Investigation of the Presence of a G-quadruplex In the Hoechst Dye Binding DNA Aptamer

by

Kyla Egerdeen

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Chemistry

Waterloo, Ontario, Canada 2022

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Aptamers are nucleic acid sequences that have been selected to bind to a ligand with high affinity and selectivity. Aptamers have great potential for use in various biotechnological and medical applications but understanding their structure is important to fully maximize their potential. In 2007 Sando *et al*. selected for an aptamer that targeted a tert-butyl (*tBu*) Hoechst dye developed in their lab. No formal studies were performed on the aptamer, but the team generated a predicted secondary structure that included a loop-stem element. When this project began, it was noted that the Hoechst DNA aptamer sequence was rich in guanines, creating the potential for the formation of a G-quadruplex structure. The research presented in this thesis aims to provide structural information regarding the Hoechst DNA aptamer.

In Chapter 2, various biomolecular methods were used to better characterize the Hoechst DNA aptamer and variations of the Hoechst DNA aptamer. First, computational methods were used to model potential G-quadruplexes. Then, fluorescence was used to prove that the aptamer was binding to its ligand and to determine dissociation constants (K_d). Next, native gels provided evidence that the Hoechst DNA aptamer was forming a multi-stranded structure. Finally, circular dichroism (CD) suggested that the Hoechst DNA aptamer formed a G-quadruplex structure.

In Chapter 3 nuclear magnetic resonance (NMR) was employed to gain structural information regarding the DNA aptamer bound to the *tBu* Hoechst dye. 1D Proton NMR experiments further proved that the DNA aptamer was binding to the *tBu* Hoechst dye, and further supported the potential presence of a G-quadruplex. Additionally, it was possible to tentatively assign the *tBu* peaks within the NMR spectra. Future work could provide additional structural information and could eventually lead to the elucidation of a solved structure for the DNA aptamer bound to the *tBu* Hoechst dye.

Acknowledgements

I would like to extend my deepest gratitude to Dr. Thorsten Dieckmann for all his guidance to make this thesis possible. His enthusiasm, knowledge, and encouragement made my graduate experience an overwhelmingly positive one.

Thank you to Dr. Elizabeth Meiering and Subha Kalyaanamoorthy for agreeing to sit on my committee, and for sharing their expertise with me. I would like to thank all members of the Dieckmann lab for their constant support. Specifically, I would like to highlight Dr. Kyle Piccolo for his guidance, patience, and friendship. I would also like to thank Avery To and the Murphy lab for their assistance in making the tert-butyl Hoechst dye, as without them I would not have been able to complete this research.

Finally, I would like to thank my family for their continued encouragement throughout my studies.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ATP	Adenosine triphosphate
AMP	Adenosine monophosphate
AT-rich	Adenosine-thymine rich
CD	Circular dichroism
cDNA	Complementary DNA
CNS	Crystallography & NMR System
COSY	Correlation spectroscopy
DEAE	Diethylaminomethyl
DFHBI	3,5-difluoro-4-hydroxybenzylidene imidazolinone
DFHO	3,5-difluoro-4-hydroxybenzylidene-imidazoli- none-2-oxime
DNA	Deoxyribonucleic acid
dsDNA	Double stranded DNA
Fmax	Maximum fluorescence
FPLC	Fast performance liquid chromatography
IDT	Integrated DNA technologies
K_d	Dissociation constant
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser enhancement
NOESY	Nuclear Overhauser enhancement spectroscopy
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
RFU	Relative fluorescence unit
RNA	Ribonucleic acid
RNase	Ribonuclease
ROE	Rotating frame Overhauser
ROESY	Rotating frame Overhauser effect spectroscopy
SELEX	Systematic evolution of ligands by exponential enrichment
t*Bu*	Tert-butyl
TMR3	Tetramethylrhodamine aptamer 3
TOCSY	Total correlation spectroscopy

Chapter 1   Literature Review

1.1   Structure and Functionality of Nucleic Acids

In 1970 Francis Crick theorized about "The Central Dogma" stating that nucleic acids could transfer important information to other nucleic acids or to proteins.[1,2] This led to the concept traditionally taught in biology classes; nucleic acids are carriers of information, where deoxyribonucleic acid (DNA) stores information before it is transcribed to ribonucleic acid (RNA).[3] In the past few decades there have been many advancements to expand this basic understanding of nucleic acids. One such advancement was an improved understanding of the structure of both RNA and DNA. There are three levels of nucleic acid structure: primary, secondary, and tertiary.[4] The primary structure refers to the sequence of nucleotides built from the three building blocks of nucleic acids; a phosphate group, a pentose sugar that is either ribose for RNA or deoxyribose for DNA, and a base.[3] The phosphate group and pentose sugars make up the backbone of the nucleic acid with the bond angles between the atoms in the backbone creating torsion angles.[5] The glycosidic bond formed between the base (N) and the sugar (C_1') is defined as the Chi angle, and can be classified as either *syn* or *anti* conformation based on the angle (Figure 1.1).[5]
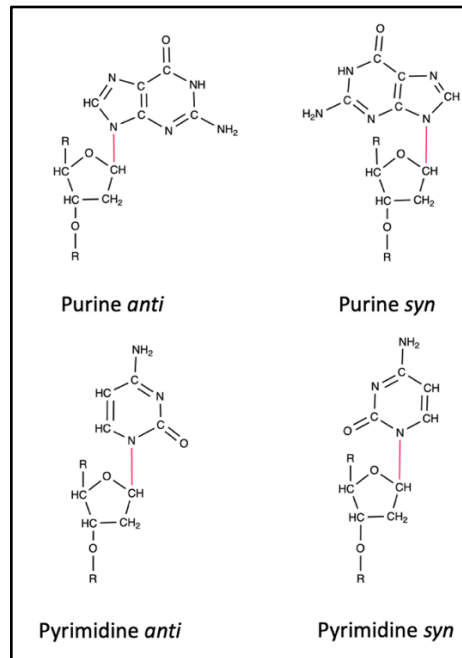
Figure 1.1 S*yn* and *anti* glycosidic angles in purines and pyrimidines.[5] Red bond indicates the Chi angle.

Nucleic acids form base pairing arrangements referred to as the nucleic acids secondary structure (Figure 1.2).[4,6] The secondary structures can then interact with each other to form tertiary structures.[4,7] Some examples of tertiary structures include triple helices created by a double helix interacting with a third strand, pseudoknots created by the loop of a hairpin interacting with a single strand, and kissing hairpins created by two hairpin loops interacting through base pairing (Figure 1.3).[7]
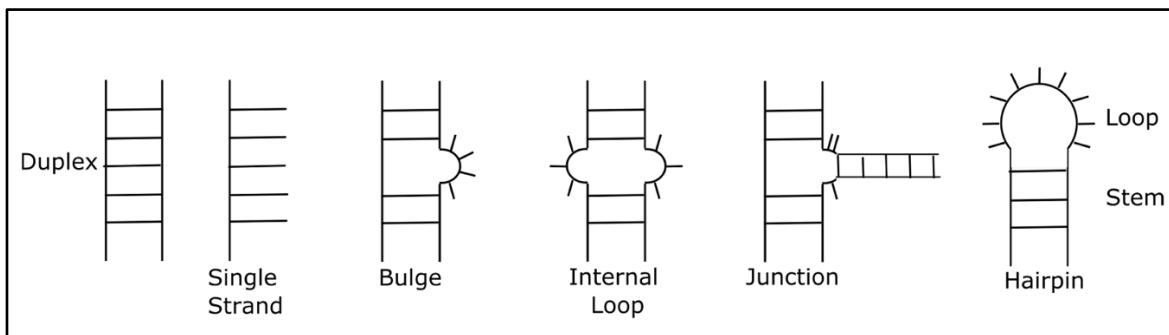
Figure 1.2 Examples of secondary structures found in nucleic acids. The examples include a duplex, a single strand, a bulge, an internal loop, a junction, and a stem and loop structure known as a hairpin.[4,7]
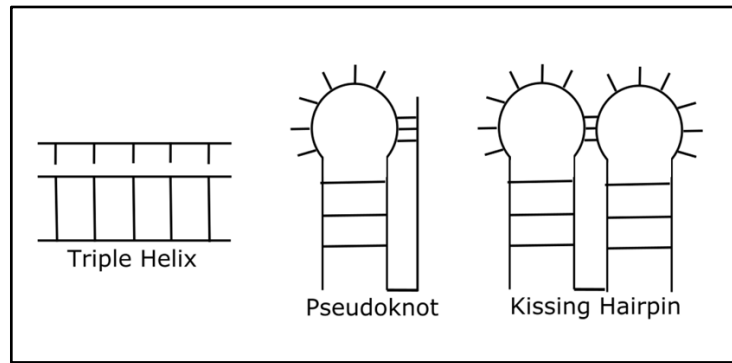
Figure 1.3 Examples of tertiary structures found in nucleic acids. The structures are formed through interactions of secondary structures and include a triple helix, a pseudoknot, and a kissing hairpin.[7]

Arguably, one of the most recognizable nucleic acid structures is the double stranded helical structure, or duplex structure. Within the classification of a double stranded helix, different forms of helices exist, with the most common being A-form RNA and B-form DNA.[7,8] The formation of a nucleic acid helix results in a minor groove and major groove, although dimensions of the grooves differ between forms (Figure 1.4).[8] These grooves are important as they allow the bases that are present inside the helix to be accessed by proteins, dyes, and other small ions.[3,8] A-form RNA has a narrow major groove that can limit the base's ability to interact with ligands.[7] B-form DNA has a wide major groove and a narrower minor groove.[5,7,8]
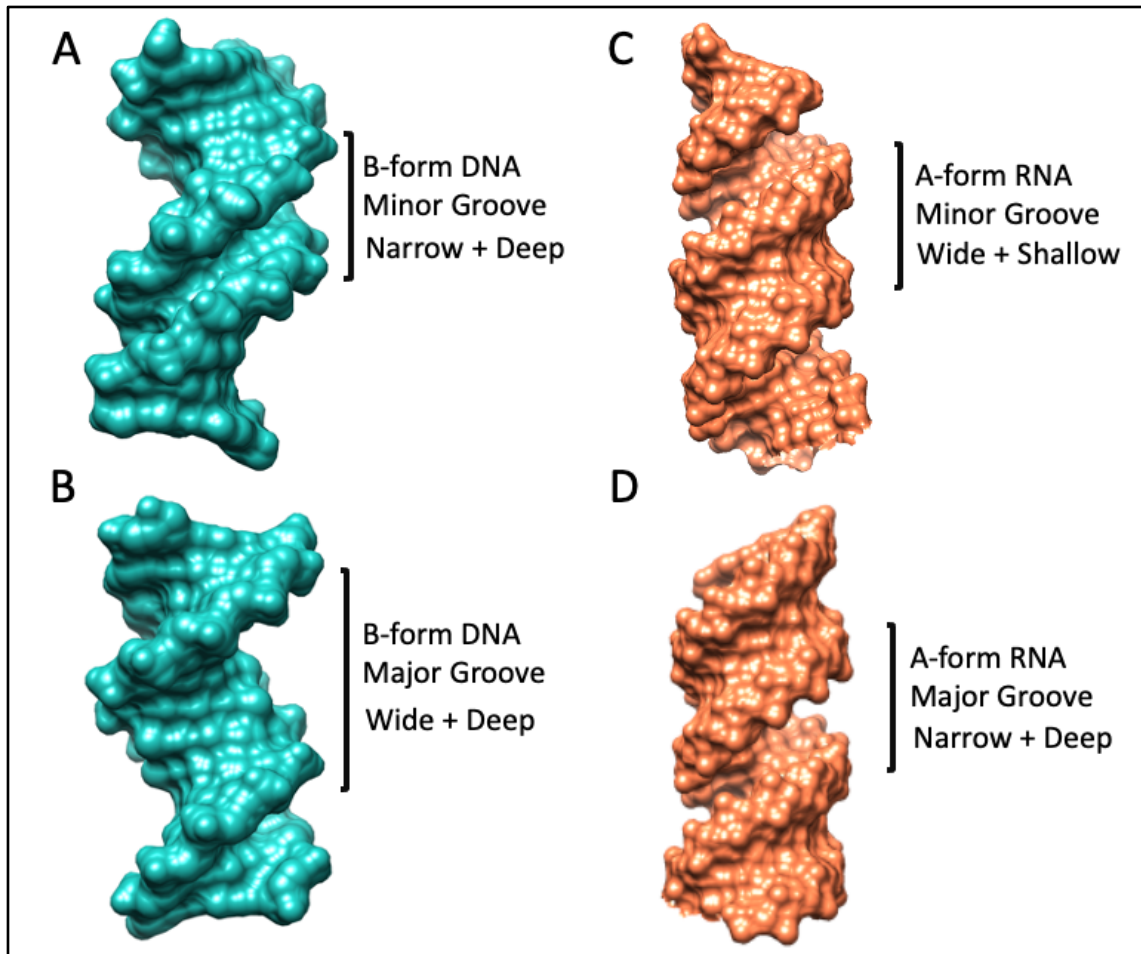
Figure 1.4 Grooves of common helices in nucleic acids. A) The narrow but deep minor groove of B-form DNA and B) the deep and wide major groove of B-form DNA (PDB:1BNA).[5,9] C) The shallow and wide minor groove of A-form RNA and D) the narrow and deep major groove of A-form RNA (PDB:1QC0).[5,10] Structures were rendered using Chimera.[11]

Advancements in understanding the structures of nucleic acids go hand in hand with advances made in understanding the expansive functionality of nucleic acids. In the 1980s Cech and Altman published separate works discovering the catalytic capabilities of RNA. Their work demonstrated that RNA has the catalytic ability to cleave RNA.[12,13] The splicing performed by the catalytic RNA was similar to protein enzymes and Cech decided to name the catalytic RNAs ribozymes to distinguish them. Research into different ribozymes has not slowed, leading to the recent discovery of a novel ribozyme in the human genome.[14]

1.2   Development of Aptamers and Their Applications

In 1990 two separate research groups, Jack Szostak's group at Harvard University and Larry Gold's group at University of Colorado Boulder, both published papers explaining their processes for developing nucleic acid sequences with high binding affinity to a target. Tuerk and Gold called their process "systematic evolution of ligands by exponential enrichment" or SELEX, whereas Ellington and Szostak called their process "*in vitro* selection".[15,16] Tuerk and Gold's research was focused on an RNA sequence that would bind to T4 DNA polymerase, whereas Ellington and Szostak's work focused on RNA aptamers binding to small dyes. Collectively, this research demonstrated the ability of modern science to select for nucleic acid sequences that target a specific ligand.

Despite the differing names, Tuerk and Gold's and Ellington and Szostak's processes follow similar steps (Figure 1.5). The processes begin with a large library pool of random oligonucleotide sequences.[15,16] The sequences consist of two constant regions that are typically 18 to 30 bases, required for the polymerase chain reaction (PCR) amplification, with a random region in the middle.[17] The random region can be quite varied in size, typically between 20 and 200 nucleotides.[17] The size of the starting library can vary but they often consist of 10^{15} to 10^{18} sequences.[16,18,19] The nucleotide library is combined with the target molecules.[15,16] Sequences with affinity for the molecule will bind, whereas sequences without affinity will not bind and are washed away.[15,16] Next the bound sequences are eluted from the target molecule.[15,16] For selections utilizing RNA the sequences are then reverse transcribed into complementary DNA (cDNA).[15,16] The DNA or cDNA is then amplified using polymerase chain reaction (PCR).[15,16] cDNA is then transcribed to generate RNA sequences.[15,16] The amplified sequences, either DNA or RNA, then act as the nucleic acid library and the selection process is repeated multiple times with selective

pressures increasing each round.[15,16] Once the desired number of selection rounds is completed, only high affinity binders for the target molecule remain.[15,16] The remaining binders are then sent for sequencing.[15,16] The term "aptamer", originating from the Latin word "aptus" meaning "to fit", was first used by Szostak's team to describe these high affinity nucleic acids and is now commonly used to describe the nucleic acids developed by this process.
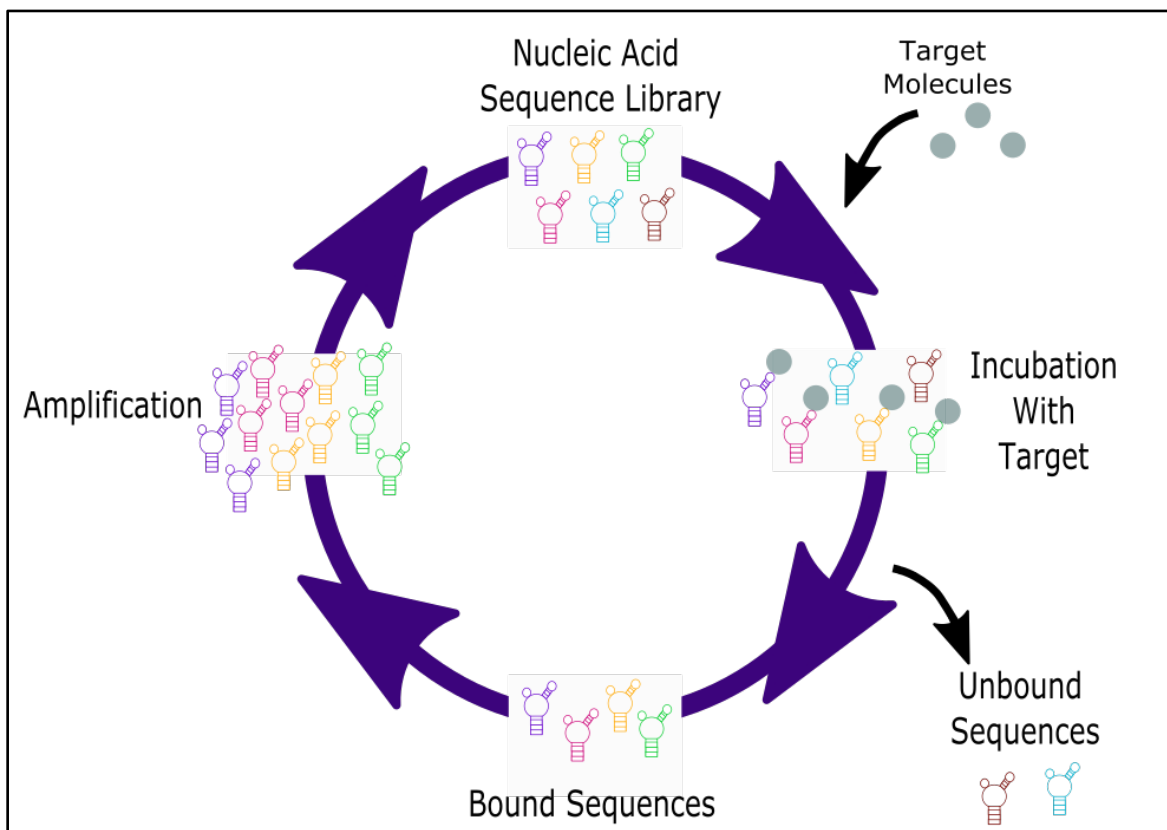
Figure 1.5 Schematic of SELEX process. First a library of random nucleic acid sequences are combined with the target molecule and incubated. The unbound sequences are then removed, and the bound sequences are separated from the target molecule. Next the bound sequences are amplified. Once the process has been repeated 6-12 times with increasingly selective conditions, the final library is sent for sequencing.

Since its discovery SELEX has been used to produce an assortment of RNA and DNA aptamers for a variety of targets.[20–26] The modern definition of aptamers has evolved to nucleic acids that can bind to target ligands with high affinity and specificity. SELEX's ability to select

for aptamers for a wide range of targets continues to create numerous possibilities for aptamer-based applications.

1.2.1 RNA and DNA Aptamer for the Same Ligand

As stated previously SELEX can be used to select for both DNA and RNA aptamers but performing two selections, one DNA and one RNA, for the same target ligand can provide an interesting model for comparing structures. One example of DNA and RNA aptamers that bind the same ligands is the adenosine triphosphate (ATP) aptamers. In 1993, Sassanfar and Szostak utilized SELEX to select for an RNA aptamer that would bind to ATP, adenosine monophosphate (AMP), and adenosine.[27] Two years later, Huizenga and Szostak published work in which they selected for a DNA aptamer that would target the same ligands.[28] Over the next few years further research was performed on both aptamers resulting in solved structures for the aptamers bound to AMP.[29,30]

Lin and Patel published the structure for the DNA aptamer one year after the solved RNA aptamer structure was released, allowing them to compare the two structures. Lin and Patel stated three notable differences between the DNA and RNA aptamers. The first difference was the number of AMP molecules bound per aptamer, with two equivalents of AMP binding to the DNA aptamer whereas the RNA is bound to only a single equivalent.[29,30] The aptamers also differ in the sequence of the binding sites.[29,30] The final difference was the secondary structure.[29,30] While the differences between the aptamers are noteworthy, the similarities may be arguably even more interesting. Both aptamers have a similar binding pocket topology with a binding recognition element involving a guanine and AMP mismatched pair (Figure 1.6).[29,30]
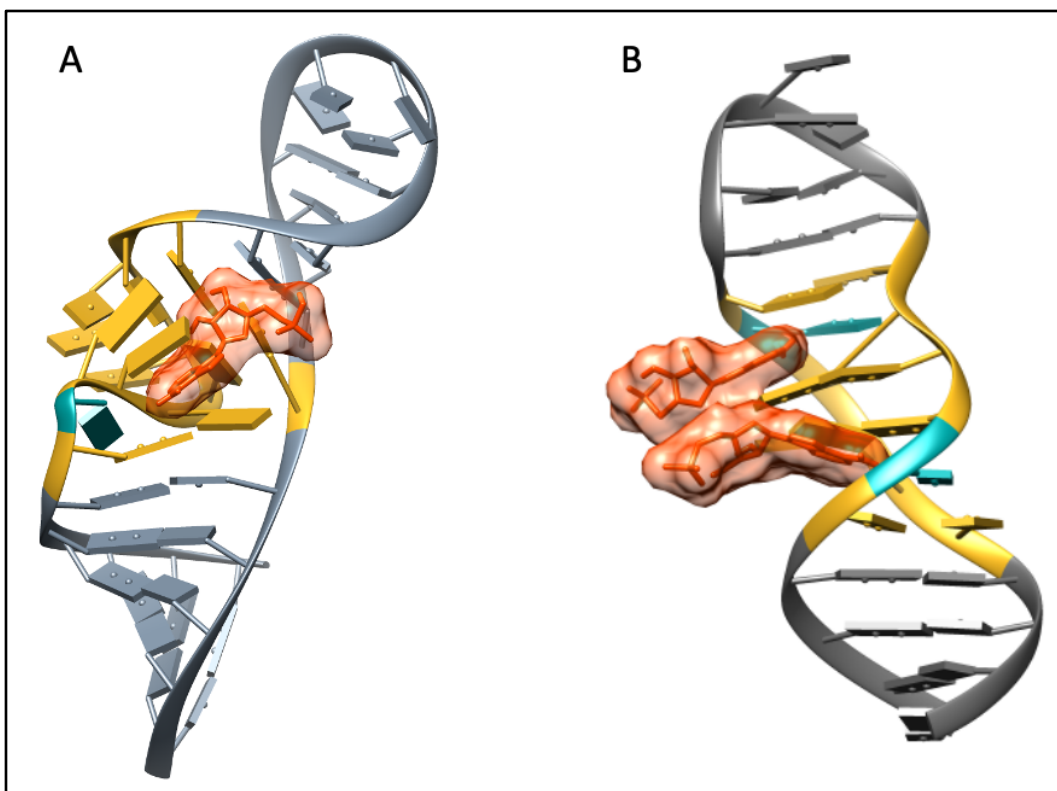
Figure 1.6 Structures of ATP binding aptamers. A) ATP-RNA aptamer (PBD: 1RAW).[29] The yellow bases represent the binding pocket with the single AMP shown in orange.[29] The blue nucleotide represents the G AMP mismatch pair.[29,30] B) ATP-DNA aptamer (PDB: 1AW4).[30] The yellow bases represent the binding pocket with the two AMP molecules shown in orange.[30] The blue nucleotides represents the G AMP mismatch pairs.[30] Crystal structures were rendered using Chimera software.[11]

The ATP aptamers share the same ligand and form a similar binding pocket but differ in other elements, such as the number of ligands binding, the secondary structure, and the sequence of the binding site, as described above. The evidence provides further testimony to the diverse nature of nucleic acids and further supports the need for structural studies. The ATP-aptamers are an interesting case study that provides evidence to support the usefulness of studying structures of multiple nucleic acid types that bind the same target.

