNA, such as the grooves, base pairs, hairpin loops, and the DNA backbone.

In order to study the structure of the ligand interaction it is essential to have an immobilized HJ. We used Cruci3HL as our primary model to study ligand binding. The majority of the ligands we investigated bound very tightly to cruciform DNA, but most lacked specificity as they bound duplex DNA to a similar extent as evidenced by CD, $T_M$, NMR, and gel electrophoresis. We devised an experiment to test for selectivity by measuring the inhibition of T7 Endo I’s cleavage on partial junctions, which showed us that LH4.14 and LH4.50 inhibit the enzymes cleavage to a lesser extent on partial junctions compared to a complete junction such as Cruci3HL or Junction3.

Looking forward, a potent drug for binding cruciforms specifically should contain 4 weak binding modules (low mM kD’s individually) held together in a semi-rigid structure of specific dimensions such that it can span the core of a Holliday junction. The binding modules could target the minor or major grooves as well as intercalate into DNA. Assuming the binding constant is multiplicative, the drug would
bind poorly to duplex DNA but bind with pM affinity if all 4 modules bind to the HJ (Figure 3.33). This customized ligand could be designed to bind to the Open-X or to the Stacked-X structure. The Stacked-X structure is more compact and it has the unique feature of having the major and minor grooves line up on the quasi-continuous strands due to coaxial helical stacking. This allows for the design of a compact ligand which is important because the linker needs to be short enough to prevent duplex cross-linking. Figure 3.34 shows the topography of the grooves of the Stacked-X conformation from an X-ray crystal structure\textsuperscript{[30]}. The major groove has a more accessible surface than the minor groove in this conformation, and thus targeting the minor groove would require a more flexible ligand for this model (Figure 3.34). The model of the Open-X structure also has a unique topology\textsuperscript{[31]}; depending on what face is exposed each of the arms extend from the core with all exposed minor or major grooves (Figure 3.35). Ligands could be designed to bind to either face of this junction.

![Figure 3.33- Model of a Cruciform Specific Ligand](image)

**Figure 3.33- Model of a Cruciform Specific Ligand.** The ideal ligand would contain 4 weak binding modules held at an angle that matches the structure of the Stacked-X conformation. The central linker should be short enough to prevent DNA cross-linking in duplex DNA.
Figure 3.34- X-ray Crystal Structure of a Stacked-X Holliday Junction. The X-ray crystal structure of a tetrameric inverted repeat HJ is shown above\textsuperscript{[30]}. The structure is rotated to show the major groove face (A) and the minor groove face (B). The major groove face of the HJ has a more accessible topography (A) than that of the minor groove face (B).
Figure 3.35 - X-ray Structure of an Open-X Holliday Junction. A-D shows the X-ray crystal structure of the Open-X conformation of a HJ bound to the CRE recombinase protein[31] (protein hidden). Depending on how the structure is rotated, each face has the minor groove or the major groove exposed on each of the arms. This is shown in A and B in the cartoon form with exposed minor and major grooves respectively. This is shown again with a space filling model in C and D.
REFERENCES


Cruci2H (Bimolecular Cruciform)

(A) 5’-TCAGGACTCCGACTCGGTTTTCGCAATGCGCATGACTG-3’

(B) 5’-CAGTCGATGCGCTGACGCTTTTGCGTCAGCGGAGTCCTGA-3’

Cruci3HL (Intramolecular Cruciform)

5’-TCAGGACTCCGACTCGGTTTTCGCAATGCGCATGACTGTTTTCGTCAGCG
GAGTCCTGA-3’

Cruci3HL-Hairpin-Loop (This is the hairpin arm from Cruci3HL closest to the 5’ end)

5’-CGCACTCGGTTTTCGCAATGCG-3’

Cruci4S

5’-(TCGGTACCGA)4-3’

Junction3 (Bimolecular Cruciform)

(A) 5’-CAGTGACTTGAGTCTTTTCTAGGTTTTCTAGCAAGGGGCTGATGTC-3’

(B) 5’-GACATGCAGCTGAGCGGTCTTTTGACCGCTCAACTCAAGTCACTG-3’

Junction3BHairpin (Hairpin from strand B of Junction3)

5’-TGACCGGTCTTTTGACCGCTCA-3’

Junction3B-HP-3’-Overhang (J3BHairpin with an extra T on the 3’ end)

5’-TGACCGGTCTTTTGACCGCTCAT-3’

Junction3B-HP-5’-Overhang (J3BHairpin with an extra T on the 5’ end)

5’-TGACCGGTCTTTTGACCGCTCA-3’

Junction3B-HP-TT-Mismatch (J3BHairpin with an extra T on both the 5’ and 3’ end)

5’-TGACCGGTCTTTTGACCGCTCAT-3’

Junction3B PolyT (Hairpin from strand B of Junction3 with poly-T tails)
**Figure SI-1. DNA Sequences Studied**

All DNA sequences were ordered from Integrated DNA Technologies. Sequences were gel filtered and analyzed on a denaturing gel at high concentration <1 OD to ensure purity of the samples. Samples were quantified by measuring their absorbance at 260nm at 85 degrees on a Cary UV-spectrometer in the presence of 10mM sodium phosphate buffer.
Imino Protons in Nucleic Acid Chemistry

Figure SI-2. Imino Proton of Guanine
The proton circled above in red is the imino proton of the nucleobase guanine. An imino proton is defined as any proton attached to a nitrogen atom adjacent to a carbonyl group. Imino protons are located on guanines at the N1 position and on thymines at the N3 position.
1D NOE Experiments to Identify Nearby Protons Through Space

Figure SI-3. 1D NOE from Irradiating the T1 and A22 Base Pair Region
The imino proton peak from T1 and A22 was irradiated at 12.45 ppm. This didn’t cause any of the other imino proton resonances to drop in intensity (not shown). The only peak to be affected was at 7.815 ppm in the adenosine H-2 range.
Figure SI-4. 1D NOE from Irradiating the A15-T8 and A3-T20 Base Pair Region
(A.) The A15-T8 imino proton was irradiated at 13.9 ppm. This caused a drop in the A-H2 range at 8.0 ppm. (B.) The A3-T20 imino proton was irradiated at 14 ppm. This caused a drop in the A-H2 range at 7.775 ppm.

Ligand Titrations onto J3B-HP and J3B-Poly-T

100μM J3B Hairpin
288K
18.3mM NaCl 10mM NaP
pH 6.78

Figure SI-5. LH4.14 Titration onto J3B-Hairpin
100μM J3B-Poly T
288K
18.3 mM NaCl 10 mM NaP
pH 6.75

Figure SI-6. LH4.14 Titration onto J3B-PolyT
Figure SI-7. LH3.123 Titration onto J3B-HP

100μM J3B-Hairpin
288K
18.3mM NaCl 10mM NaP
pH 6.80

Figure SI-8. LH3.123 Titration onto J3B-PolyT

100μM J3B-Poly T
288K
18.3mM NaCl 10mM NaP
pH 6.78
Figure SI-9. LH4.50 Titration onto J3B-HP
Figure SI-10. LH4.50 Titration onto J3B-PolyT