1.2.2 Fluorophore binding Aptamers

There is a specific subclassification of aptamers known as fluorophore binding aptamers. As the name suggests fluorophore binding aptamers target fluorescent molecules.[20,21,31–33] Most small

molecule binding aptamers fold themselves around their ligands, creating a tertiary structure that incorporates the ligand, this is known as adaptive binding.[29,34–37] Currently, there are several fluorophore binding aptamers that have solved structures.

1.2.2.1 Spinach Aptamer

The spinach aptamer was originally selected to bind to 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI), a ligand that closely reassembles the fluorophore present in green fluorescent protein.[38] Upon binding a green fluorescent signal is produced.[38] In 2014 Warner *et al.* elucidated a crystal structure of the Spinach aptamer bound to the ligand. The structure showed that the ligand was bound in a site consisting of a G-quadruplex, a base triple, and a single unpaired guanine (Figure 1.7).[39] It was demonstrated that the G-quadruplex was likely required but other aspects of the structure could be altered.[39] In 2017 Fernandez-Millan *et al.* created what they described to be "an improved Spinach molecule" that they called iSpinach. iSpinach had various mutations that increased the brightness, improved thermal stability, and reduced the aptamer's sensitivity to salt.[40] iSpinach included five mutations and is a shorter version of Spinach. The mutations do not occur in the ligand binding site, therefore the team concluded the changes must be important to other aspects of the aptamer's folding.[40] A crystal structure was then determined for iSpinach.[40] Achieving the structure of the original spinach aptamer allowed for the development of the improved iSpinach, providing testimony to the importance of understanding the structure of aptamers.
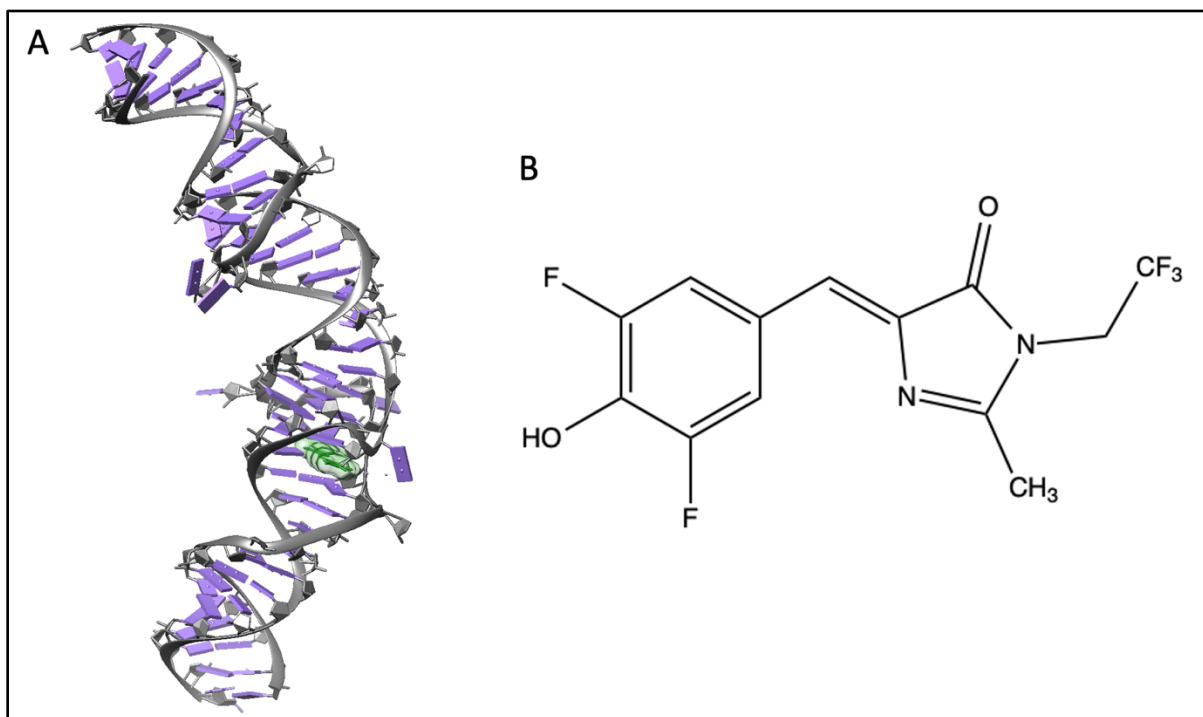
Figure 1.7 Original spinach aptamer structure. A) The original spinach aptamer with bases shown in purple, with DFHBI, green, in the binding site (PDB: 4TS2).[39] In the PBD file there were two additional DFHBI not bound in the binding site, these were removed to simplify the figure.[39] B) Chemical structure of DFHBI. Crystal structure was rendered using Chimera.[11]

1.2.2.2 Malachite Green Aptamer

In 1999 when Grate and Wilson first utilized SELEX to select for the malachite green aptamer they hoped to develop a gene targeting methodology.[26] Research for utilizing the aptamer for gene targeting was abandoned but further structural studies continued until eventually both a crystal structure and a solution structure were solved for the malachite green aptamer (Figure 1.8). The solution-structure shows the aptamer bound to malachite green,[41] whereas the crystal structure shows the aptamer bound to tetramethylrosamine.[42] Comparison of the structures showed one flexible binding site used by both aptamer conformations.[41] This binding site consists of a quadruple base pairing at the base of the binding pocket and various other pairs and triple base pairs to form an asymmetrical loop.[41,42]
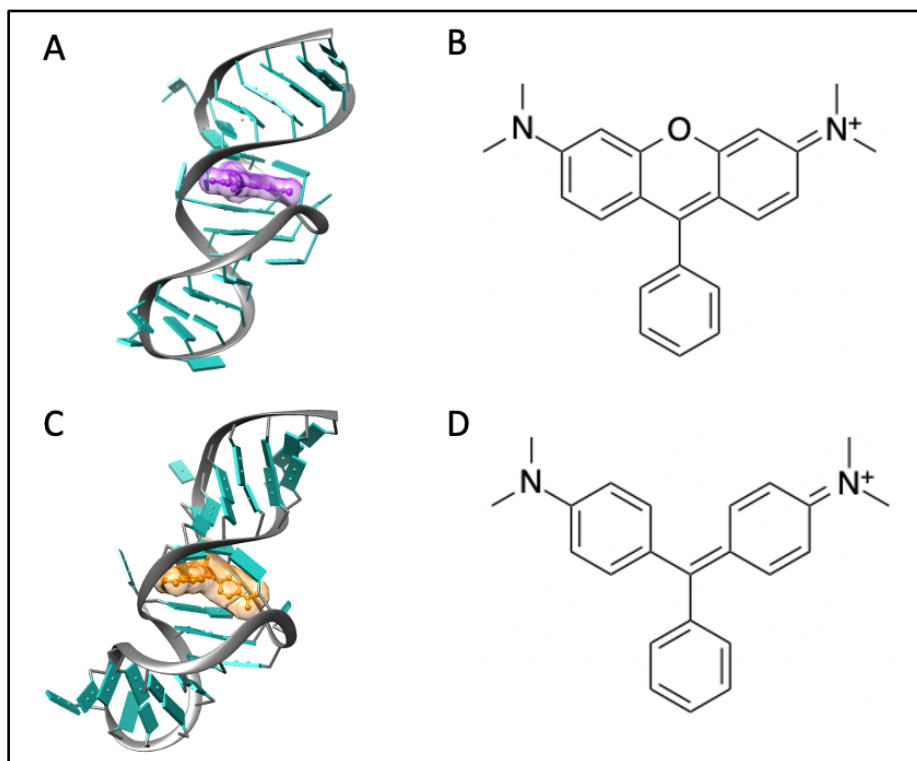
Figure 1.8 Malachite green aptamer structure. A) The malachite green aptamer, bases in teal, bound to tetramethylrosamine, purple (PDB: 1F1T).[42] B) Chemical structure of tetramethylrosamine. C) Solution structure of malachite green aptamer, bases in teal, bound to malachite green, orange (PDB: 1Q8N).[41] D) Chemical structure of malachite green. Crystal and solution structures were rendered using Chimera.[11]

1.2.2.3 Tertamethylrhodamine Aptamer 3

In 2020, Duchardt-Ferner *et al.* determined a solution structure for the tetramethylrhodamine aptamer 3 (TMR3) (Figure 1.9). Tetramethylrhodamine is a fluorescent dye that emits light around 570nm.[33] When TMR3 binds to the ligand the fluorescence emitted by the dye is quenched.[33] The solution structure demonstrated that the 48-nucleotide aptamer consists of three helical stems meeting in a junction.[33] Within the core of the junction sat the bound ligand.[33] Based on the three-stems determined in the solution structure Duchardt-Ferner *et al.* hypothesized that the size and sequence of the stem could be altered as long as the junction core is maintained.
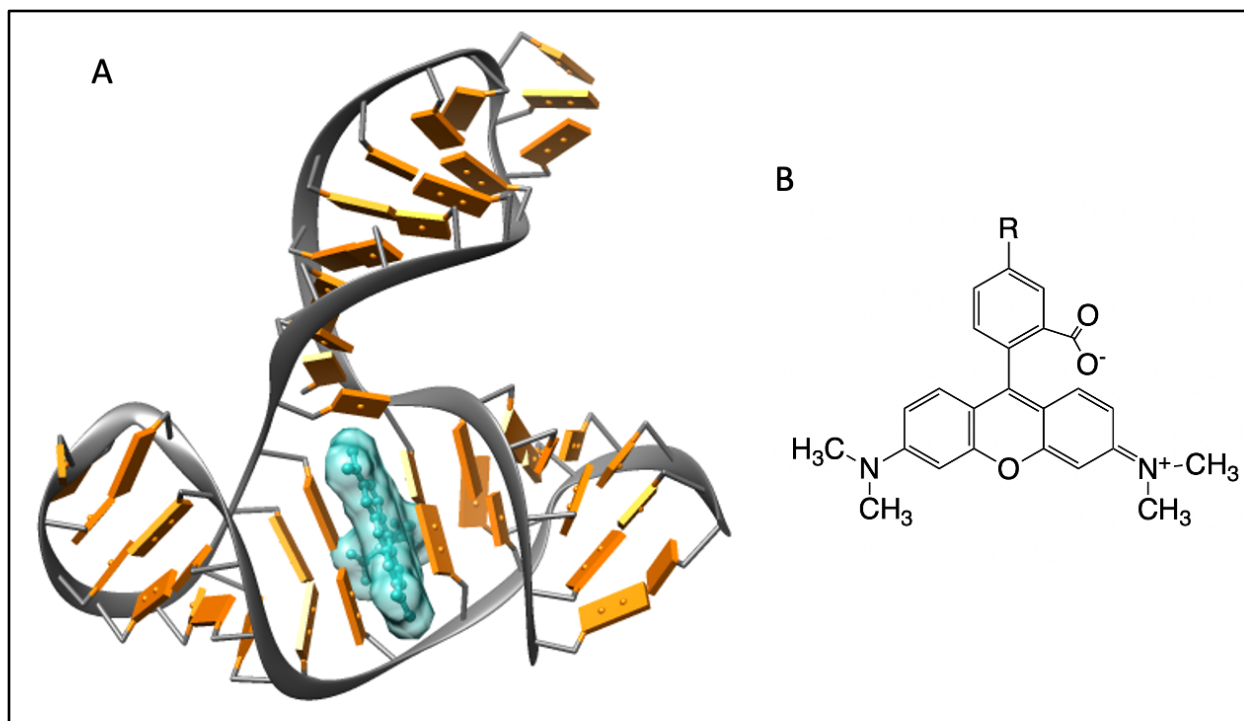
Figure 1.9 Tetramethylrhodamine aptamer 3 structure. A) TMR3, bases in orange, bound to tetramethylrhodamine, teal (PDB: 6GZR).[33] B) Chemical structure of tetramethylrhodamine. Solution NMR structure was rendered using Chimera software.[11]

1.2.2.4 Corn Aptamer

The corn aptamer is another example of a fluorophore binding aptamer with multiple solved structures. Originally in 2017 Warner *et al*. published a crystal structure of the corn aptamer bound to 3,5-difluoro-4-hydroxybenzylidene-imidazoli- none-2-oxime (DFHO).[43] Shortly after, in 2019, Sjekloca and Ferre-D'Amare published crystal structures for an aptamer bound to thiazole orange and an aptamer bound to thioflavin.[44] The corn aptamer is set apart from other aptamers as its structure consists of a homodimer (Figure 1.10).[43,44] In the corn aptamer the ligand is located between two G-quartets, one from each strand of DNA. The solved structure of the corn aptamer further demonstrates the diversity of aptamer-ligand complexes.
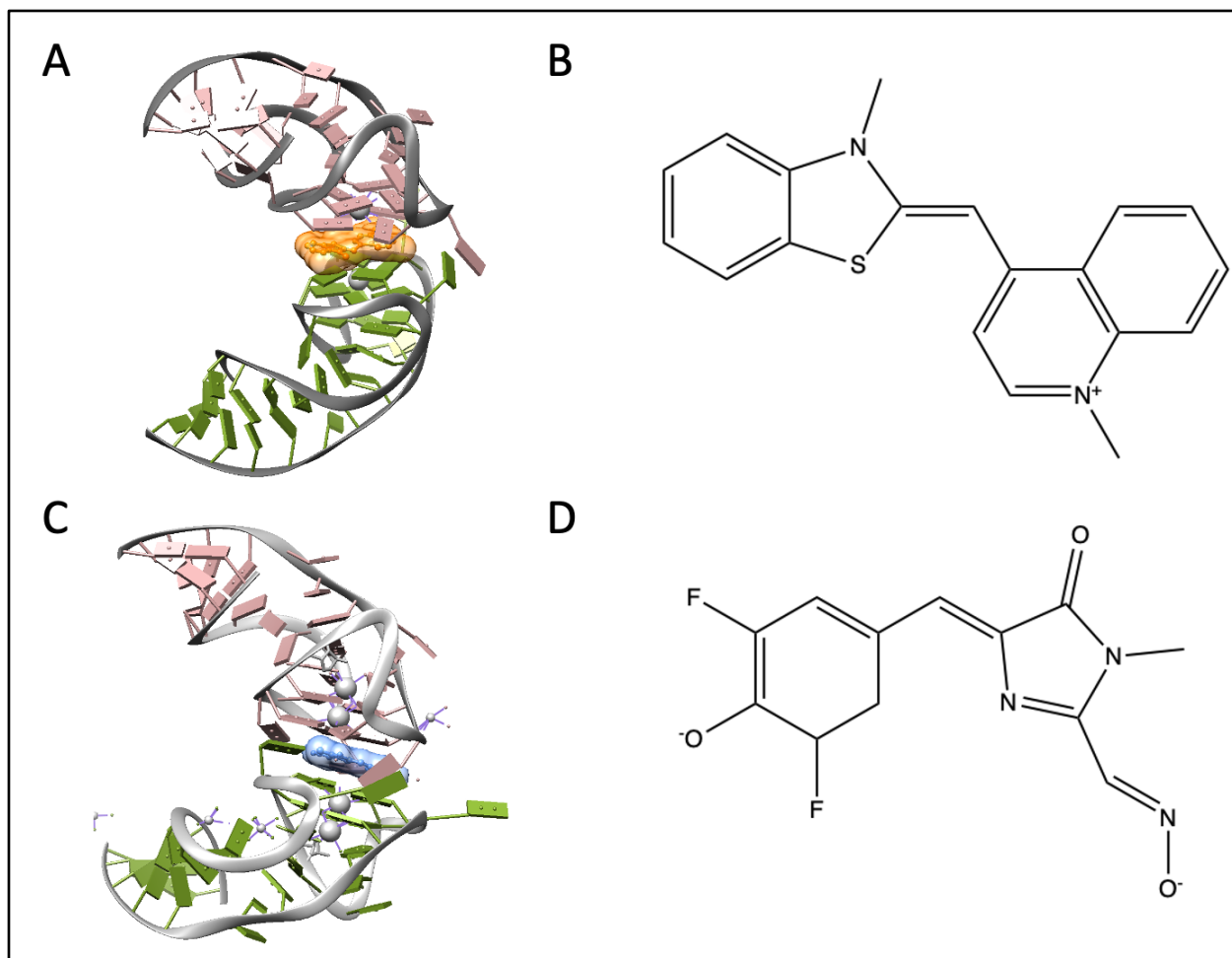
Figure 1.10 Structure of dimer corn aptamer complexed with various ligands. A) The corn aptamer, bases in green for one corn aptamer and pink for the second corn aptamer, bound to thiazole orange, orange (PDB: 6E84).[44] B) Chemical structure of thiazole orange. C) The corn aptamer, bases in green for one aptamer and pink for another, bound to DFHO, blue (PDB: 5BJO).[43] D) Chemical structure of DFHO. Solution NMR structures were rendered using Chimera software.[11]

1.2.3 Aptamers and Their Application in Biosensors

Aptamer's high affinity and selectivity for their targets make them excellent candidates for potential use in biosensors. Typically, biosensors contain two key parts: a biological recognition element, used to target a specific molecule, and a transducing element, used to create a detectable signal from a binding event.[45] Aptamers are good candidates for the biological recognition element required in a biosensor design. Aptamers are also appealing for use in biosensors as the process of

SELEX begins with a desired target and then the aptamer is developed for that target, creating the potential for endless aptamer ligand pairs. In addition, it is possible to add functional groups, such as a biotin molecule, to the ends of aptamers in order to immobilize them on surfaces such as beads or biochips.[22,46,47]

When designing a biosensor, it is important to consider the stability and by proxy the shelf life of the aptamer. DNA aptamers are considered to be highly chemically stable which is advantageous for biosensors, as it increases the shelf life of the aptamers and creates potential for the DNA aptamer to be reused many times.[22,24,48,49] In comparison RNA aptamers are less stable, are more likely to degrade, and are susceptible to ribonuclease (RNase) degradation, therefore it is recommended that RNA aptamers be used only for single short measurements.[50,51] But, the stability of RNA can be improved for use in biosensors through various methods. One method to reduce the likelihood of RNase degradation is with the addition of RNase inhibitors.[50] Modifying the 2'-OH of the ribose group, with either an amino group or a fluoro group, can also be used to reduce RNase degradation without affecting the RNA aptamers binding properties.[50–52] As long as one considers the stability of the aptamers and makes required modifications for RNA, both RNA and DNA aptamers have potential for utilization in commercial applications.

1.2.3.1 Examples of the Application of Fluorophore Binding Aptamers in Biosensors

Fluorophore binding aptamers can be particularly useful when engineering biosensors. As previously discussed, fluorophore binding aptamers bind to fluorescent molecules and upon forming an aptamer-ligand complex, produce a change in fluorescence intensity. The fluorescence produced when the aptamer binds to their ligand can be utilized to produce a detectable signal in a biosensor.

The malachite green aptamer is one example of a fluorophore binding aptamer that has been used in biosensors.[53,54] In 2004 Stojanovic and Kolpashchikov were the first to develop what they called an "modular aptameric sensor". The sensor utilized two different aptamers connected through a stem, the malachite green aptamer, to act as a signal, and a recognition aptamer, to detect the target.[53] The binding of the recognition aptamer to its target stabilizes the stem, resulting in stabilization of the malachite green aptamer-ligand complex to produce a detectable signal.[53] The sensor was successful in detecting flavin mononucleotide, adenosine, and theophyilline.[53] In 2010 Xu and Lu created a biosensor that utilized the malachite green aptamer and adenosine aptamer to detect adenosine.[54] This sensor begins with the two aptamers attached through a bridging strand to prevent the binding of malachite green.[54] Once the adenosine aptamer binds to its ligand, adenosine, the bridging strand is removed, and the malachite green aptamer can bind to malachite green resulting in a detectable increase in fluorescence signal.[54] Both biosensors demonstrate the potential to combine the malachite green aptamer with other aptamers to create label-free biosensors.

The Spinach Aptamer has also been utilized in biosensors. In 2015, Kellenberger *et al*. published a protocol wherein they described two methods, flow cytometry and fluorescence microscopy, in which the Spinach aptamer could be utilized for live cell imaging. Both methods utilized the original spinach aptamer and DFHBI to detect cyclic di-guanosine monophosphate levels.[55] The protocol was created with the hope that others could make minor adjustments to create different biosensors with the Spinach aptamer, or with other versions of the Spinach aptamer.[55] Despite the promising research into aptamer based biosensors, at the time of this thesis, research did not suggest any widely commercially available biosensors with this technology. The

general protocol developed by Kellenberger offers a glimpse into the ever-expanding potential of fluorophore binding aptamers.

1.3 Hoechst Dye

In the early 1970's Hoechst AG, a German chemicals company, developed fluorescent bisbenzimide dyes for the purpose of staining double stranded DNA.[56,57] Hoechst AG created multiple bisbenzimide dyes and named the products with the company name followed by the product number, for example Hoechst 33342.[56] On one terminal end of the Hoechst dye is a methyl piperazine ring that is connected to two linked benzimidazole rings, which are further connected to a phenol group on the other terminal end (Figure 1.11).[58] The variable R group on the phenol end of the dye differs between types of Hoechst dyes.

Hoechst Dye	R Group
Hoechst 33342	-OCH2CH3
Hoechst 33258	-OH
Hoechst 34580	-N(CH3)$_2$
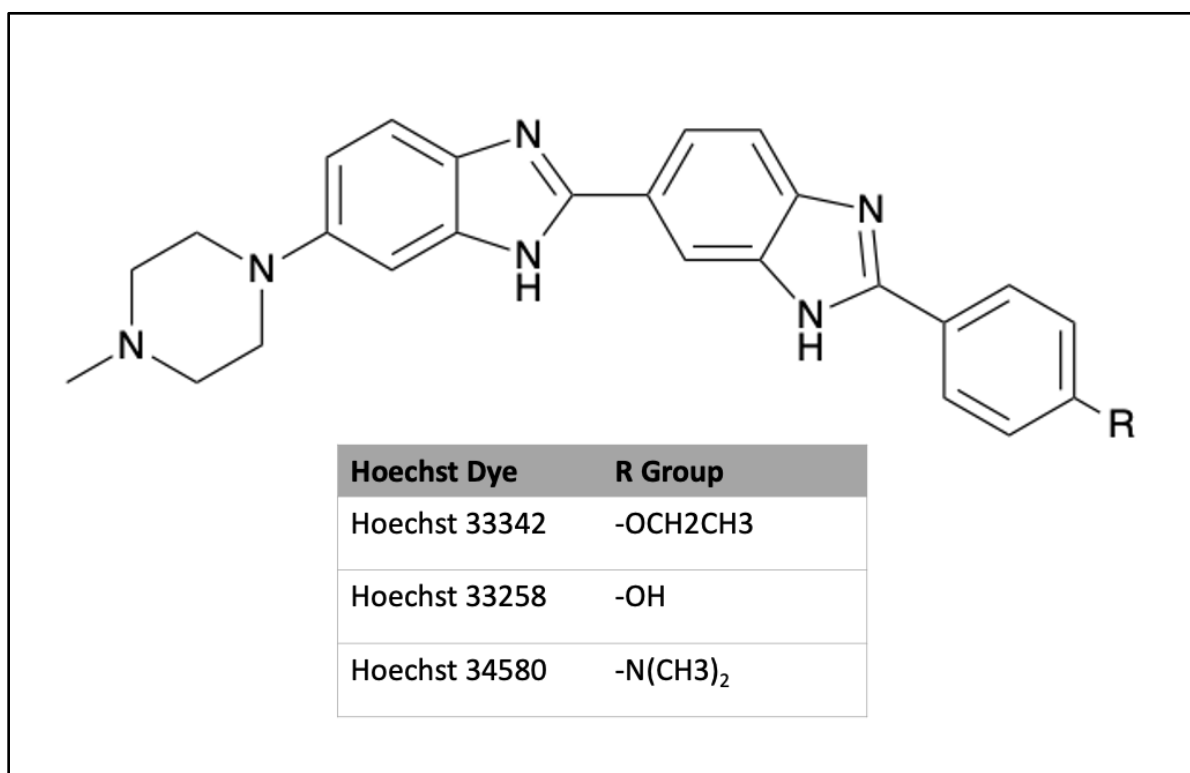
Figure 1.11 Hoechst dye chemical structures. The table represents the different R groups that can be added to the terminal phenol of the Hoechst dye to create different variations of the dye.[57]

Hoechst dyes bound to DNA are excited at a wavelength of 345 nm and emit a fluorescence maximum peak at approximately 460 nm.[56,59] Hoechst dyes target B-form, double stranded DNA (dsDNA) in adenosine-thymine rich (AT-rich) regions (Figure 1.12) and upon binding approximately 30-fold increase in fluorescence signal is observed.[56,59–62] B-form DNA is the most commonly occurring form of DNA therefore Hoechst dyes are quite versatile and not overly selective with regards to staining AT-rich dsDNA.[7,8,57,63] Despite the versatility with regards to DNA it should be noted that Hoechst dyes are not expected to bind to RNA.[63] Solved structures have improved the understanding of Hoechst dyes specificity towards AT-rich regions.[58,62,64] The thymine O2 and the adenine N3 (Figure 1.12B, Yellow) act to stabilize the complex through hydrogen bonds to the benzimidazoles NH groups (Figure 1.12B, Blue).[58,62,64] Adenine's H2 and the deoxyribose O4' have van der Waals interactions to the dye (Figure 1.12B, Green).[58,62,64] These interactions allow the dye to fit into the minor groove of the DNA (Figure 1.12).
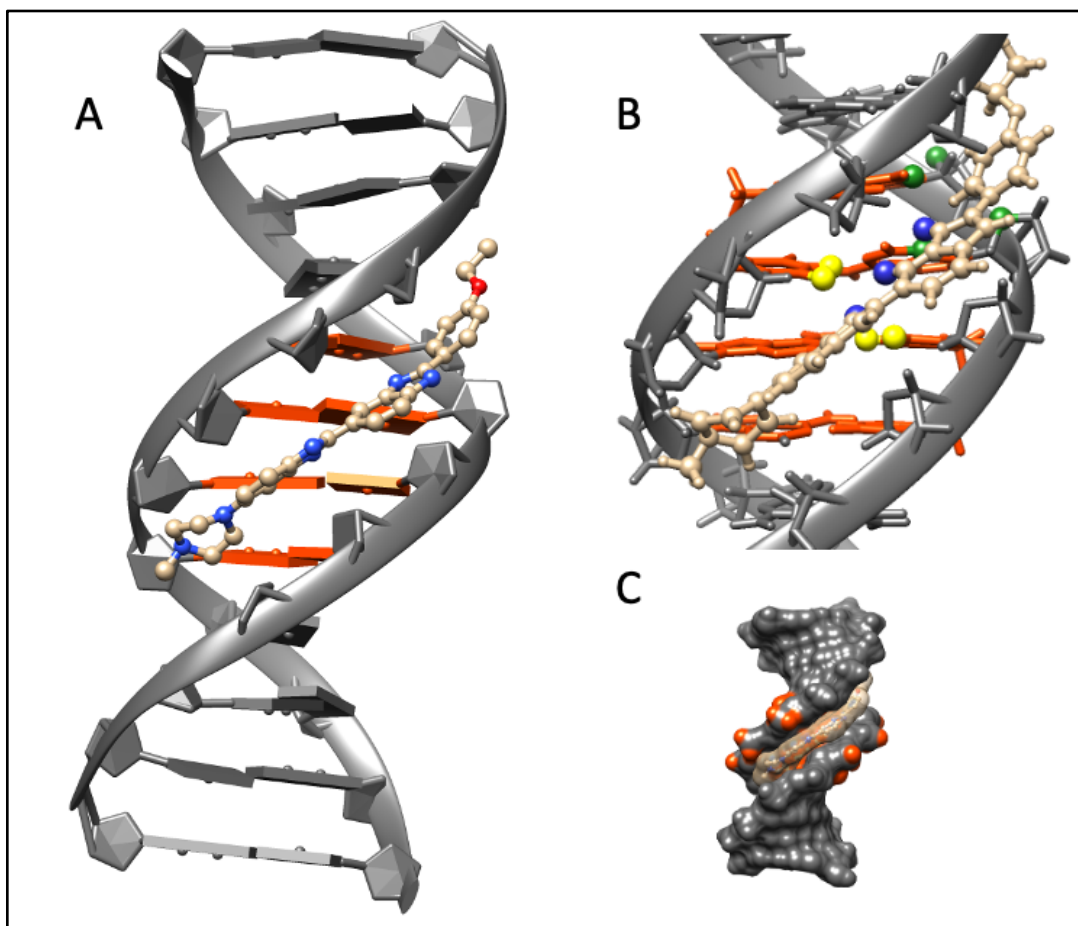
Figure 1.12 B-DNA bound to Hoechst 33342.[62] A) A dsDNA dodecamer bound to Hoechst dye 33342 (PDB: 129D).[62] Orange bases represent AT base pairs. Grey base pairs represent GC base pairs. B) Representation of the atoms interacting between DNA and dye. Green atoms show adenine 5 and adenine 6 H2 and O4'. Blue atoms show benzimidazoles. Yellow show thymine 7 and thymine 19 O2 and N3. C) Shows the helix with surface rendered showing the placement of the dye in the minor groove (PDB: 129D).[62] Structures were rendered using Chimera.[11]

In industry Hoechst dyes are commonly utilized in fluorescence microscopy, flow cell cytometry, and other cell staining applications for visualization of dsDNA.[65] Hoechst dyes are advantageous for experiments that aim to visualize DNA non-specifically, but these dyes can have limitations when it comes to detecting a more specific target.[63] Hoechst dyes have also been used in more novel experiments including their use in label-free aptamer based biosensors.[66–68]

1.3.1 Development of a tert-Butyl Hoechst Dye and Aptamers

The versatility of the original Hoechst dyes is useful for staining cells and for some biosensor applications, but a more selective dye creates potential for more diverse applications, including uses in different types of biosensors. Sando *et al*. began experimenting with altering the structure of the Hoechst dye to create a ligand with a more selective target.[20] The team developed a dye with a similar structure to Hoechst dye 33342, but they modified the terminal ring by adding two tert-butyl (t*Bu*) groups (Figure 1.13). The team tested the binding affinity of the t*Bu* dye through a melting temperature analysis, with a greater increase in melting temperature indicating binding of the dye to dsDNA. The melting temperature for the t*Bu* dye with dsDNA was altered by a very small value, 0.1°C, therefore it was concluded the t*Bu* dye was not binding to the dsDNA.[20] Sando *et al*. concluded that the addition of these bulky t*Bu* groups prevented the dye from fitting inside the helix of DNA and therefore the dye could no longer bind B-form DNA.[20]
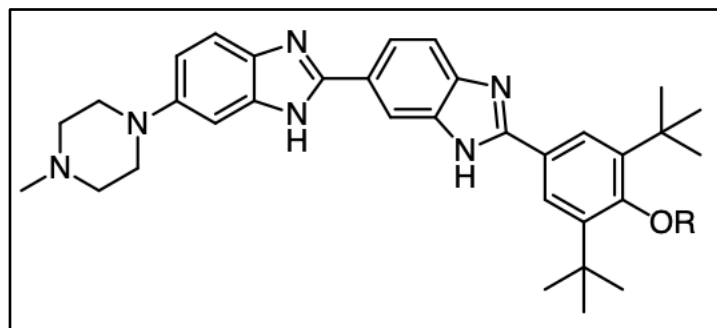
Figure 1.13 Chemical Structure of t*Bu* Hoechst dye. The dye developed by Sando *et al*.[20] The dye includes two t*Bu* groups added to the terminal end of the dye to prevent the dye from binding to dsDNA.[20]

Next, Sando *et al*. utilized SELEX to reselect for a DNA aptamer that would bind to the newly developed t*Bu* Hoechst dye.[20] The aptamer was 25 base pairs long and although the team produced a predicted secondary structure, no formal structural studies were performed on the aptamer. The predicted secondary structure developed by Sando *et al*. shows a AT-rich stem and a loop structure (Figure 1.14).[20] Sando *et al*. performed various modifications to the aptamer's stem and loop elements to observe the impact on ligand binding.[20] It was concluded that both

hypothesized structural elements were important for the binding of the ligand. They predicted that the AT-rich stem provides a binding site for the dye while the loop provides accommodation for the t*Bu* groups.[20]

```
            A   A   A
         T           G
      A                 G
   G                       G
   G                       G
     G                   G
       C               C
     T   T - A   C   C
         A - T
         A - T
         C - G
        5'      3'
```
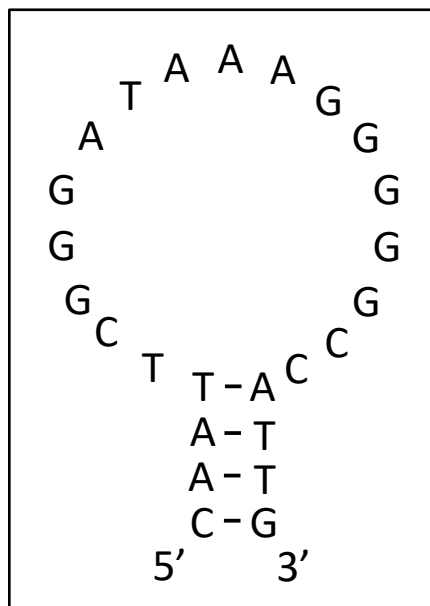
Figure 1.14 Predicted secondary structure of Hoechst DNA aptamer. The predicted structure includes two elements, a AT-rich stem and a loop structure. It is predicted that the stem is required to provide a binding site for the dye and the loop is required to provide space for the t*Bu* group. Figure modified from Sando *et al*.[20]

The following year Sando *et al*. published additional work wherein they used SELEX to reselect for an RNA aptamer that would bind to the same t*Bu* Hoechst dye.[21] The development of an RNA aptamer that binds to the t*Bu* Hoechst dye is particularly interesting as the unmodified dye targets DNA and typically does not bind to RNA.[63] In addition, the team had now created both a DNA and an RNA aptamer that would bind the same target. Sando *et al*. determined the dissociation constant (K_d) values for both aptamers and found the Hoechst RNA aptamer has a smaller K_d value compared to the Hoechst DNA aptamer, 35nM and 878nM respectively.[20,21] Similarly to the Hoechst DNA aptamer, Sando *et al*. developed a predicted secondary structure for the Hoechst RNA aptamer, but no formal structural studies were performed.[21] In DNA, the Hoechst

dyes bind to AT-rich regions therefore Sando *et al*. predicted if an RNA aptamer were able to bind to the dye it would need to contain an adenine-uracil rich stem.[21] However, based on the predicted secondary structure developed by Sando *et al*. this does not appear to be true.[21] The predicted structure for the RNA contains a head-stem and bulge structural element (Figure 1.15).[21] The prediction software used by Sando *et al*. was RA structure Version 4.4 program which is considered to be relatively accurate, but the predicted structure is not guaranteed.[69]
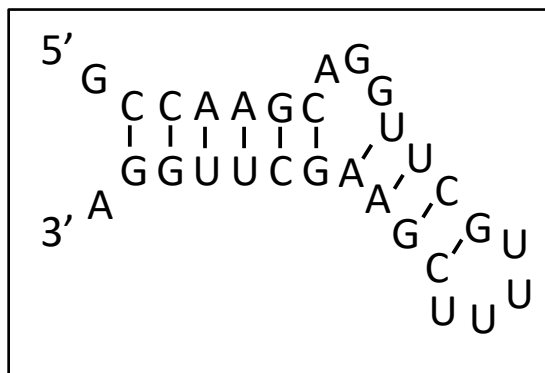
Figure 1.15 Predicted secondary structure of Hoechst RNA aptamer. The predicted structure shows a stem-loop region and a bulge element present in the RNA aptamer. Figure modified from Sando *et al*.[21]

As previously discussed, RNA and DNA aptamers have differing properties that can impact the applications of the aptamer. Creating a Hoechst dye that no longer binds to double stranded DNA but instead has both a DNA aptamer and an RNA aptamer, offers an interesting dynamic that can be further researched. Formal structural studies of the Hoechst t*Bu* aptamers bound to the ligand could aid in the deployment of the aptamers in practical applications. Based on the precedent set by the ATP aptamers (Section 1.2.1), investigation of both the DNA and RNA Hoechst aptamers could lead to interesting discoveries and a better understanding of how the ligands are bound and recognized by the aptamers.

1.4   G-quadruplexes

In the early 1960s it was observed that guanine-rich nucleic acids could form a gel, leading to the hypothesis of structures consisting of four hydrogen bonded guanines.[70] These four—stranded structures, seen in guanine rich nucleic acid sequences, would later become known as G-quadruplexes.[71–73] Since their discovery quadruplexes have been characterized in a variety of biological functions including telomeres,[71,73,74] DNA replication,[75] anti-cancer therapeutics,[76,77] and transcription.[78]

A tetrad consists of four bases hydrogen bonded to form a planar square (Figure 1.16).[71,72] Tetrads can form from one strand of nucleic acid or they may include up to four different guanine rich strands.[79,80] Tetrads that contain only one strand of nucleic acids create unimolecular quadruplexes, whereas tetrads with multiple strands can create dimeric, trimeric, or tetrameric quadruplexes.[79,80] Quadruplexes are formed when at least two tetrads stack, but quadruplexes can also consist of more tetrads.[71,72] Many structures of quadruplexes consist of GGGG tetrads, but quadruplexes can also include tetrads with alternative bases, such as GGGT tetrads (Figure 1.16).
[71,73,81–83]
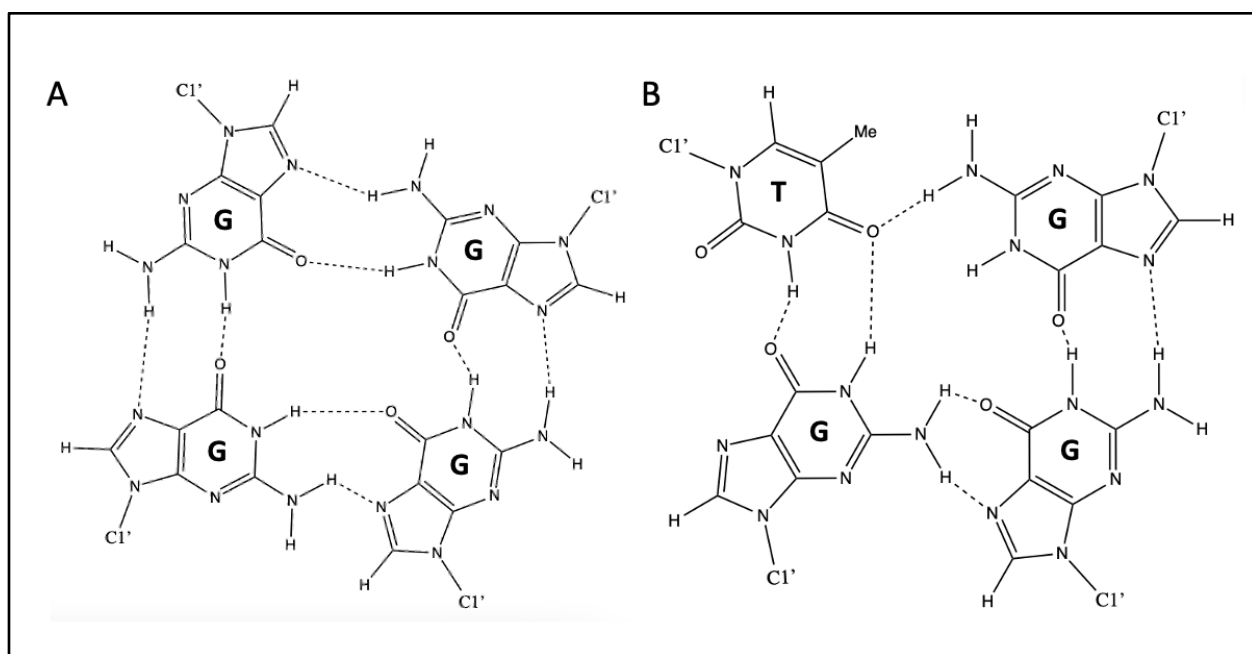
Figure 1.16 Examples of tetrads. A) GGGG tetrad.[82,83] B) GGGT tetrad.[82,83]

Varying tetrads is just one element leading to quadruplexes diversity. Quadruplexes can also differ in strand directionality and in glycosidic conformation, with both parameters being directly related to each other.[71,72,80,84,85] The strand directionality can be either parallel or anti parallel with the glycosidic conformation as either *syn* or *anti*.[71,72] To form quadruplexes additional base pairs must create "loops" linking the tetrads.[71,72] The loops can differ in length and sequence.[71,72] Quadruplexes can vary dramatically based on how the different elements described above are combined.

The formation of quadruplexes requires cations to help stabilize the structure.[72,86,87] The guanines carbonyl groups are positioned such that they are pointed to the inside of the tetrads.[72] Cations, often potassium or sodium, are needed to help stabilize the quadruplex.[72] The cations can be positioned between tetrads, or within the same plane as the tetrads.[72,88] Generally potassium ions are better for stabilizing the quadruplex.[89–91] When studying a G-quadruplex, designing experiments wherein the type of salt is altered, either potassium or sodium, can produce interesting dynamics to gain additional knowledge into the quadruplex.[89] The importance of cations to stabilize the quadruplex must be considered when designing experiments aiming to study such structures.

Another interesting advantage that G-quadruplexes have over other nucleic acid structures, is that G-quadruplexes are very thermodynamically stable.[92] This stability can be important for the binding of ligands to the aptamer.[92] The thermal stability of G-quadruplex also allows for experiments such as melting curves, wherein the type of salt present can be further varied for additional information.[93–95]

Cumulative research into quadruplexes has demonstrated that quadruplexes are incredibly diverse in many aspects including their applications, their strand orientation, the number of strands

involved, and their *syn*/*anti* conformation. The amount of diversity within quadruplexes can complicate structural prediction as there are many possible solutions to the quadruplex structures.

1.4.1 G-Quadruplex Aptamers

G-quadruplexes have some qualities that make them particularly valuable in aptamers applications. G-quadruplexes are very structurally stable, which can be beneficial for aptamers. Guanine-rich aptamers have been employed in a variety of applications.[96–98] One example is the thrombin-binding aptamer, which contains two stacked tetrads.[99] A key characteristic of G-quadruplexes is that they do not produce an immune response in patients, therefore, the thrombin-binding aptamer has potential therapeutic applications.[92,100] A second example of an aptamer containing a G-quadruplex is the insulin-binding aptamer.[96] In 2019 Wu *et al.* developed an aptasensor that could detect insulin using this aptamer.[96] This electrochemical aptamer based sensor utilizes a redox labelled insulin aptamer and upon binding of the aptamer to its ligand a change in structure of the aptamer results in a decreased electron transfer.[96] The stability of G-quadruplexes can play a key role in aptamer structure, stability, and their interactions with their ligands.

1.5   Research Objectives

Fluorophore binding aptamers are nucleic acids that have been selected to bind specifically to a fluorophore and produce an observable change in fluorescence.[20,21,31–33] Sando *et al.* developed a novel t*Bu* Hoechst dye that no longer bound double stranded DNA, and selected for two different nucleic acid aptamers, one DNA and one RNA, that would bind the t*Bu* Hoechst dye. Two different aptamers that both target the same t*Bu* Hoechst dye offer an interesting dynamic for research potential. Observing the similarities and differences between the two aptamer structures could

provide information regarding the binding of the dye. Both aptamers that were developed observed an increase in fluorescence intensity when bound to the t*Bu* Hoechst dye.[20,21] The published work focused on fluorescence studies and, although they presented potential structures for the aptamers, no formal structural studies were performed.[20,21] As stated previously, aptamers are able to perform adaptive binding wherein they incorporate their ligand into their tertiary structure.[29,34–36] Performing biochemical experiments in addition to structural studies will aid in better understanding of the Hoechst DNA aptamer bound to its ligand.

The research presented here aimed to characterize the binding of the Hoechst DNA aptamer to the t*Bu* Hoechst dye developed by Sando *et al*. A combination of biochemical methods and nuclear magnetic resonance (NMR) experiments were used to observe variations of the Hoechst DNA aptamer-ligand complex, with a goal to gain knowledge into the structure. The biochemical methods include fluorescence, to confirm binding, native gels to observe the approximate size, and circular dichroism to test for G-quadruplexes. NMR data acquired for the DNA aptamer bound to the t*Bu* Hoechst dye allowed for initial resonance assignments and modeling to provide a foundation for future studies. Future goals for this project would include obtaining a structure for the Hoechst DNA aptamer bound to the t*Bu* Hoechst dye. Other projects are focused on obtaining a structure of the Hoechst RNA aptamer, which would complement this research and allow for comparison of structures.

Chapter 2   Biomolecular Characterization

2.1 Introduction

Characterizing a biomolecular system requires intensive investigation from various avenues. Some characteristics that can be explored include the binding affinity for the ligand and the potential presence of structural elements. In 2007 Sando *et al*. hypothesized a stem and loop structure based on the sequence of the Hoechst DNA aptamer, but since no formal structural studies were performed, there was an opportunity to further explore this ligand-complex. The methods utilized in this chapter to characterize the Hoechst DNA aptamer include computational modeling, fluorescence experiments, native gels, and circular dichroism.

As discussed in Chapter 1, fluorophore binding aptamers are expected to produce a change in fluorescence upon binding to their target. This property can be incredibly useful for several applications, such as biosensors, but the change in fluorescence can also provide valuable information regarding the binding of the aptamer to its ligand. Based on the work published by Sando *et al*., it was expected that there would be an observable increase in fluorescence intensity at a wavelength of approximately 460nm when the aptamers are bound to the ligand.[20] The increase in fluorescence observed upon the addition of the aptamer to *tBu* Hoechst dye can be used to confirm the binding of the aptamer to ligand and to determine the K_d for the aptamer.[101]

When observing the sequence of the Hoechst DNA aptamer it was noted that many guanine bases were present. As discussed in Chapter 1, guanine rich sequences have the potential to form quadruplex structures. As the interest in quadruplexes grew, various computational methods were developed to aid in predicting the presence of a quadruplex within a given sequence.[102–104] ImGQfinder is one such software for predicting the presence of quadruplexes based on the sequence of a nucleic acid.[105] The ImGQfinder is unique as the software allows for "imperfect"

quadruplexes containing single defects, wherein a tetrad contains three guanines and another base (either thymine, cytosine, or adenosine).[105] Although ImGQfinder provides an excellent starting point for G-quadruplex prediction, this software only highlights the potential bases that may be involved in the quadruplex, and does not give information regarding the orientation of the quadruplex. As previously discussed in Chapter 1, G-quadruplexes can take on many diverse structures; therefore, the use of additional modeling software, such as Crystallography & NMR System (CNS), can aid in providing additional information regarding the folding of the quadruplex.

As the name suggests, CNS, was developed for structural determination of biomolecules from both NMR and crystallography research.[106,107] The software allows its user to input a nucleic acid sequence and various structural constraints, including data retrieved from NMR experiments, before the software returns a 3D structure output.[106,107] The software can also be utilized to model systems to gain further insights.[106,107] In this chapter CNS was utilized in combination with ImGQfinder to model potential structures of G-quadruplexes within the sequence. The models developed from the software were then used to guide further experiments.

Once a predicted structure of the aptamer is achieved through computational software, key bases within the predicted quadruplex can be substituted for inosine. Inosine substitutions can be beneficial for NMR of quadruplexes and can test the validity of the models, as will be discussed in Chapter 3. Inosine has a similar chemical structure to guanine and can often be incorporated into a structure seamlessly, but in other cases inosine substitutions have been observed to disrupt nucleic acid structures.[108,109] When performing inosine substitutions on a fluorophore binding aptamer, fluorescence emission scans can be used as an initial check for potential structural disruptions. Therefore, all inosine substituted samples were subjected to a fluorescence intensity scan prior to NMR studies.

When performing structural studies on DNA, it is important to consider the ability of nucleic acids to fold into a variety of multi-stranded structures, including dimers, trimers, or tetramers.[44,110] Native polyacrylamide gels can be utilized to compare the structures of different sequences, and may suggest whether a multistranded structure is present.[87,111] Native gels can provide insight into the charge, shape, and size of a biomolecule.[112–114] Importantly, native gels alone cannot confirm the exact number of strands in a nucleic acid structure, but they offer a comparison between the desired sample and a known ladder or other samples.[112–114] The clues obtained from native gels, in combination with other biomolecular methods, can provide great insight into the structure of an aptamer and its ligand.

Circular dichroism (CD) is an optical method used to gain information regarding conformation of nucleic acids in solution.[105,115,116] CD can give structural insights and can indicate the presence of different elements, such as a G-quadruplex or B-form DNA.[115–118] CD experiments have become common practice when studying G-quadruplexes as quadruplexes produce a recognizable output in CD.[77,105,110] Both parallel and anti-parallel G-quadruplexes produce a positive peak at 210nm, with the parallel quadruplex also having a positive peak at 260nm and the antiparallel quadruplexes having a negative peak at 260nm and a positive peak at 290nm.[115,117] CD is a method that works well in combination with other investigations as it can provide some conformational information.

In this chapter the original DNA aptamer and various alterations of the Hoechst DNA aptamer were characterized through a combination of biomolecular methods. Firstly, computational methods were used to predict the presence of a possible quadruplex. Next, fluorescence was used to confirm binding of sequences to the ligand, check for structural disruptions in inosine samples, and to determine K_d values. Native gels were used to predict the

presence of multimer structures. Finally, CD was used to investigate secondary structures of the aptamers and to check for the presence of quadruplexes. The biomolecular characterization performed in this chapter in combination with the NMR studies presented in Chapter 3, provided insight into the Hoechst DNA aptamer bound to its ligand.

2.2   Methods and Materials

2.2.1 Molecular Graphics

Molecular graphics created using UCSF Chimera software developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco with support from NIH P41-GM103311.[11]

2.2.2 G-quadruplex Structural Prediction and Modelling

ImGQfinder is an online search tool that allows users to input a sequence with parameters, and the software will output bases that could potentially be involved in a G-quadruplex.[105] ImGQfinder does not specify details regarding orientations or bond angles of the quadruplex, therefore CNS version 1.3 was used to model different variations of the quadruplex to determine if the structures were possible.[106,107] The URL for the CNS website is included in Appendix C.

CNS allows for additional restraints to guide the folding of the nucleic acid. For the quadruplex, hydrogen bonding restraints were added to mimic the GGGT quadruplex seen in literature.[82,83] Restraints were also added to create planarity within each of the tetrads. During NMR experiments, nuclear Overhauser enhancements (NOE's) are produced when protons are within ~5Å of each other.[119–121] Since CNS was created to calculate NMR structures it is possible to include NOE restraints within the model to specify distances between the protons. In the models presented here, NOE's were added to the model between the two stacked quadruplexes. The

hydrogen bond distances and NOE restraints were created based on a solved structure of a human telomeric G-quadruplex (PDB:2HY9).[122] Based on the input restraints, the CNS program simulates various annealing conditions, and if the structure is energetically possible, the software will output 2D coordinate set and the structures energy values (Figure 2.1).[106,107]
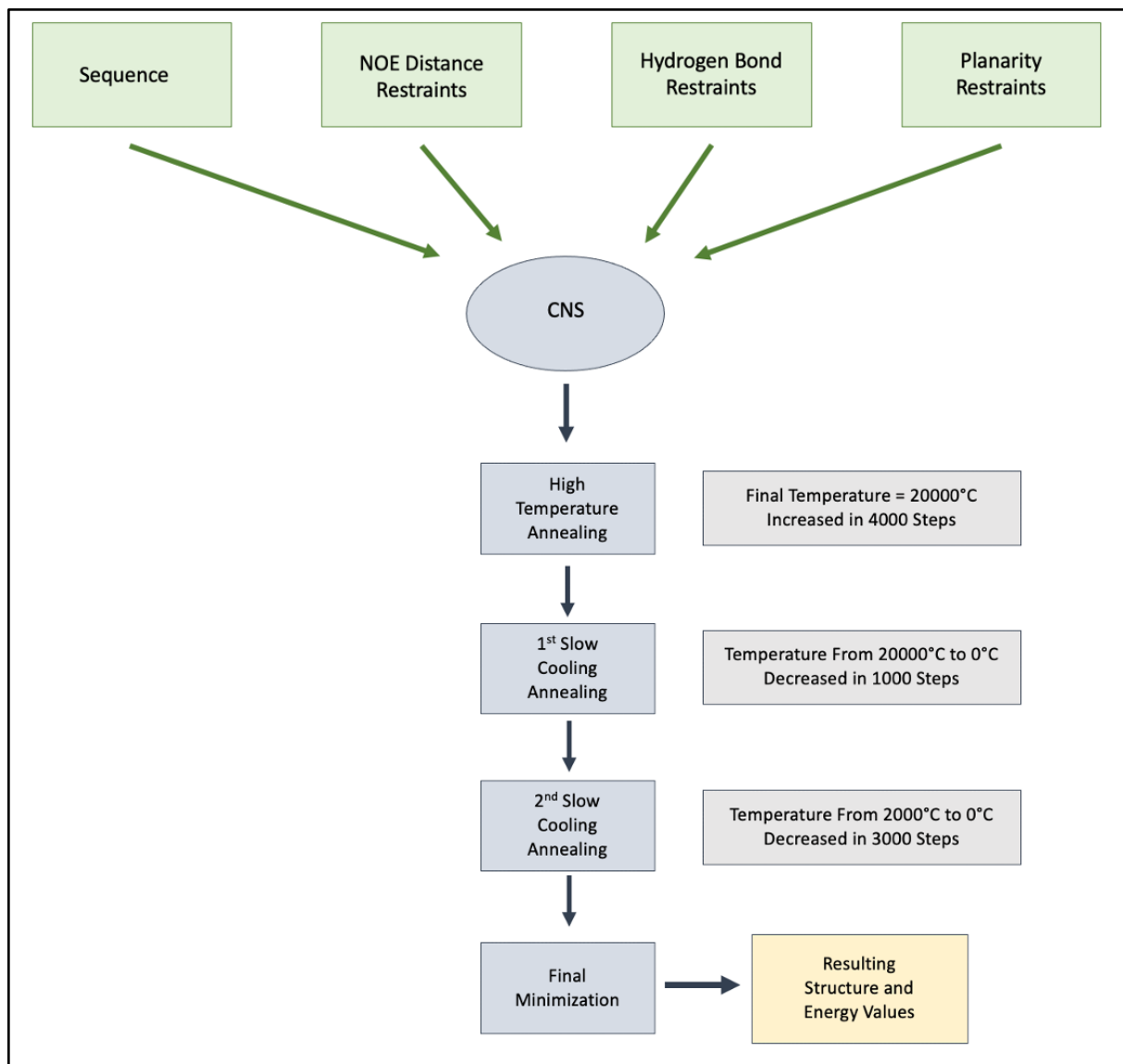
Figure 2.1 General flowchart of CNS. CNS accepts inputs including a sequence, distance restraints based on NOE data, hydrogen bond restraints, and planarity restraints. CNS then calculates structures through a series of annealing steps for the nucleic acid.

2.2.3 Purification of Nucleic Acids

For this project all DNA was custom synthesized by Integrated DNA Technologies (IDT) in Coralville Iowa. Various alterations of the 25nt DNA aptamer sequence developed by Sando *et al.* were used (Table 2.1).[20] All ordered DNA was desalted and deprotected by IDT.

Table 2.1 DNA Sequences Utilized in Project. Inosine substitutions represented as I.

Sequence Name	Sequence
Original DNA Aptamer	5' CAATTCGGGATAAAGGGGGCCATTG 3'
I_{16} DNA Aptamer	5' CAATTCGGGATAAAGI_{16}GGGCCATTG3'
I_{25} DNA Aptamer	5' CAATTCGGGATAAAGGGGGCCATTI_{25} 3'
Extended Stem Aptamer	5' GCCGCAATTCGGGATAAAGGGGGCCATTGCGGC 3'
I_{16} Extended Stem Aptamer	5' GCCGCAATTCGGGATAAAG I_{16}GGGCCATTGCGGC 3'

Once the nucleic acids were acquired the next step was to purify the samples using denaturing polyacrylamide gel electrophoresis (PAGE).[123] Urea was added to the nucleic acids for a final concentration of 5M. The DNA aptamer was heated at 90°C for 10 minutes, before it was run on a 15% gel.[123] The desired band on the gel was cut out and soaked in 300mM NaCl for 36 hours to extract the nucleic acid.[123]

The next step in the purification process of nucleic acids is to perform an ethanol precipitation for at least one hour at -20°C.[123] MgCl was added to the ethanol precipitation of the DNA, resulting in a final concentration of 10mM MgCl, to improve precipitation yield.[124,125] The sample was then centrifuged using the Sorvall Legend RT+ centrifuge (ThermoFisher Scientific, Waltham, MA) at a speed of 13200 rpm at 4°C for 45 minutes. The supernatant was removed, and the pellet was lyophilized before it was resuspended in water. Anion-exchange chromatography was then performed using fast performance liquid chromatography (FPLC) with the HiPrep 16/10

diethylaminomethyl (DEAE) FF anion-exchange column (GE Healthcare, Uppsala, Sweden). For the anion-exchange a low salt solution (150mM NaCl) and high salt solution (2M NaCl) were used. Another ethanol precipitation was then performed, followed by centrifugation, and the lyophilized pellet was again dissolved in water. The sample was then desalted with the FPLC using the HiPrep 5mL Desalting column (GE Healthcare, Uppsala, Sweden). The Nanodrop 2000 (ThermoFisher Scientific, Waltham, MA) was then used to determine the final nucleic acid yield.

2.2.4 Hoechst Dyes

The bisBenzimide H 33342, or Hoechst 33342, was purchased from BioVision (Milpitas, CA). The t*Bu* Hoechst dye was synthesized by Avery To in Dr. Graham Murphy's lab at the University of Waterloo. To improve the solubility of the t*Bu* Hoechst dye in water, a polyethylene glycol (PEG) tail was added to the C_1 of the phenol group (Figure 2.2). The PEG tail used in this project was selected due to its similarity to the PEG tail used during the aptamer selection by Sando et al.[20]
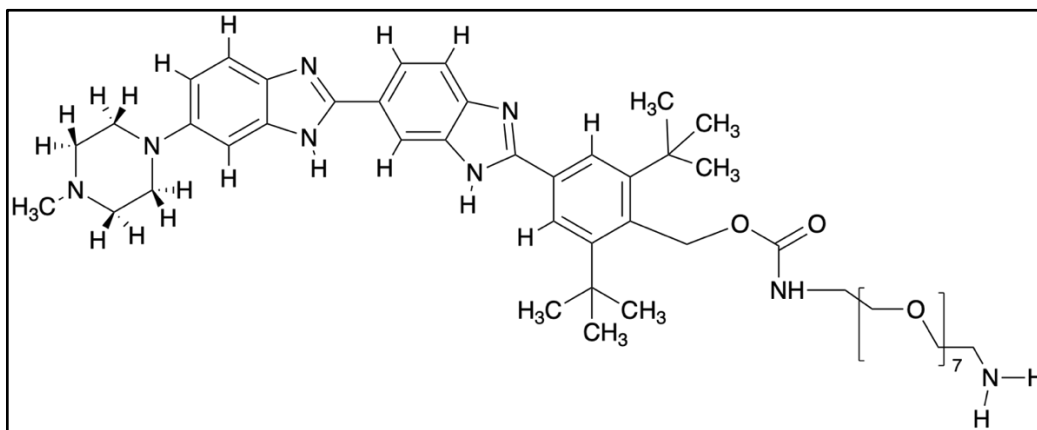
Figure 2.2 Chemical structure of t*Bu* Hoechst dye synthesized by Avery To.

2.2.5 Fluorescence Intensity

Samples were prepared for fluorescence intensity scans by combining 40μM of the nucleic acid aptamer, 10μM Hoechst dye, and a PBS buffer (pH=7.4) similar to the selection buffer used

by Sando *et al.*, consisting of 137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, and 2.5mM Mg^{+2}.[20] The fluorescence experiments were performed using Molecular Devices SpectraMax M5 microplate reader (Molecular Devices, San Jose, CA) and a Greiner Bio-One 96 well microplate (Greiner Bio-One, Kremsmünster, Upper Austria). The samples were excited by a wavelength of 345 nm and the emission was measured from 400 nm to 550 nm in 1 nm intervals.[56,59]

2.2.6 Fluorescence Titration Experiments

Fluorescence titration experiments were performed using Molecular Devices SpectraMax M5 microplate reader (Molecular Devices, San Jose, CA) and a Greiner Bio-One 96 well microplate (Greiner Bio-One, Kremsmünster, Upper Austria). The titration was performed with varying concentrations of DNA (0.4μM to 30μM), 1μM of t*Bu* Hoechst dye, and a PBS buffer (pH=7.4) consisting of 137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, and 2.5mM Mg^{+2}. The samples were excited at 345 nm, and endpoint readings were taken at 465 nm. The endpoint was utilized due to the maximum relative fluorescence (RFU) observed in the fluorescence scans. Three triplicates of the titration for each aptamer were run. The triplicates were used to calculate an average and standard deviation for each aptamer. The resulting data was then fitted using non-linear least squares regression with the equation

$\% Fmax = Fmax \times \frac{[Nucleic\ Acid]}{K_d+[Nucleic\ Acid]}$, where Fmax is the maximum fluorescence at a large concentration of the nucleic acid and K_d is the dissociation constant.[101]

2.2.7 Native PAGE Experiments

Native PAGE experiments were utilized to observe non-denatured nucleic acids. Approximately 0.7μg of the nucleic acids were loaded on a 15% PAGE with 50% glycerol.[123]

Samples that include dye have approximately a 1:1 ratio of dye to DNA. A Low Weight Molecular Weight DNA Ladder (New England Biolabs, Ipswich, MA) with a size range of 25 base pairs to 766 base pairs was utilized as a standard. The gel was directly transferred to the FluorChem FC2 Imaging System (Alpha Innotech, San Leandro, CA) and excited by an ultra-violet light to observe the Hoechst dye fluorescence. Another native gel was run with the same conditions, but instead of immediately viewing, the gel was placed in a container with a 1x stain solution, diluted in tris-borate-EDTA buffer from 10 000x SYBR Safe DNA gel stain (Invitrogen, Waltham, MA), and the container was wrapped in tinfoil.[126] The gel was then incubated with the SYBR safe solution for 30 minutes on a shaker at 50rpm.[126] The gel was then visualized using a green filter at 527nm on the FluorChem FC2 Imaging System and blue filter on the DRC III Fiber Optic Light Source Illuminator (Schott, Mainz, Germany).[126]

2.2.8 Circular Dichroism Experiments

Samples were prepared by combining 40μM relevant DNA, 50μM t*Bu* Hoechst dye, and PBS buffer (137mM NaCl, 2.7mM KCl, 10mM Na2HPO4, 1.8mM KH2PO4, pH 7.4) with 2.5mM Mg^{2+}. The samples were heated at 90°C for 5 minutes before cooling. The samples were then incubated at 4°C for at least 24hours prior to CD. CD experiments were performed using a Jasco J-815 spectropolarimeter (Jasco Inc., Easton, MD) and a 0.1cm pathlength quartz cuvette. A scan was taken from 320nm to 200nm, with a data interval of 0.5nm, response time of 1 second, band width of 0.5nm, and a scanning speed of 30nm minute[-1]. Four scans were accumulated. Baselines were established using two blanks: either a sample consisting of only buffer or a sample with buffer and 50μM t*Bu* Hoechst dye. The baselines were then subtracted from their corresponding data. The resulting data was smoothed in python using Scipy's Savitzky-Golay filter.[127,128] All raw data can be found in Appendix A.

2.3 Results and Discussion

Based on the guanine-rich sequence, it was important to address the potential presence of a G-quadruplex within the original DNA aptamer.[104] To investigate, the sequence was input into ImGQfinder's online software. ImGQfinder identified two possible G-quadruplexes in the Hoechst DNA aptamer, represented by the yellow highlight:

Option 1: 5' CAATTCG_7G_8GATAAA$G_{15}G_{16}$GG$_{18}G_{19}$CCATT$_{24}G_{25}$ 3'

Option 2: 5' CAATTCGG_8G_9ATAAA$G_{15}G_{16}$GG$_{18}G_{19}$CCATT$_{24}G_{25}$ 3'

Based on literature, three different, commonly occurring, single-stranded quadruplex orientations were modelled with in CNS (Figure 2.3).[79,80,84] ImGQfinder suggested two possible sets of bases for the G-quadruplex and literature provided three potential orientations, resulting in six possible quadruplexes to be tested in CNS.
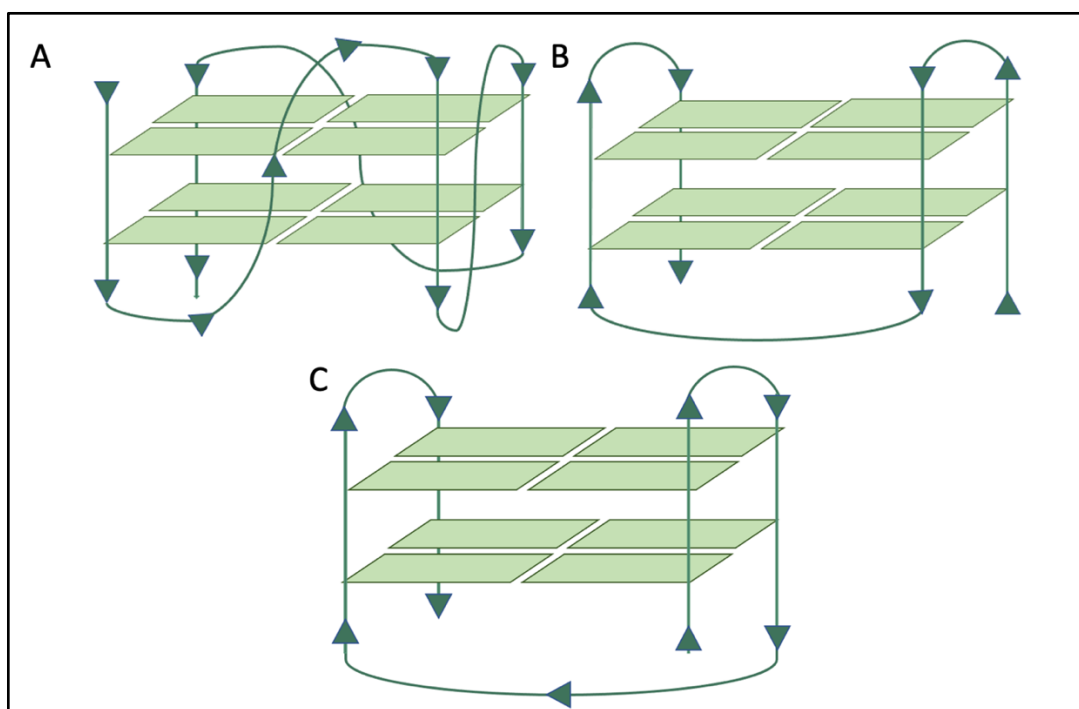
Figure 2.3 Possible architectures of G-quadruplex tested in CNS. Green arrows indicate the directionality of the nucleic acid strand from 5' to 3' and green boxes represent the bases. A) A

monomeric double chain reversal loop.[72,80] B) A monomeric edgewise loop.[72,80] C) A monomeric diagonal loop.[72,80]

Within CNS the three different quadruplex architectures determined from literature were modelled with each of the sequences suggested by ImGQfinder, resulting in a total of six models (Table 2.2). CNS was able to produce structures for all six of the quadruplexes indicating that each suggested quadruplex structure was theoretically possible. Additional modelling information presented in Appendix C.

Table 2.2 Quadruplexes Tested In CNS

Quadruplex Name	Quadruplex Bases	Quadruplex Orientation
Quadruplex A	G_7G_8 $G_{15}G_{16}$ $G_{18}G_{19}$ $T_{24}G_{25}$	Diagonal Loop
Quadruplex B	G_8G_9 $G_{15}G_{16}$ $G_{18}G_{19}$ $T_{24}G_{25}$	Diagonal Loop
Quadruplex C	G_7G_8 $G_{15}G_{16}$ $G_{18}G_{19}$ $T_{24}G_{25}$	Double Chain
Quadruplex D	G_8G_9 $G_{15}G_{16}$ $G_{18}G_{19}$ $T_{24}G_{25}$	Double Chain
Quadruplex E	G_7G_8 $G_{15}G_{16}$ $G_{18}G_{19}$ $T_{24}G_{25}$	Edgewise Loop
Quadruplex F	G_8G_9 $G_{15}G_{16}$ $G_{18}G_{19}$ $T_{24}G_{25}$	Edgewise Loop

After performing a structure calculation, CNS outputs required energies for various elements such as bond energy and NOE energy. A table with all the energies is included in Appendix D. The total energy for each structure was compared using a heatmap (Figure 2.4). The red, orange, and yellow colours for Quadruplex A and B structures indicate they have the highest energy, whereas Quadruplex E and F generally had lower energies, represented by green colours. The results from the models suggest that Quadruplex E and F are the most likely configurations. In addition to the energies, CNS outputs coordinates for the quadruplexes in a PDB file that can be visualized in Chimera, allowing for a better understanding of the predicted structure (Figure 2.5). Although the modelling indicates that Quadruplex E and F are most likely, it does not confirm

that the quadruplexes are present. If CNS was able to produce a structure, it was concluded that

the quadruplex was possible, and the next step was to apply information gained to experimental

methods.
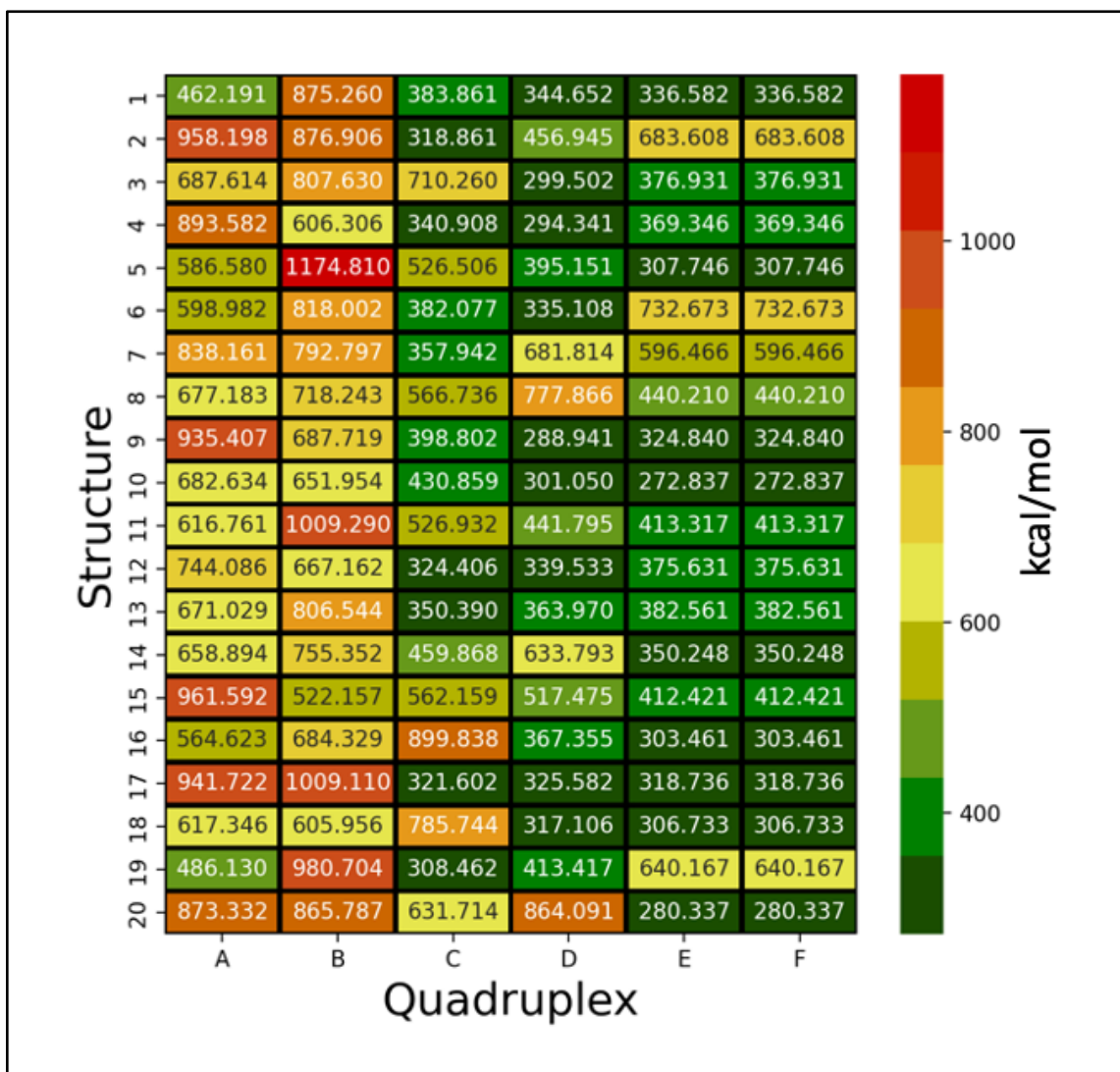
Figure 2.4 Heatmap for Quadruplex Energies. The energies in the table are total energy (kcal/mol) that are summed from energies for bonds, angles, improper dihedral, van der Waals, and NOEs. Red represents the quadruplexes that require the most energy. The high energy indicated more strain on the bonds and therefore these structures are less likely to form. As the energy decreases the colour changes to green, with the darkest green indicating the lowest energy.
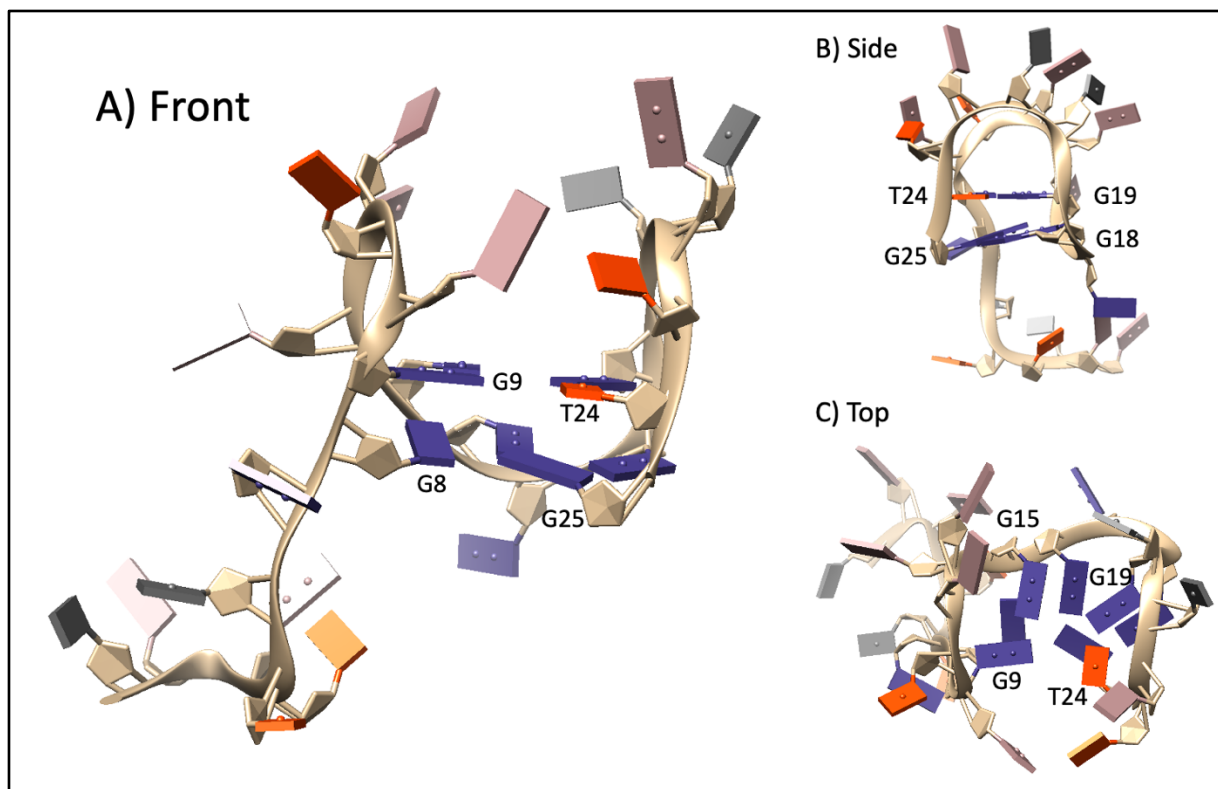
Figure 2.5 PDB Output of Quadruplex F. From the 20 structures calculated by CNS, Quadruplex F Structure 10 was included as a representation of the quadruplex based on its low energy. Thymines are represented with orange bases, guanines are dark blue, adenines are pink, and cytosines are grey. A) Front view. B) Side view. C) Top view. Structures were rendered using Chimera.[11]

2.3.1 Generation of Inosine Substitution in the Hoechst DNA Aptamer

If there is a G-quadruplex present in the Hoechst DNA aptamer inosine substitutions may provide additional information, regarding the orientation of bases within the quadruplex. Literature and software suggest that there are multiple orientations that the quadruplex could adopt. [71,72,79–83,105–107] Single inosine substitutions could provide important information if an appropriate guanine is substituted. An elimination process was then used to determine the guanines that were most suited for inosine substitution (Figure 2.6). ImGQfinder provided an output showing potential quadruplexes each consisting of two quadruples, eight bases total (Figure 2.6: Step 1). The

software suggested two possible sets of eight bases, resulting in nine potential bases that may be involved in the quadruplex. Based on this information guanine 7 and guanine 9 were eliminated as possible inosine substitutions as they may or may not be involved in the quadruplex, although they may be interesting for future studies (Figure 2.6: Step 2). One of the bases that may potentially be involved in the quadruplex is a thymine, that is not suitable for the inosine substitution as thymine is a pyrimidine and its structure varies too greatly from inosines (Figure 2.6: Step 3). The remaining six bases were then assessed based on the CNS generated structures.

To optimize the NMR results acquired from a single inosine substitution, the substitution should produce a quadruple with the H2 of the inosine pointed towards the N_8 of another guanine (Figure 2.7). Therefore, any inosine substitutions that may result in the H2 pointed to the thymine were eliminated. Combining the previous eliminations with the structures generated in CNS, two guanines, G_{16} and G_{25}, were selected as possible candidates for inosine substitutions because in every model, substituting either G_{16} or G_{25} would result in H2 of the inosine pointing to a N_8 of a guanine (Figure 2.6: Step 4). Once the bases were selected for inosine substitutions the next steps were to perform fluorescence emission scans to ensure the inosine was not disrupting the structure, which will be discussed in Section 2.3.3.2.
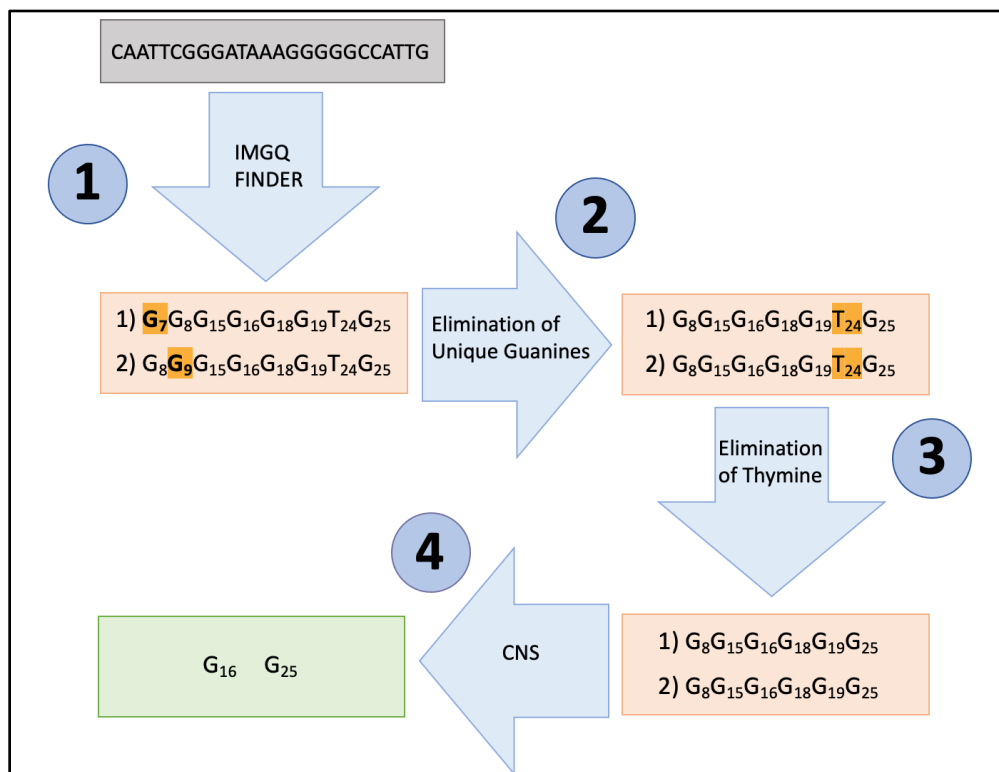
Figure 2.6 Flow chart for development of inosine samples. The original sequence (grey box) was input into ImGQFinder (Step 1). ImGQfinder produced two different sequences, with two unique bases (G_7 and G_9 highlighted in orange). Once the unique bases were eliminated (Step 2), T_{24} could also be eliminated as it was not suitable for substitution (Step 3). Based on the structures produced by CNS it was possible to narrow down to two bases (Step 4) that would be suitable for inosine substitution: G_{16} and G_{25}.
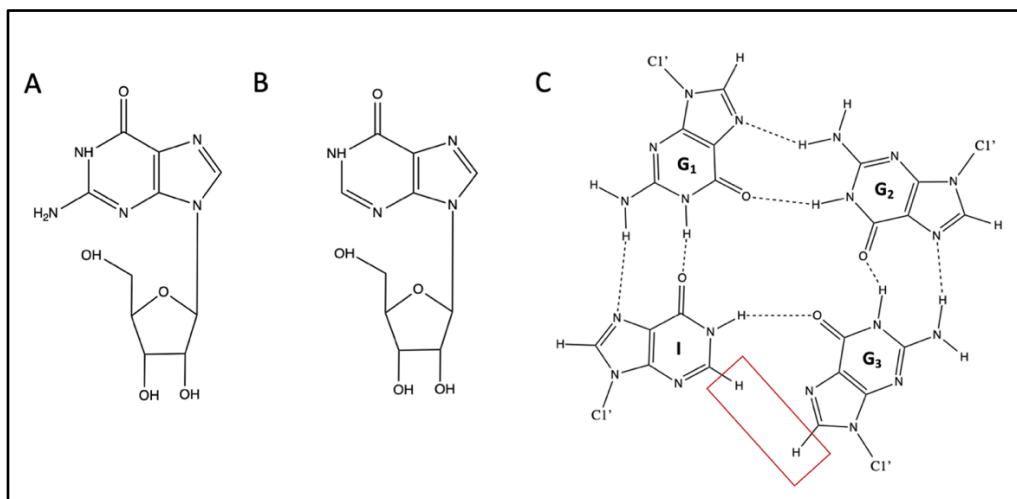
Figure 2.7 Ideal Inosine Substitution. A) Structure of a guanine base. B) Structure of inosine. C) An ideal inosine substitution with the red box indicating how the H2 of the inosine (I) is pointed towards a guanines N_8 (G_3).

2.3.2 Generation of an Extended Stem Aptamer

Following the selection of the Hoechst DNA aptamer, Sando *et al.* generated a predicted secondary structure that consisted of a four base pair stem with a loop.[20] Based on this predicted secondary structure, Sando *et al.* performed modifications to what they predicted to be a stem.[20] All the modifications presented produced a decrease in relative fluorescence enhancement, further supporting the team's predicted structure.[20] While the changes in relative fluorescence are consistent with the stem and loop structure proposed by Sando *et al.* in 2007, the changes to the sequence could also be consistent with a disruption of the single-stranded quadruplex structure predicted in this chapter.

First, Sando *et al.* performed a series of sequence modifications on the eight bases believed to make up the four base pair stem.[20] Two of these modifications included the addition of extra base pairs in the stem (Figure 2.8A and B).[20] Additional base pairs within the stem may increase the stability of the stem and therefore prevent the formation of the G-quadruplex. Another modification to the sequence included changing T_{24} to an adenosine (Figure 2.8C).[20] Again, this alteration would prevent the formation of the quadruplex as T_{24} is believed to be directly involved in the quadruplex. The final stem alteration removed bases 1 through 4 to remove one side of the stem (Figure 2.8D).[20] While this alteration is not as obviously related to the potential G-quadruplex, it may be possible that these bases play a role in guiding the secondary structure formation or are involved elsewhere in the structure.
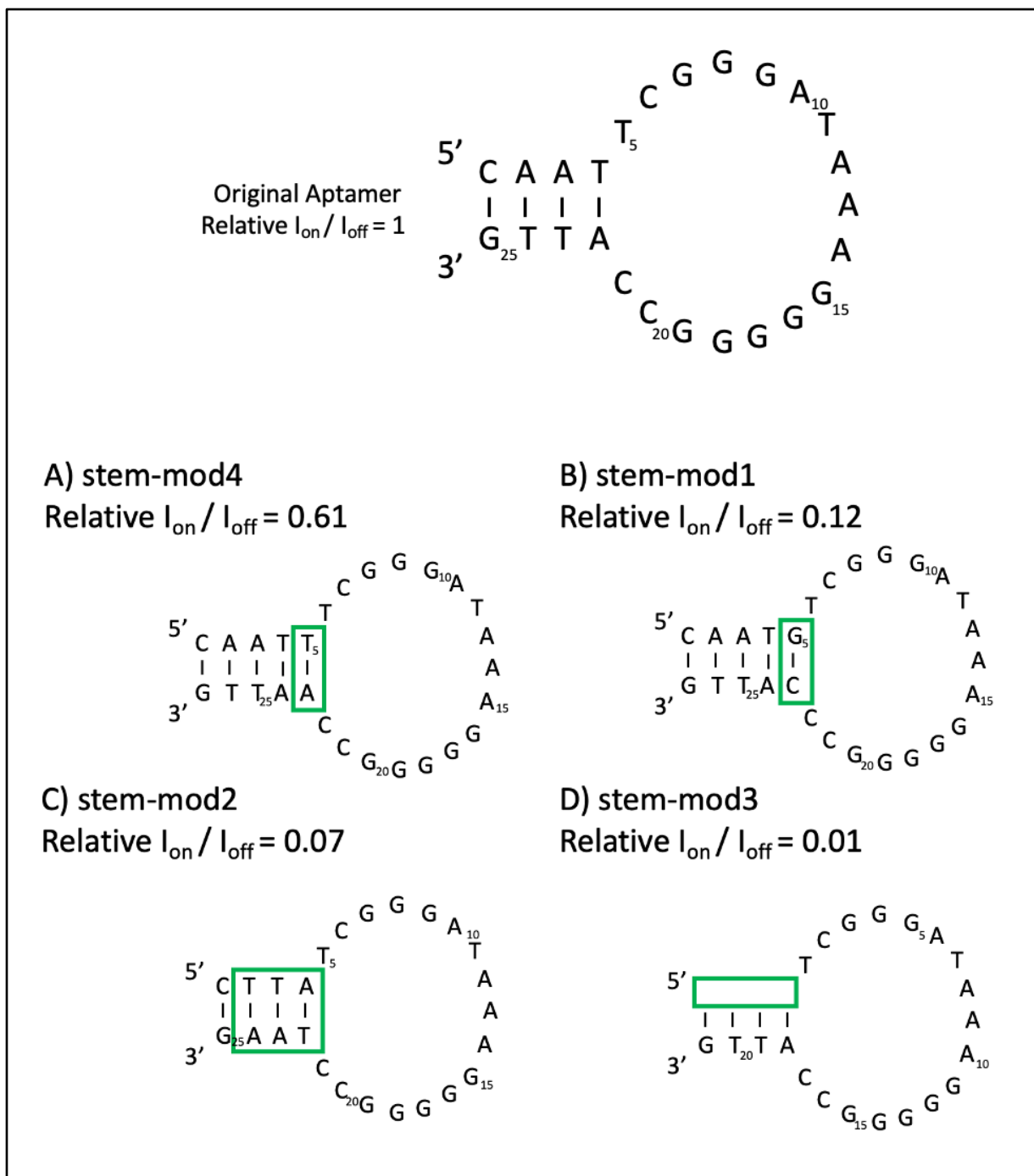
Figure 2.8 Stem sequence modifications tested by Sando *et al.*[20] Figure based on figure in Sando *et al.*[20] Green boxes indicate where the alterations to the sequence were made. A) stem-mod4 shows the addition of a AT base pair at the top of the stem, where the stem connects to the loop. B) stem-mod1 shows the addition of a GC base pair at the top of the stem, where the stem connects to the loop. C) stem-mod2 shows the alterations made to the bases in the stem. D) stem-mod3 shows the removal of one side of the stem. Relative I_{on}/I_{off} values were taken from Sando *et al.* where I_{on}/I_{off} represents the enhancement in fluorescence and the values were set as relative values to compared them to the I_{on}/I_{off} of the original aptamer.

In addition to the four base pair stem, the predicted secondary structure for the DNA aptamer included a 17 base pair loop.[20] In 2007 Sando *et al.* performed sequence modifications to the 17 bases involved in the loop, but once again, the presented data could also support the hypothesis of a G-quadruplex. The first sequence modification reduced the number of bases in the loop and changed all the bases in the loop to thymine (Figure 2.9A).[20] The second modification kept all 17 bases in the loop but changed them all to thymines (Figure 2.9B).[20] Although the guanines in the stems were not altered both loops alterations eliminated most of the guanines that are predicted to make up the G-quadruplex.
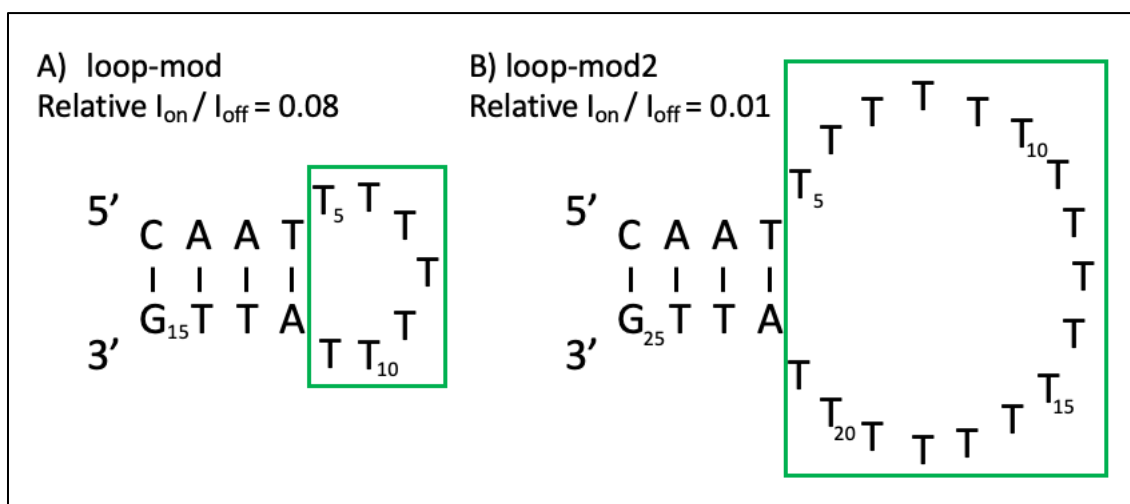
Figure 2.9 Loop sequence modifications tested by Sando *et al.*[20] Green boxes indicate the changes made to the predicted loop of the DNA aptamer. A) loop-mod shows the smaller loop with all the bases substituted for thymine. B) loop-mod2 shows all the bases in the loop substituted for thymine. Relative I_{on}/I_{off} values were taken from Sando *et al.* where I_{on}/I_{off} represents the enhancement in fluorescence and the values were set as relative values to compared them to the I_{on}/I_{off} of the original aptamer.

The observed decrease in relative fluorescence intensity observed by Sando *et al.* in 2007 following their sequence modifications could be interpreted as additional evidence for a stem-loop structure, but it could also provide evidence for the G-quadruplex theory presented here. To expand on the work performed by Sando *et al.* in 2007, another sequence alteration was developed for the research in this thesis. The alteration included four additional base pairs at the end of the predicted

stem (Figure 2.10). The hypothesis was that the formation of stable base pairing at the ends of the sequence would decrease the aptamer's ability to form a G-quadruplex, resulting in a decrease in fluorescence intensity. This hypothesis was explored in 2.3.3.3. The aptamer with a long stem was denoted as "extended stem aptamer", for the purpose of this work.
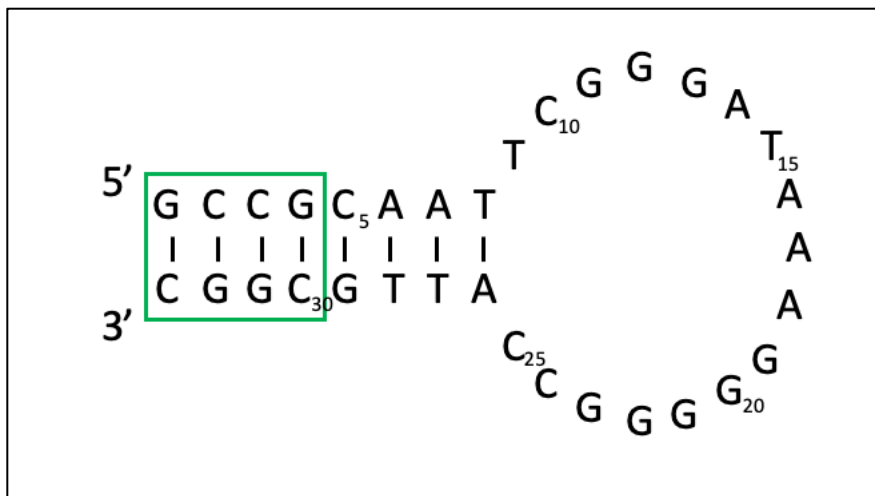
Figure 2.10 Extended Stem Aptamer. The green box indicates the extra bases added. Adding base pairs to the end of the stem was predicted to encourage the stem-loop structure therefore decreasing the ability of the aptamer to form a quadruplex.

2.3.3 Confirming Binding of Dyes to Samples by Fluorescence Intensity Scans

2.3.3.1 Fluorescence Scan of Original DNA Aptamer and Hoechst 33342

The first experiment that was performed with the original DNA aptamer and Hoechst 33342 was a fluorescence emission scan. Considering that the dye produces an increase in fluorescence intensity upon binding, the fluorescence scan can provide an indication as to whether the dye was binding to the aptamer.[20,56,59] The fluorescence experiment compared the fluorescence of the Hoechst 33342, without any nucleic acid, to the Hoechst 33342 with excess original DNA aptamer (Figure 2.11). The Hoechst 33342 dye with original DNA aptamer had approximately a 21-fold increase in fluorescence at its maximum peak at 465nm, compared to dye on its own,

indicating that the aptamer was binding to the Hoechst 33342 dye. The Hoechst 33342 dye was

originally developed to bind to DNA therefore it was not surprising that the dye would bind to the

original DNA aptamer to some degree.
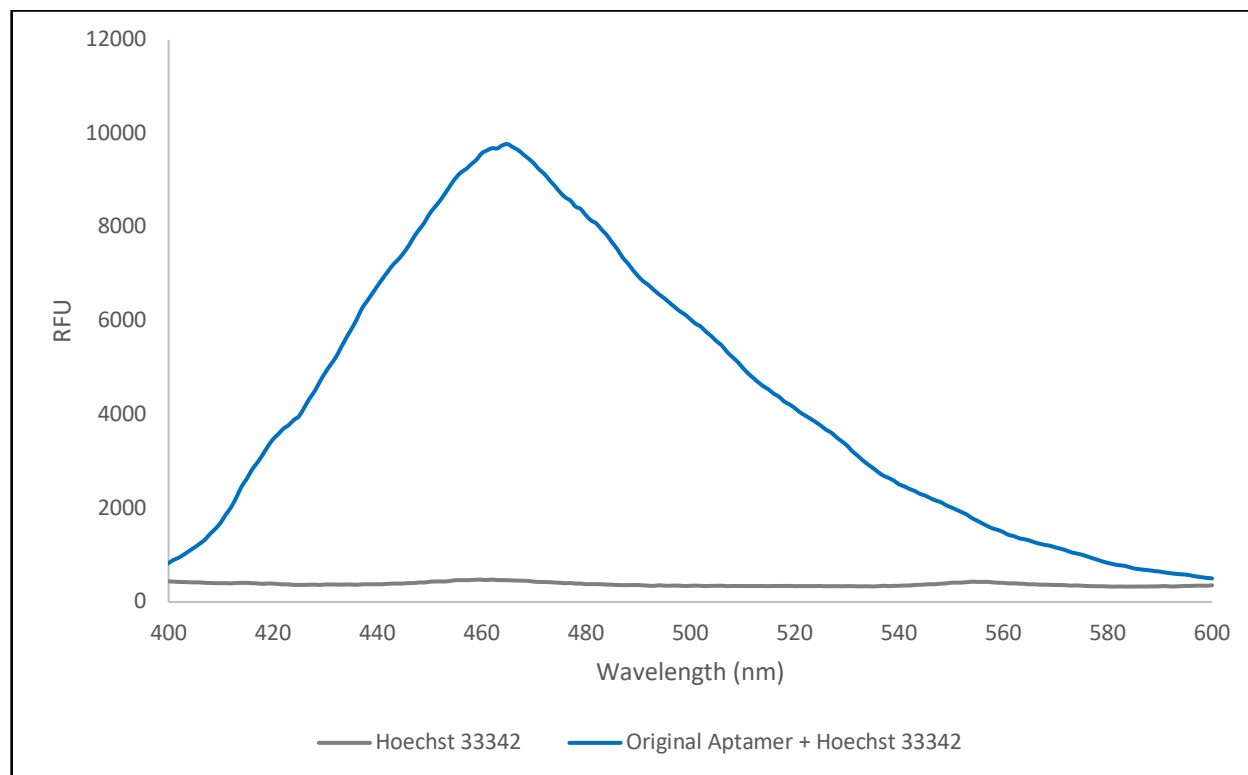
Figure 2.11 Fluorescence data from microplate reader for Hoechst 33342 and original DNA aptamer. Both samples contain a PBS buffer (pH=7.4) with Mg^{2+}. The control sample contains 10μM Hoechst 33342 (grey) and the Hoechst DNA sample (blue) contains 40μM of the original DNA aptamer and 10μM Hoechst 33342 dye.

2.3.3.2 Fluorescence Scan of Hoechst Aptamer Variations and t*Bu* Hoechst dye

The original DNA aptamer was selected by Sando *et al*. to bind to the t*Bu* Hoechst dye, but

prior to any other research it is important to confirm the formation of the aptamer-ligand complex

through fluorescence.[20,56,59] Once again, the binding was confirmed through a fluorescence scan.

Scans were performed on one sample containing only t*Bu* Hoechst dye, and another sample

containing t*Bu* Hoechst dye and original DNA aptamer. At 465mn, the sample with the original

DNA aptamer and t*Bu* Hoechst dye showed approximately a 16-fold increase in RFU, compared

to the samples with only dye (Figure 2.12). Again, this indicated that the original DNA aptamer was binding to the t*Bu* Hoechst dye, as expected.

Additionally, fluorescence scans were performed for the I_{16} DNA aptamer, the I_{25} DNA aptamer, the extended stem aptamer, and the I_{16} extended stem aptamer (Table 2.3). All samples containing aptamers and t*Bu* Hoechst dye showed an increase in fluorescence compared to the sample with only t*Bu* Hoechst dye. It was concluded that all variations of the Hoechst aptamer were capable of binding to the t*Bu* Hoechst dye in some capacity (Figure 2.12).

Table 2.3 Increase in RFU for Altered Hoechst Aptamer Sequences

Sample	x-fold increase in RFU
Original DNA Aptamer + t*Bu* Dye	~16
I_{16} DNA Aptamer + t*Bu* Dye	~13
I_{25} DNA Aptamer + t*Bu* Dye	~10
Extended Stem Aptamer + t*Bu* Dye	~21
I_{16} Extended Stem Aptamer + t*Bu* Dye	~12
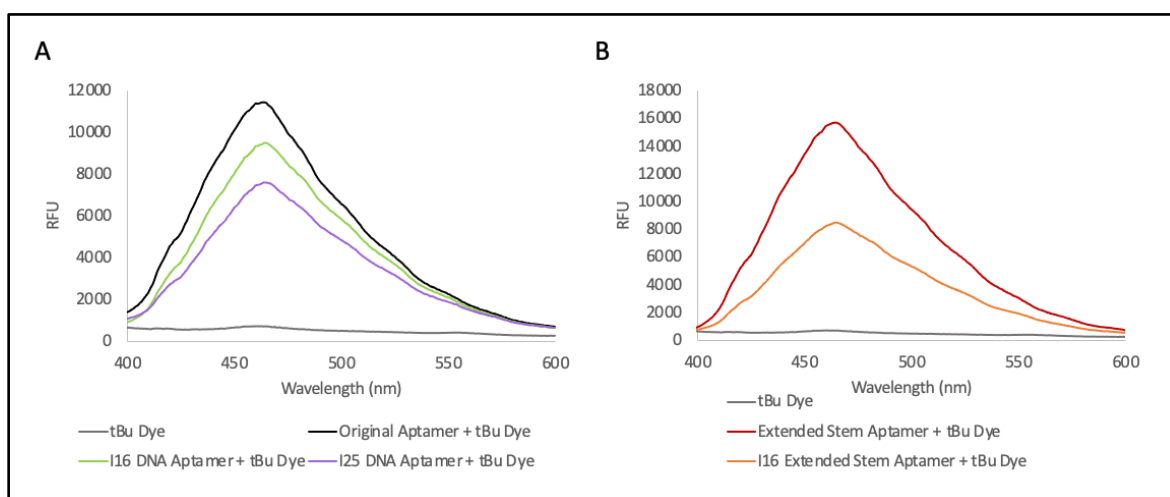
Figure 2.12 Fluorescence data from microplate reader for confirming binding of aptamers to t*Bu* Hoechst dye. A) A fluorescence scan for original DNA aptamer, I_{16} DNA aptamer, and I_{25}

46

DNA aptamer with t*Bu* Hoechst dye. B) A fluorescence scan for the extended stem aptamer and the I_{16} extended stem. All samples include 10µM dye and PBS buffer (pH=7.4) with Mg^{2+}. Samples were excited at a wavelength of 345nm. Where applicable samples contained 40µM DNA.

2.3.3.3 Comparison of K_d Values by Fluorescence Titration

Fluorescence titration experiments can be utilized to determine the K_d value for aptamer-ligand complexes. As discussed previously in Section 2.3.2, based on the predicted G-quadruplex structure it was theorized that extending the stem could cause a disruption in the structure, resulting in a decreased binding affinity. Experimentally it was shown that the extended stem did not reduce the binding affinity (Figure 2.13). In fact, the K_d for the extended stem aptamer was smaller than the K_d of the original DNA aptamer. The extended stem aptamer was created with the hope of stabilizing the stem and preventing the formation of a quadruplex. Based on the fluorescence scan and the fluorescence titrations there was not a decrease in binding for the extended stem aptamer, contradicting the hypothesis. However, this does not necessarily disprove the presence of a G-quadruplex as the extended stem could still form a G-quadruplex.

Aptamer	Calculated K_d (µM)	95% Confidence Interval for K_d (µM)
Original Hoechst Aptamer	6.04	3.16 to 6.99
Extended Stem Hoechst Aptamer	1.38	0.65 to 2.76
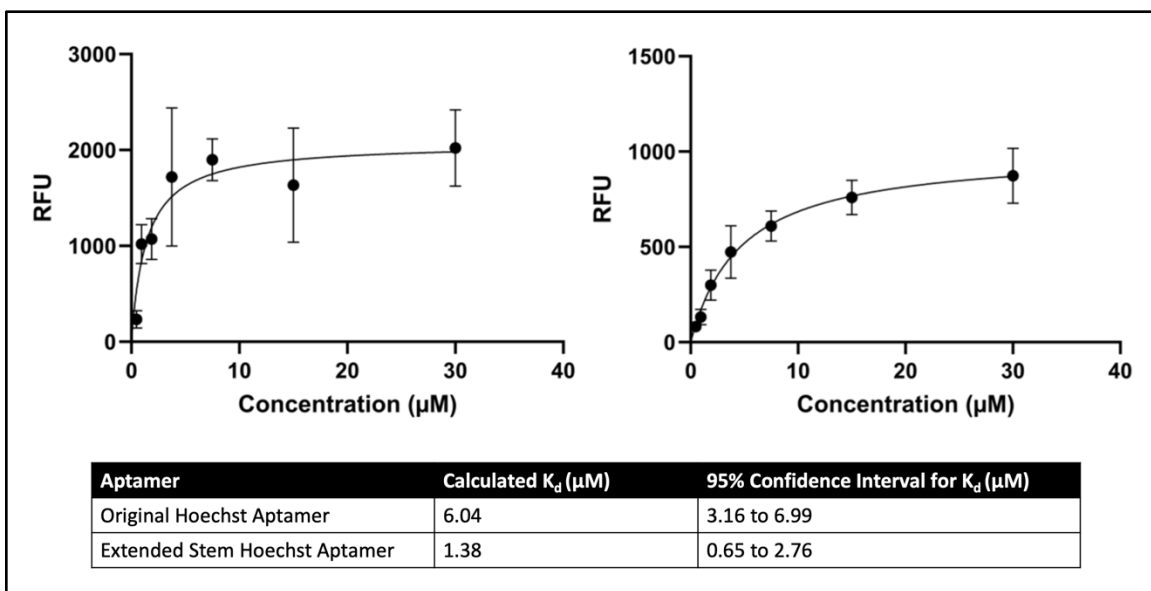
Figure 2.13 Fluorescence titration of Hoechst aptamers. Averages were taken from three triplicates and the resulting data was fitted using non-linear least squares regression. Samples contained varying concentrations of DNA with 1µM t*Bu* Hoechst dye and PBS buffer (pH=7.4)

with Mg^{2+}. Samples were excited at 345 nm and fluorescence was measured at 465nm. The equation $\% \boldsymbol{Fmax} = \boldsymbol{Fmax} \times \frac{[\boldsymbol{Nucleic\ Acid}]}{\boldsymbol{K_d} + [\boldsymbol{Nucleic\ Acid}]}$, where Fmax is the maximum fluorescence at a large concentration of the nucleic acid and K_d is the dissociation constant, was used to fit the curve. Error bars were calculated utilizing three triplicate samples and a standard deviation equation. Graphing, K_d calculation, and confidence interval calculation were all performed using GraphPad Prism.[129]

2.3.4 Native PAGE Results

To investigate the general size of the structures, a native PAGE was run (Figure 2.15). The size, shape, and charge of a molecule affect the migration of nucleic acids through the gel, therefore one cannot directly interpret the distance migrated as the size of the structure.[112–114] First, the gel was observed without stain using the FluorChem FC2 Imaging System and a UV excitation (Figure 2.14). It would be expected that any Hoechst dye present would emit a fluorescence signal, therefore viewing the gel prior to staining confirmed the presence of a bound aptamer-ligand complex. Wells without dye (Well 1, 2, 4) show no bands as would be expected. The original DNA aptamer with t*Bu* Hoechst dye (Well 3) shows a single band high on the gel that is consistent with a high molecular weight molecule. The extended stem aptamer with t*Bu* Hoechst dye (Well 5) shows a single, very bright band lower on the gel that is consistent with a monomeric structure. Since both Well 3 and Well 5 show a fluorescent t*Bu* Hoechst dye band in the gel, it was concluded that at least some of the dye stayed bound as the DNA migrated into the gel. The high band in Well 3 indicated that the original DNA aptamer was likely a multistranded structure because it moved slower in the gel. A sample may form a multistranded structure for various reasons but a high percent of guanines in a sequence or a high concentration of nucleic acid could result in a multistranded structure.[94,130] The lower band in Well 5 indicated that the extended stem aptamer was likely a single stranded structure because it moved faster through the gel. Well 6 does not show a fluorescent band within the gel but it does appear to have some dye present on the surface

of the well, that may be a result of the Hoechst 33342 dye separating from the original DNA

aptamer. The wells containing only t*Bu* Hoechst dye (Well 7) and only Hoechst 33342 dye (Well

8), did not show any migration of the dyes into the gel.
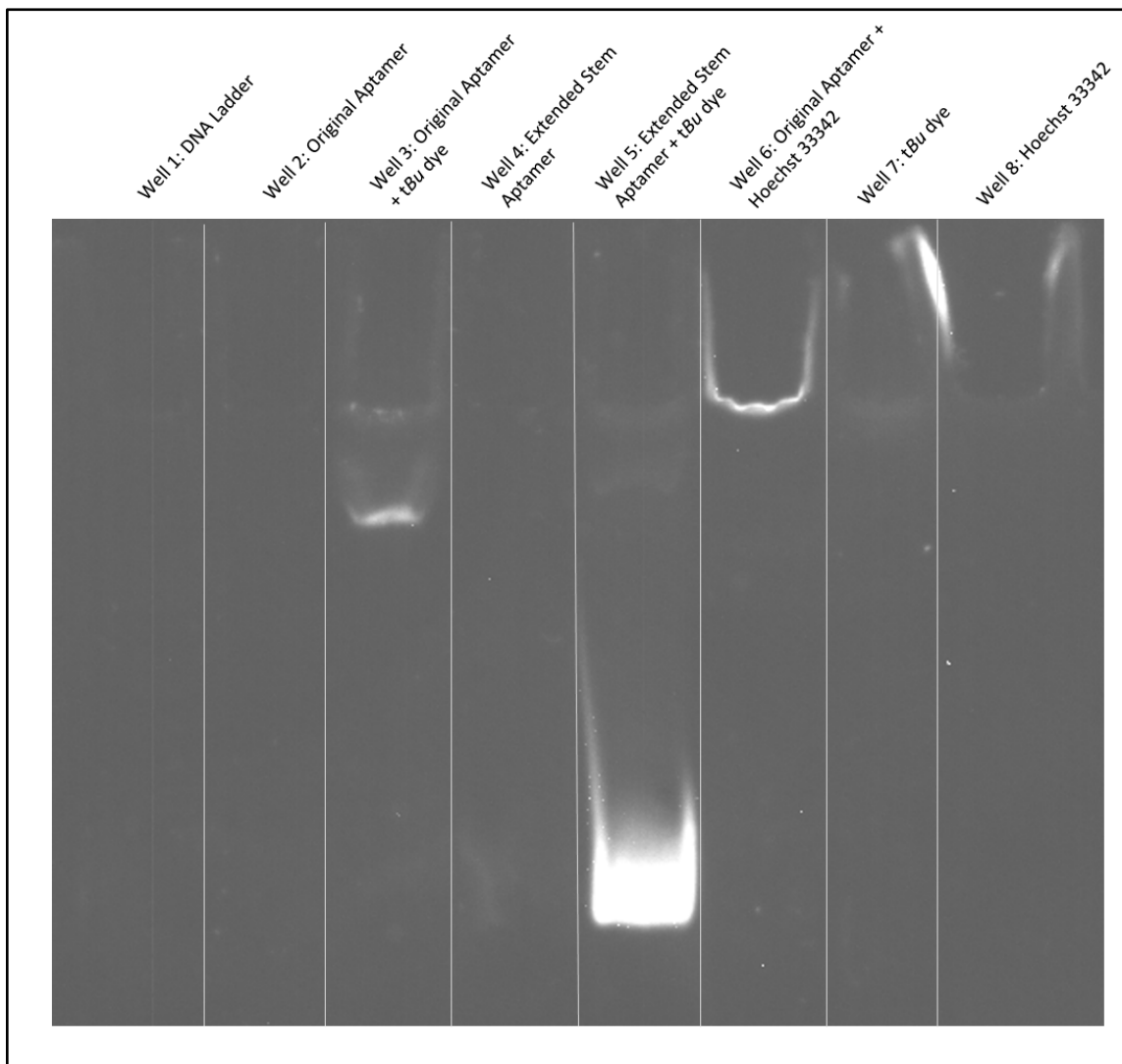
Figure 2.14 Visualization of 15% PAGE without stain by UV excitation. Well 1 (DNA Ladder) contains the Low Weight Molecular Weight DNA Ladder (New England Biolabs, Ipswich, MA) ranging from 766bp to 25bp. Each well contains approximately 0.7µg of the nucleic acid and if applicable approximately a 1:1 concentration ratio of dye. Well 2 contains only the original DNA aptamer. Well 3 contains the original DNA aptamer and the t*Bu* Hoechst dye. Well 4 contains only the extended stem aptamer. Well 5 contains the extended stem aptamer and t*Bu* Hoechst dye. Well 6 original DNA aptamer and the Hoechst 33342. Well 7 contains the t*Bu* Hoechst dye and Well 8 contains the Hoechst 33342 dye.

Next a gel was stained with SYBR Safe, and bands of DNA were clearly visible (Figure 2.15). Well 2 showed two bands present for the original DNA aptamer without dye. In the next well over (Well 3), the original DNA aptamer with t*Bu* Hoechst dye, showed two bands present. One brighter band appeared to migrate around the same distance as the 350bp ladder, and one less bright band migrated past the 25bp ladder. The higher band corresponded to the t*Bu* Hoechst dye band seen in the gel that was visualized prior to staining (Figure 2.14). The presence of two bands appeared to indicate that under these conditions the original DNA aptamer has two different conformations, with the major form consisting of a multistranded structure that is bound to the t*Bu* Hoechst dye.[110] The extended stem without dye showed a smeared banded (Well 4), but once the dye was added the band becomes much more resolved (Well 5). The extended stem aptamer with t*Bu* Hoechst dye (Well 5) produced a single bright band past the 25bp ladder, corresponding to the band seen in the gel prior to staining (Figure 2.14).
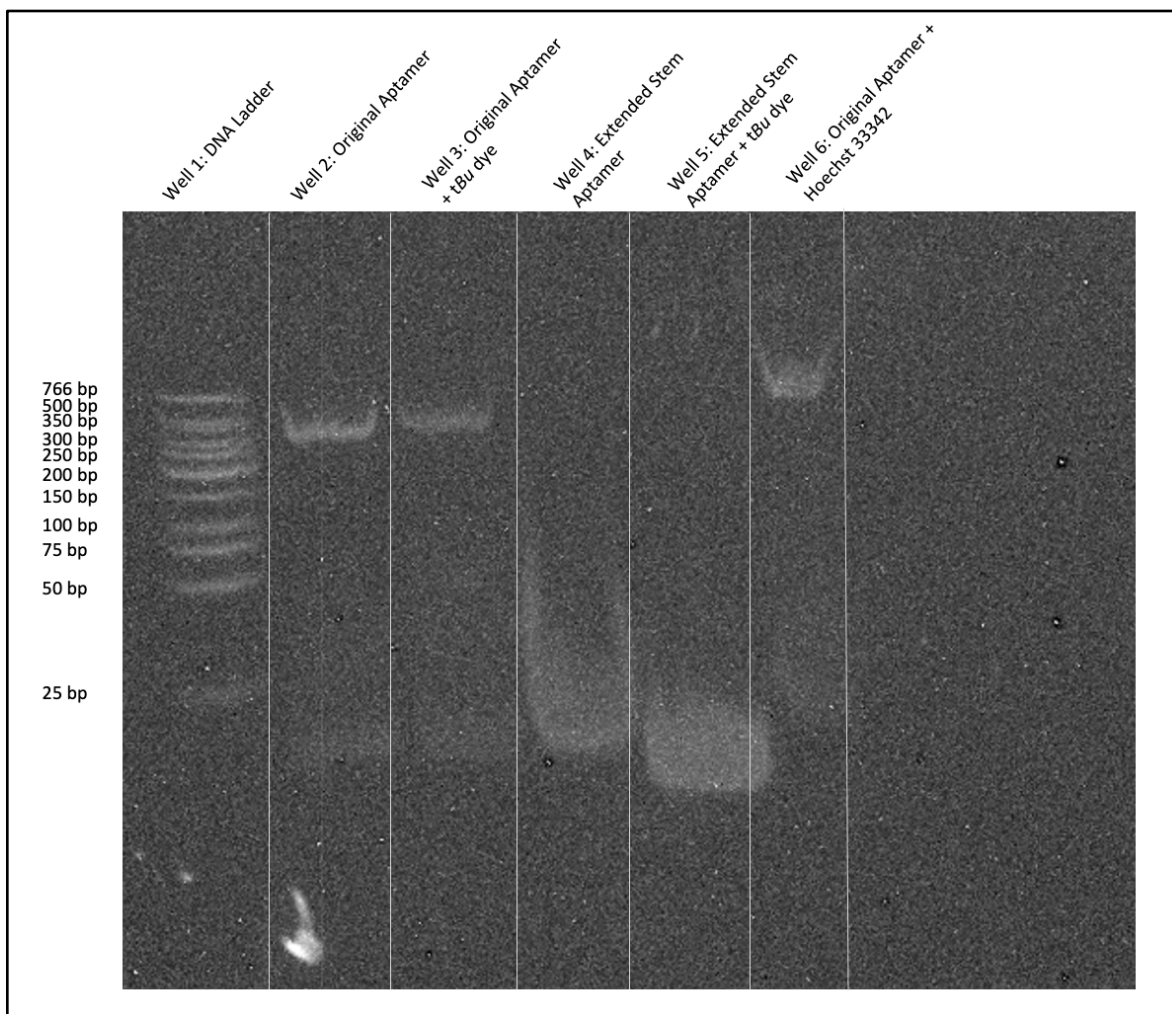
Figure 2.15 Photo of 15% PAGE stained with SYBR Safe. Well 1 (Low MW Ladder) contains the Low Weight Molecular Weight DNA Ladder (New England Biolabs, Ipswich, MA) ranging from 766bp to 25bp. Each well contains approximately 0.7μg of the nucleic acid and if applicable approximately a 1:1 concentration ratio of dye. Well 2 contains only the original DNA aptamer. Well 3 contains the original DNA aptamer and the t*Bu* Hoechst dye. Well 4 contains only the extended stem aptamer. Well 5 contains the extended stem aptamer and t*Bu* Hoechst dye. Well 6 original DNA aptamer and the Hoechst 33342.

2.3.5 Circular Dichroism Studies

CD spectra produce maxima and minima at key wavelengths that can be indicative of secondary structures within a nucleic acid.[105,115,116] Table 2.4 contains literature values for known CD peaks as well as the peaks observed in the CD spectra obtained for the original Hoechst aptamer and the extended stem aptamer. The CD spectrum for the original DNA aptamer without

t*Bu* Hoechst dye and with t*Bu* Hoechst dye were relatively similar (Figure 2.17). Both had maximum peaks around 263 nm and 205 nm and a minimum peak near 240 nm. The original DNA aptamer with dye had an additional minimum at 280 nm. Based on the literature values listed in Table 2.4, the 263 nm and 240 nm peaks in the CD spectrum of the original DNA aptamer are consistent with a parallel G-quadruplex. The minima at 280 nm and maxima at 205 were both enhanced when the t*Bu* dye was added therefore the dye was likely enhancing a feature that is not generally characteristic of a quadruplex. The peak at 280 nm is also similar to the absorbance scan of the t*Bu* Hoechst dye (Appendix A). The CD spectrum of the extended stem aptamer has similar maximum and minimum peaks suggesting that extending the stem does not prevent the formation of a G-quadruplex under these conditions (Figure 2.17). Although the bases were added with the hope of stabilizing the stem, the presence of the G-quadruplex in CD, proves these bases did not achieve the prevention of a quadruplex. Additionally, the extended stem aptamer with t*Bu* Hoechst dye also contained a peak near 280 nm that was enhanced with the addition of the t*Bu* Hoechst dye.

Table 2.4 Expected CD Peaks Corresponding to Nucleic Acid Structures

Secondary Structure	Maximum	Minimum
B-form DNA[116]	275 nm	240 nm
Anti-parallel Quadruplex[115,117,118]	290 nm + 210 nm	260 nm
Hybrid[118]	295 nm + 290 nm	245 nm
Parallel[115,117,118]	264 nm	245 nm
Original DNA Aptamer	266 nm + 220 nm + 200 nm	247 nm
Original DNA Aptamer + t*Bu* Hoechst dye	256 nm + 205 nm	241 nm + 286 nm
Extended Stem Aptamer	260 nm + 221 nm + 200 nm	242 nm
Extended Stem Aptamer + t*Bu* Hoechst Dye	256 nm + 201 nm	241 nm + 278 nm
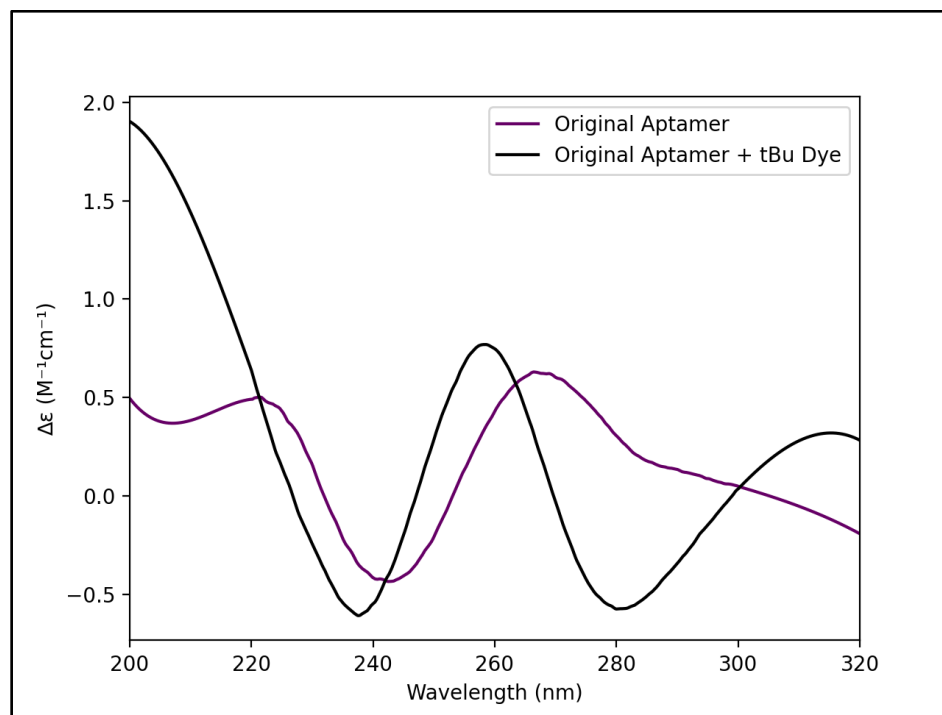
Figure 2.16 CD scan of original DNA aptamer. Samples contain 40µM original DNA aptamer, 50µM t*Bu* Hoechst dye, and PBS buffer (pH 7.4) with Mg^{2+}. Samples measured with 0.1cm pathlength quartz cuvette from 320nm to 200nm, with a data interval of 0.5nm, response time of 1 second, band width of 0.5nm, and a scanning speed of 30nm minute^{-1}. Four scans were

accumulated. Baselines were established using either only buffer or buffer with 50μM t*Bu* Hoechst dye. The curve was smoothed in python using Scipy's Savitzky-Golay filter.[127]
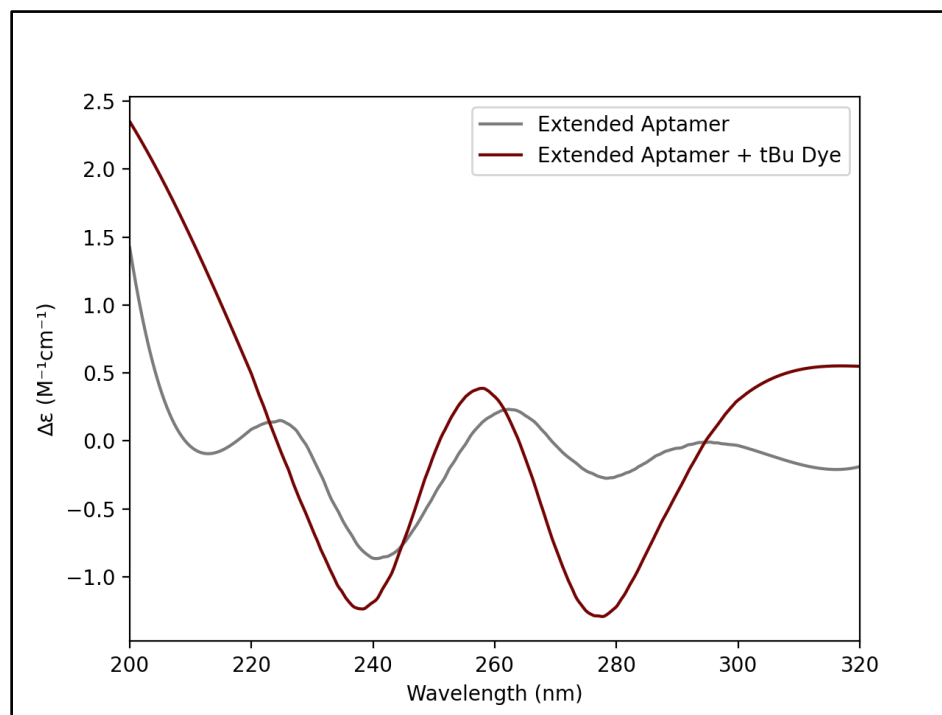
Figure 2.17 CD of extended stem aptamer. Samples contain 40μM extended stem aptamer, 50μM t*Bu* Hoechst dye, and PBS buffer (pH 7.4) with Mg^{2+}. Samples measured with 0.1cm pathlength quartz cuvette from 320nm to 200nm, with a data interval of 0.5nm, response time of 1 second, band width of 0.5nm, and a scanning speed of 30nm minute[-1]. Four scans were accumulated. Baselines were established using either only buffer or buffer with 50μM t*Bu* Hoechst dye. The curve was smoothed in python using Scipy's Savitzky-Golay filter.[127]

2.4 Conclusions

The biomolecular methods presented in this chapter focused on exploring the presence of a G-quadruplex and on modeling the system to determine suitable NMR studies. The presence of a quadruplex was hypothesized due to the large number of guanines in the original DNA aptamer sequence. Computational modeling paved the way for the design of inosine substitutions and the development of an extended stem aptamer. At the beginning of this research, it was assumed that if a quadruplex was forming it would be a single stranded structure, therefore the models were built for a single stranded quadruplex. Fluorescence experiments demonstrated all sequences

appeared to be binding to the t*Bu* Hoechst dye. The native gel indicated that the original DNA aptamer was mostly forming a multi-stranded structure under the native gel conditions, but the extended stem aptamer appeared to form a single stranded structure. Armed with this knowledge CD was performed on the original DNA aptamer and the extended stem aptamer. Within CD the original DNA aptamer followed trends that resemble a parallel G-quadruplex. The buffer used throughout this chapter was the same as the buffer utilized by Sando *et al.* to ensure similar conditions to the selection of the aptamer. Based on the biomolecular information presented here it appears that the original Hoechst DNA aptamer is forming a multi-stranded G-quadruplex structure under certain conditions, instead of the hairpin structure that was initially suggested.

Chapter 3   Characterization   of   Hoechst   DNA Aptamer by NMR

Fluorophore binding aptamers provide the opportunity to develop novel biosensors that are relatively affordable, sensitive, and highly selective. As previously discussed, there is large variety in biosensor design and because aptamers can fold into a large variety of structures, it is safe to say that there is a multitude of possibilities for aptasensor design. In Chapter 1, several solved structures for aptamers bound to their ligands and biosensors developed using such aptamers were introduced. Solving the structure of an aptamer-ligand complex provides invaluable information for any application of the aptamer-ligand complex.

High-resolution NMR spectroscopy is an excellent tool for aiding in structural studies of biomolecules, including nucleic acids, in solution.[131] In NMR-spectroscopy a sample is placed into a magnetic field and the spectrometer uses radiofrequency pulses to simultaneously excite all nuclei.[131] The raw output is difficult to interpret as the signals are a function of time.[131] To overcome this difficulty a mathematical method, known as a Fourier transformation, is applied to convert the data to a function of frequency.[131,132] Once a Fourier transformation is applied, the data in a 1D spectrum is presented as frequency vs peak intensity, allowing for analysis to be performed on the spectrum.[131,132]

The usefulness of Fourier transformation is not limited to one-dimensional spectra; it can also be applied to multidimensional NMR. After the Fourier transformation is applied to a 2D spectrum the data is presented with two frequency axis and the intensity displayed as a contour plot.[131,132] There are a variety of different NMR experiments that can provide information about different parts of a nucleic acid structure, such as exchangeable protons vs non-exchangeable protons.[120,121,133,134] Acquiring spectra from multiple types of experiments can produce spectra that

can then be overlaid to combine information. However, there are limitations to the information that can be retrieved from overlaid spectra. Nucleic acids are built from only four bases with very similar structures.[7] Unfortunately, this results in many of the peaks within a nucleic acid NMR spectrum to be overlapped in a small region (Figure 3.1).[135,136] For longer oligonucleotides, overlapping signals can make unambiguous assignment of the spectra difficult. As the nucleic acid grows in size the problem of overlapping peaks is also increased.[135,136] In cases where the peaks cannot be assigned other methods may be utilized, such as base specific labelling, wherein all bases of one type can be fully or partially labelled, or isotope labelling of segments.[136,137]
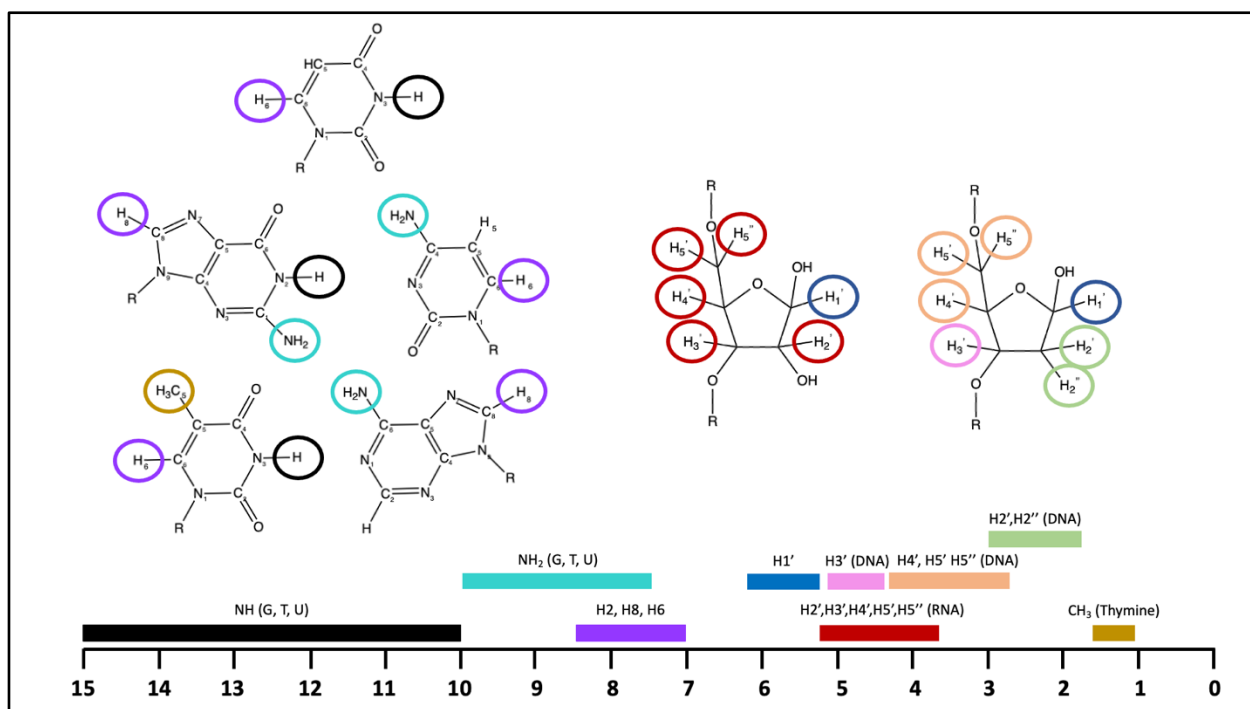
Figure 3.1 Approximate chemical shifts for nucleic acid protons in a NMR spectrum.[138] As nucleic acids contain only four bases with similar structures there can be overlap within the NMR spectrum that can make assignment difficult.

1D proton NMR experiments can be helpful for initial feasibility studies. 1D NMR experiments can be run quickly at lower concentrations, compared to 2D experiments.[121] In particular a 1D titration experiment, wherein aliquots of dye are added until a 1:1 ratio of aptamer

to ligand is achieved, can be utilized to observe if binding is occurring.[120,121] In particular the 11-14ppm region of the 1D spectrum contains peaks resulting from imino protons of base paired nucleotides.[121] Therefore based on the sequence and predicted secondary structure of the nucleic acid, one can estimate the expected number of peaks present.[121] Due to their quick and affordable nature, 1D experiments can be a very useful first step for providing initial information regarding the structure of a nucleic acid when bound to its ligand.

While 1D experiments can provide some insight into the binding of aptamers to ligands, 2D experiments are important for obtaining more in-depth structural information. A commonly used 2D NMR experiment that can provide useful structural information is a 2D nuclear Overhauser enhancement spectroscopy (NOESY) experiment. NOESY experiments provide information regarding longer range interactions within a molecule by producing NOEs (nuclear Overhauser enhancement) signals from through-space correlations for protons that are within ~5Å of each other.[120,121,133,134] 2D NOESY performed in 90% H_2O/10% D_2O are useful for assigning the imino protons.[120,121] 90%H_2O/10%D_2O NOESYs can be utilized to identify bases that are stacked on top of each other in a helix (Figure 3.2).[120,121] NMR experiments can also be performed in 100% D_2O. 100% D_2O NOESYs can allow for sequential assignments of the H8/H6 and H1' protons (Figure 3.3).[120,121] Another 100% D_2O experiment that is frequently used for nucleic acid NMR is a total correlation spectroscopy (TOCSY) that produces cross-peaks from through bond correlations allowing for the identification of the H5/H6 of cytosines and uracils (Figure 3.4).[120,121,139] By combining information from these experiments one can gain additional insights into the structure of a nucleic acid.
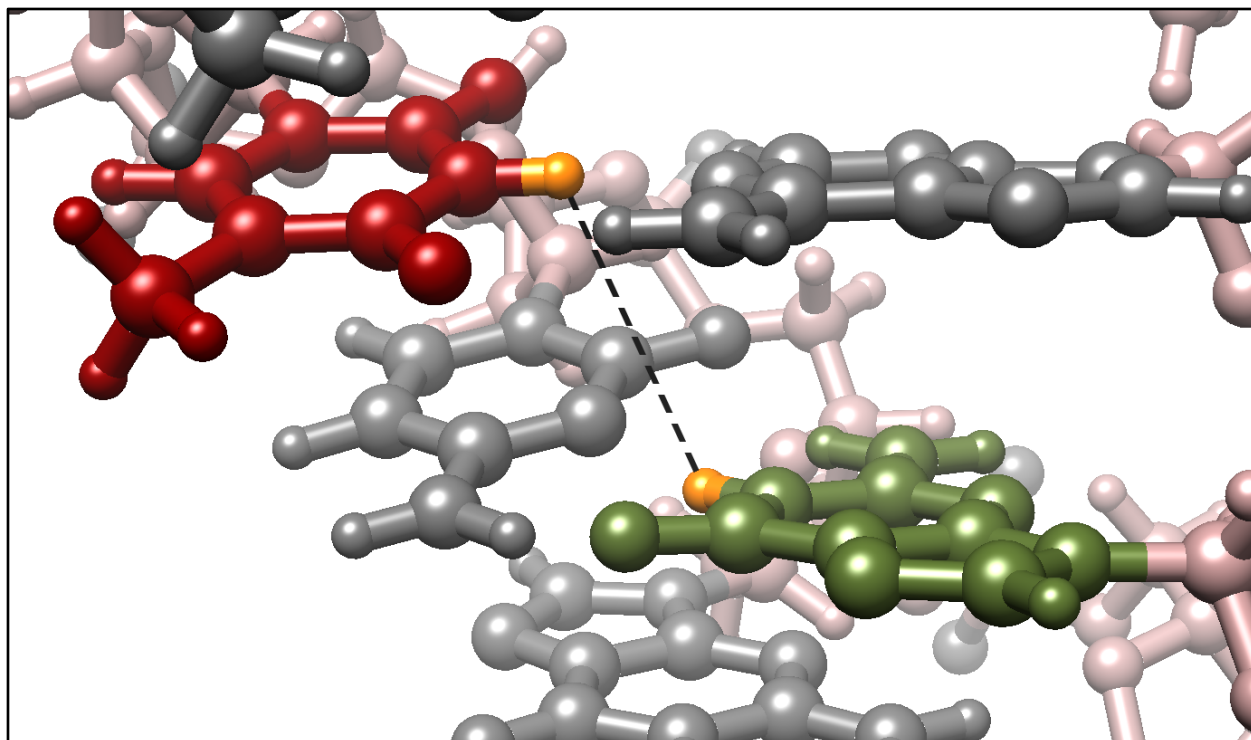
Figure 3.2 Representation of the expected imino correlation proton in a 90%H₂O/10%D₂O spectrum. The black dotted line represents the NOEs that are expected between base pairs that are stacked (PDB:1BNA), the distance from the thymine H3 to the guanine H1 is ~3.9Å.[9,121] Structures were rendered using Chimera.[11]
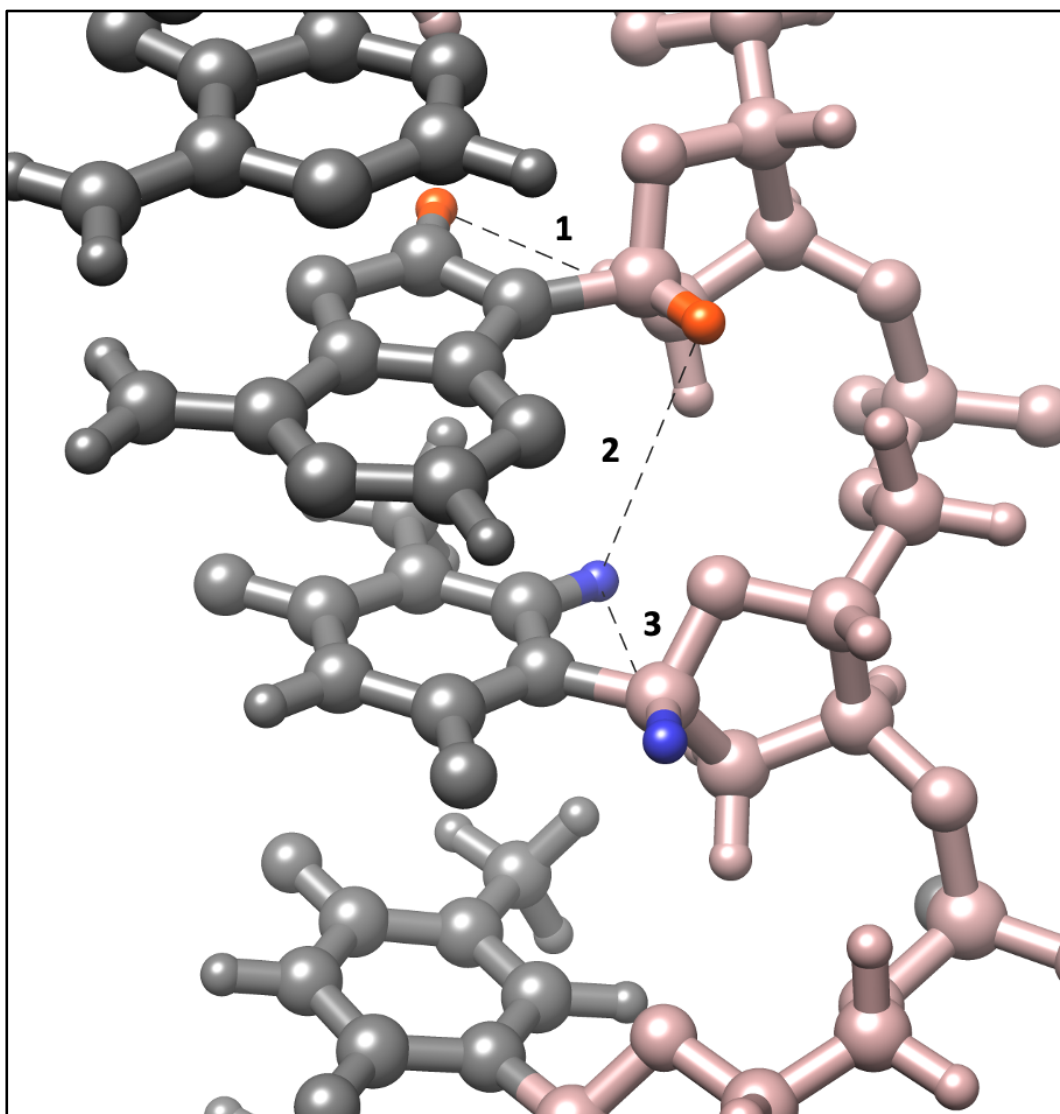
Figure 3.3 Representation of the expected resonances allowing for a sequential walk in a 100% D₂O Spectrum. Each dotted line represents a different NOE that can be used for sequential walking along the backbone (PDB:1BNA).[9] 1) Here, the walk begins at the H8-H1' of a adenine with a distance of ~3.9Å. 2) Next, the H1' of the adenine to the H6 of the thymine with a distance of ~4.0Å. 3) Finally, the H6 of the thymine to the H1' of the thymine with a distance of 3.7Å.[120] This process can be repeated throughout the backbone of the nucleic acid. Structures were rendered using Chimera.[11]
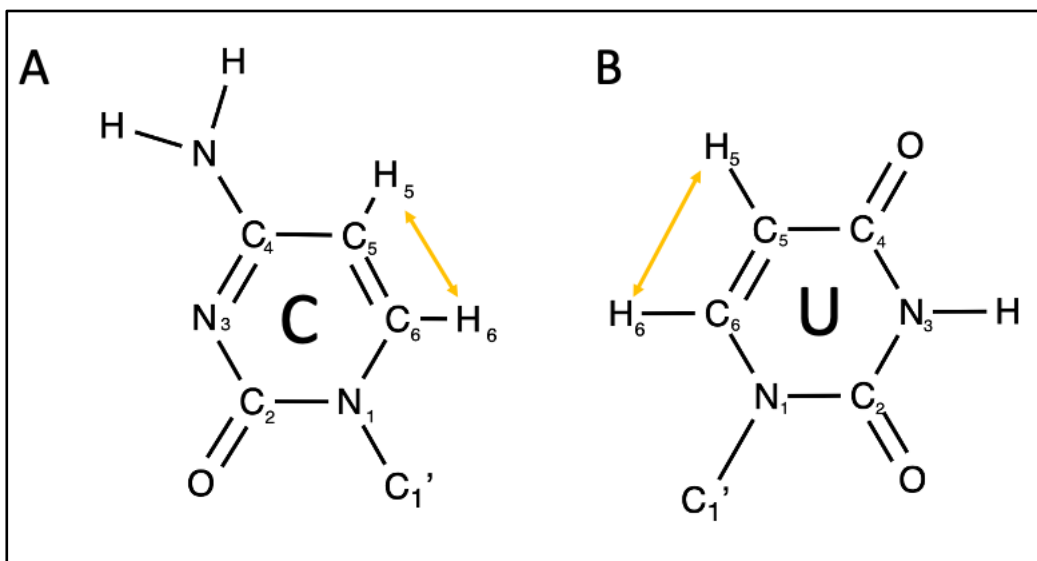
Figure 3.4 Expected NOEs within cytosine and uracil in TOSCY Spectrum. The yellow arrows represent the intramolecular NOEs seen in nucleic acid base pairs.[120,121,139] A) The intramolecular NOEs in a cytosine. B) The intramolecular NOEs in uracil.

As discussed in Chapter 2, the large number of guanines present in the original DNA aptamer sequence raised the possibility of a G-quadruplex structure. There are certain methodologies that can be applied for studying quadruplex structures. As discussed, NOESY experiments can be used to correlate protons that are within ~5Å of each other.[119–121,133,134] Within a quadruplex there are certain NOEs that can be expected. Similarly to bases that are stacked within a backbone, we would expect to see NOEs for H_8-H_1' between stacked tetrads.[85] The distance between the H_8-H_1' differs between *syn* and *anti* conformations, with the *syn* conformation resulting in a shorter distance.[71,82,140] The distances between protons are directly related to the intensity of the NOEs therefore intensity can indicate the glycosidic conformations, with strong intensity peaks indicating *syn* and more moderate intensity peaks indicating *anti*.[71,82,85,140] Additionally within the tetrad it is expected that there will be NOEs between the imino and H8 of the guanines (Figure 3.5A).[85,141] For tetrads containing a thymine the same NOEs would exist for the guanines but the thymine would have additional peaks (Figure 3.5).[82]

61
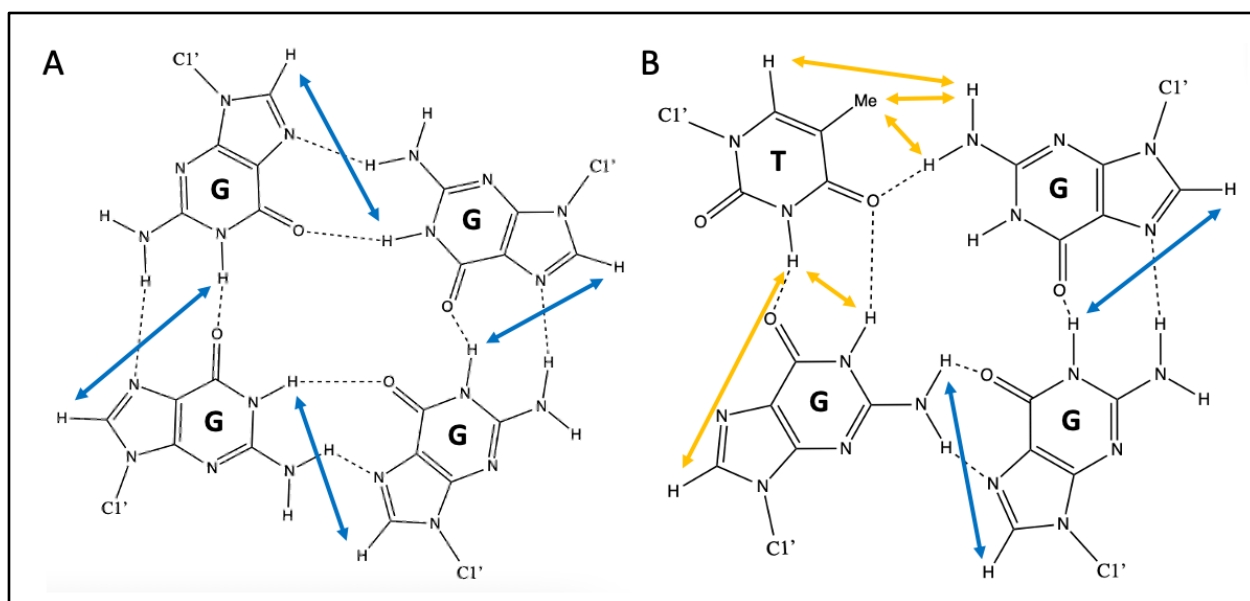
Figure 3.5 Intermolecular NOEs of Exchangeable Protons in a G-quadruplex. A) GGGG tetrad NOEs.[141] B) GGGT tetrad NOEs.[82] Yellow arrows represent the NOEs between thymine and a guanine. Blue arrows represent NOEs between two guanines, based on PDB:3T5E and PDB:1XAV structures these protons are slightly larger than ~5Å apart.[142,143]

When studying quadruplex structures it is common to substitute a guanine within the quadruplex with a inosine to aid in unambiguous assignment of NMR spectra.[34,73,74,85,111,141,144–146] Inosine substitution can be used to provide additional structural information regarding quadruplex structures. Inosine is a modified nucleic acid with a structure similar to guanine, but the amino group at N2 is replaced with an imino (Figure 3.6).[73,97,147] However, it is important to note that despite its similar structure to guanine, some past research has demonstrated that the substitution of inosine at particular bases can completely disrupt the quadruplex structure.[108,109] It has been theorized that the amino group of certain guanines may be critical for ligand recognition and for nucleic acid folding, therefore if a critical guanine is substituted for an inosine the structure will not properly fold.[108,109] Although it is possible for the inosine to disrupt the structure, there are many examples of successful inosine substitutions.[34,73,74,85,111,141,144–146] If the selected guanine is part of a quadruplex, and the quadruplex is able to fold, there will be a characteristic peak in the

13-14ppm region a 1D spectrum resulting from the inosines imino proton.[85,111,141,145,146] The characteristic peak produced by the substitution of inosine can be combined with other NMR data to aid in assignment.
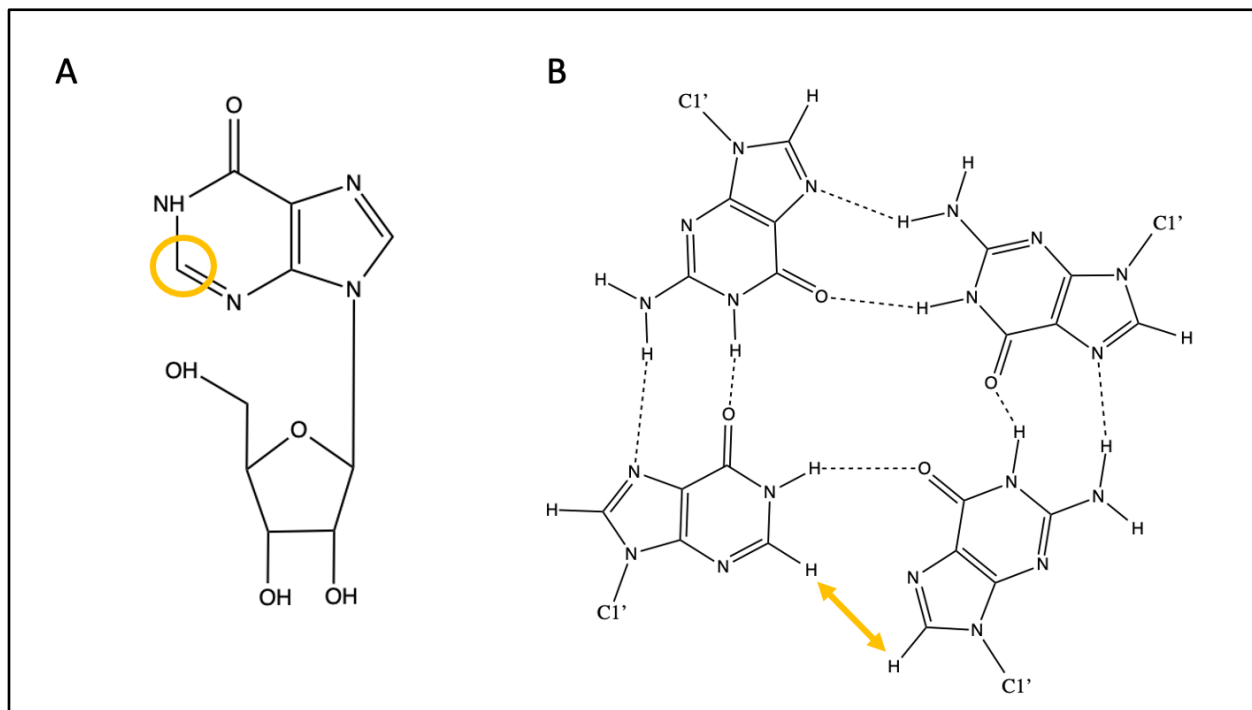
Figure 3.6 Inosine Substitution into Tetrad. A) The structure of inosine with the substituted imino group circled in yellow. B) Yellow arrow represents the NOE between the H2 of inosine and the H8 of guanine in a G-tetrad.[73,97]

The work presented in this chapter aimed to explore the binding of the Hoechst DNA aptamer through NMR.[20,21] The Hoechst 33342 dye and the t*Bu* Hoechst dye were utilized for NMR studies. In addition, different variations of the Hoechst DNA aptamer were explored. These variations include an extended stem aptamer and inosine substituted samples. Once the initial studies were completed further studies were performed utilizing a t*Bu* Hoechst dye synthesized by Avery To in Dr. Graham Murphy's lab at the University of Waterloo.

3.1 Methods and Materials

3.1.1 Molecular Graphics

Performed as described in 2.2.1.

3.1.2 Purification of Nucleic Acids

Performed as described in Section 2.2.3.

3.1.3 Hoechst Dyes

Performed as described in Section 2.2.4.

3.1.4 NMR Spectroscopy Experiments

Samples were prepared for NMR by dissolving the lyophilized nucleic acid into a 10 mM potassium phosphate buffer (pH=6.9) with 10 mM KCl. All NMR experiments were run using 5mm NMR tubes with a total sample volume of 500µL. A Bruker DRX-600 spectrometer equipped with an HCN triple-resonance, triple-axis PFG probe was used to obtain all spectra. Non-exchangeable protons were observed in 99.9% D_2O (Cambridge Isotopes, Tewksbury, MA) using presaturation solvent suppression.[148] D_2O experiments include 2D CITY-TOCSY with a mixing time of 50 ms and 2D NOESY with a mixing time of 150 ms.[149] For the t*Bu* Hoechst dye a correlation spectroscopy (COSY) was acquired in D_2O.[150] Additionally rotating frame Overhauser effect spectroscopy (ROESY) experiments were performed in D_2O with a mixing time of 80ms to observe any exchange peaks present.[151] All D_2O spectra were acquired at 298K. Exchangeable protons were observed in 90% H_2O/10%D_2O using 1-1 spin echo solvent suppression.[152] Most of the 90% H_2O/10%D_2O experiments, including 1D proton NMR and 2D NOESY, were acquired at 277K.[153] Some samples precipitated out at 277K, therefore they were run at 283K. Additionally, in samples that precipitated the salt concentration was increased to 80mM KCl. 1D proton NMR

experiments were utilized to observe a titration of the aptamer with dye in a step wise fashion. Dye was added in aliquots until a ratio of 1:1 was achieved. For 2D NMR experiments lyophilized aliquots of dye were added to the aptamer samples to yield a 1:1 ratio with the nucleic acid, unless otherwise stated. All relevant pulse programs are included in Appendix B.

3.2 Results and Discussion

3.2.1 Gaining Context through Proton Distances Calculated in Chimera

As stated above, NOESY experiments allow for the visualization of protons that are within ~5Å of each other, therefore predicting the distances between protons can aid in understanding what NOEs will be present. Chimera contains a tool to provide the distance between atoms in a structure.[11] In Chapter 1 the structure of a dodecamer duplex bound to Hoechst DNA, solved by Sriram *et al*., was discussed.[62] By importing that structure into Chimera, it is possible to observe distances between protons (Figure 3.7). Distances within ~5Å would be expected to produce a peak in a NOESY experiment, with shorter distances producing peaks with greater intensities.[131,138] Observing the distance between protons in the solved duplex can provide some context regarding the binding of Hoechst 33342 to DNA.
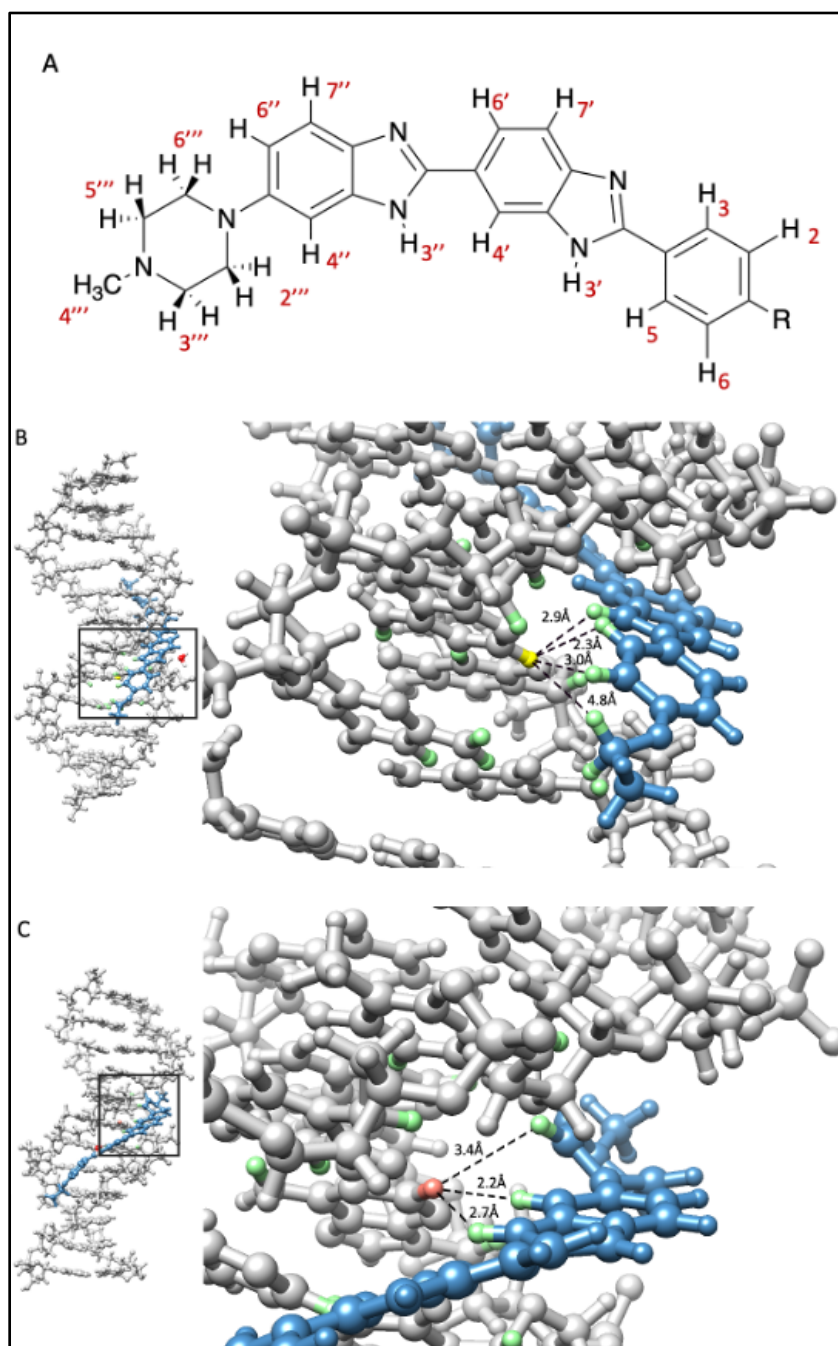
Figure 3.7 Distance of Protons within Solved Crystal Structure of DNA dodecamer [d(CGCGAATTCGCG)] bound to Hoechst 33342. Protons were added to the structure for the purpose of distance measurements within the software (PDB:130D).[62] A) A5H2 (yellow) with all protons within 5Å coloured green. Distances to DH3'/DH3/DH2/DCH2. B) A18H2 (pink) with all protons within 5Å coloured green. Distances to DH4''/DH3''/DH2'''. Molecular graphics created using UCSF Chimera software.[11]

3.2.2 Original DNA Aptamer and Hoechst 33342

1D proton NMR experiments can provide information regarding the folding of the nucleic acid and are a useful first step when performing structural studies on nucleic acids. Signals within the 11-14ppm region of a 1D spectrum originate from imino protons of base paired nucleotides, therefore it is possible to estimate the number of conformations present in the sample.[120,121] Initially a 1D spectrum was acquired on a sample containing only the original DNA aptamer (Figure 3.8). There are very few peaks present in the imino region of the spectrum, indicating that the free aptamer has very minimal structure present. A titration was then performed by adding aliquots of Hoechst 33342 dye to the nucleic acid in a stepwise fashion. Changes were observed in the spectra indicating the formation of additional secondary and tertiary structures. As dye is added more peaks were observed at 11-12ppm and above 14ppm, but the peaks were still very broad and noisy (Figure 3.8D). This indicates that the Hoechst 33342 dye is binding non-specifically to the original DNA aptamer and there could be multiple conformations present, as the broad peaks are likely a result of conformational exchange.
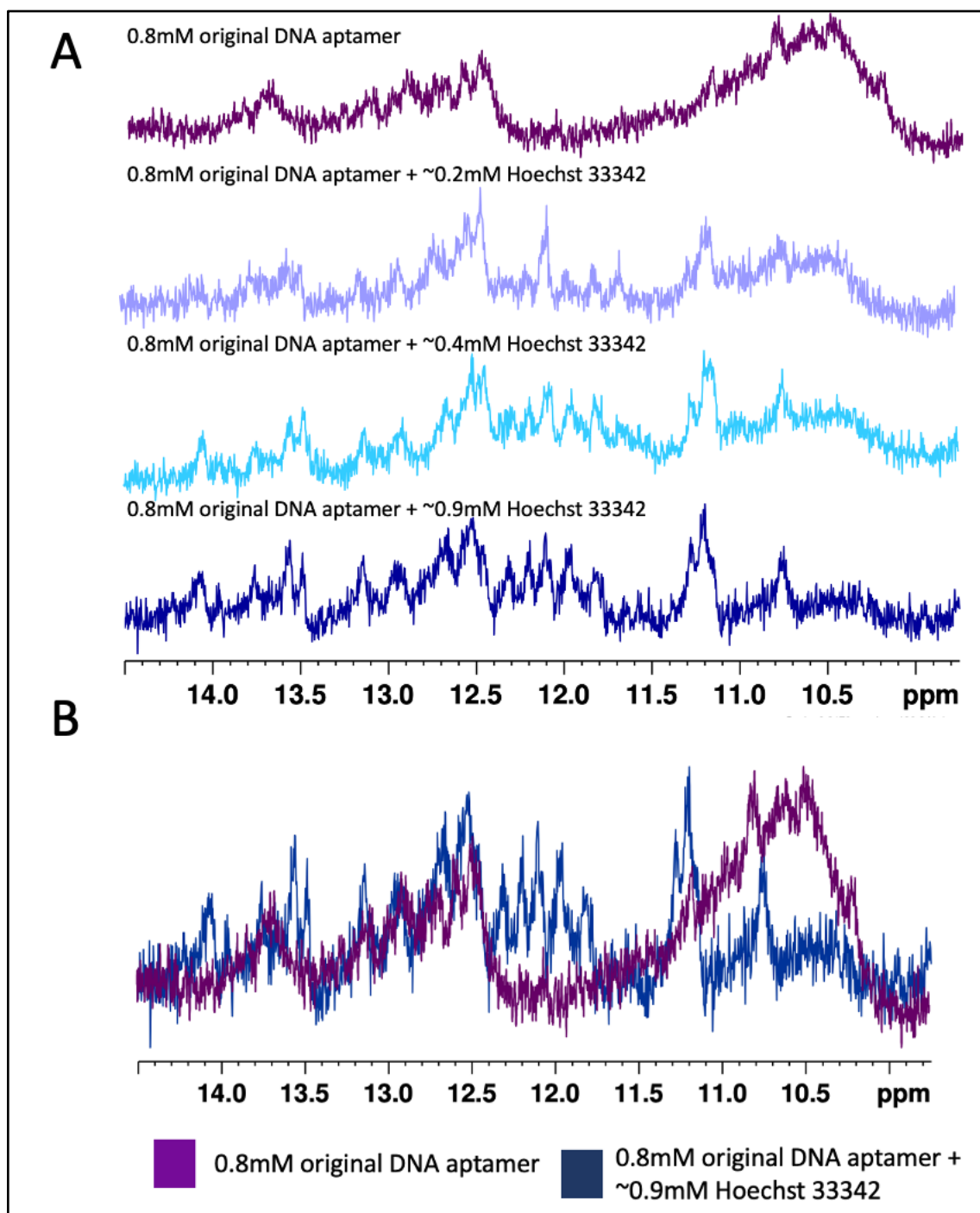
Figure 3.8 1D Proton NMR of original DNA aptamer with Hoechst 33342 dye. 1D spectra in 90% H_2O/10%D_2O using 1-1 spin echo solvent suppression, pH 6.9, at 277K. A) Stepwise titration of the original DNA Aptamer with Hoechst 33342. The dark purple sample is 0.8mM original DNA aptamer, the light purple sample is 0.8mM original DNA aptamer and 0.2mM Hoechst 33342, the light blue sample is 0.8mM original DNA aptamer and 0.4mM Hoechst 33342, and the dark blue sample is 0.8mM original DNA aptamer and 0.9mM Hoechst 33342 dye. B) Overlay of

the original DNA aptamer with no dye (dark purple) and the original DNA aptamer fully bound to the Hoechst 33342 dye (dark blue).

Although the peaks in the 1D spectra were quite broad, a 2D NOESY in 90% H_2O/10%D_2O at 277K, was performed with the original DNA aptamer bound to the Hoechst 33342 (1:1 ratio) to gain further information about the DNA-ligand complex (Figure 3.9). The spectrum showed far less peaks present than what would be expected for a sequence that is 25 nucleotides long. The evidence presented in the 2D in 90% H_2O/10%D_2O NOESY further supports the hypothesis that the original DNA aptamer is binding in multiple conformations. Based on the evidence provided in the 1D proton NMR and the 2D 90% H_2O/10%D_2O NOESY it appeared that assignments would not be possible for the original DNA aptamer and Hoechst 33342 complex therefore no further NMR studies were performed. At the time of this thesis no literature was found regarding NMR studies for the Hoechst DNA aptamer.
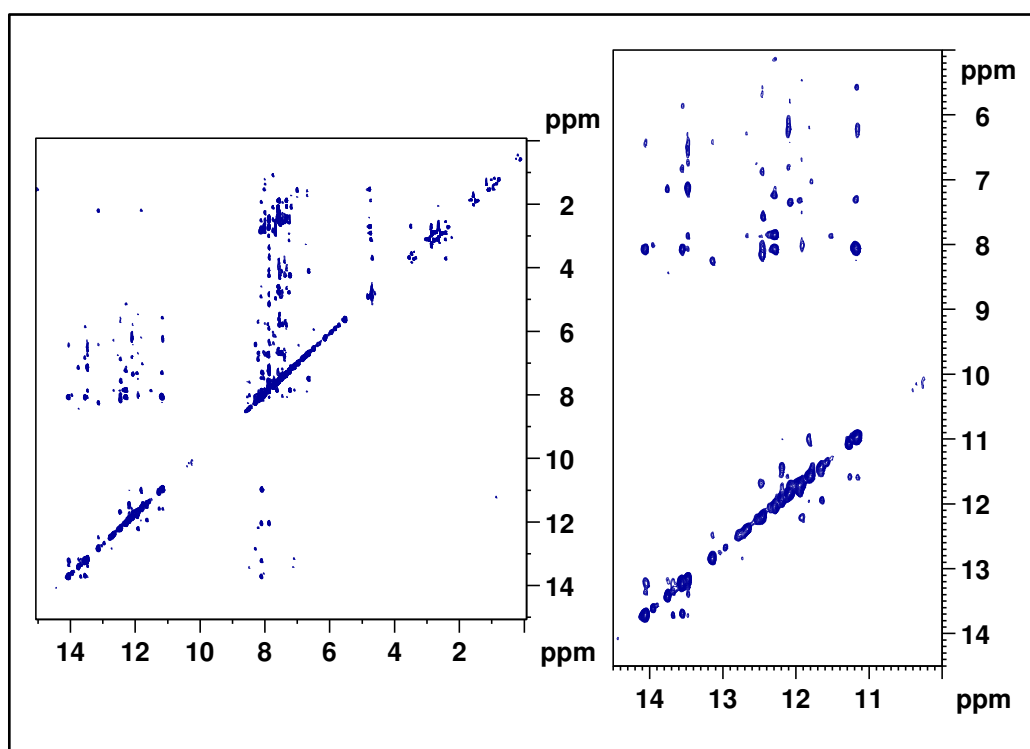
Figure 3.9 2D NOESY of original DNA aptamer bound to Hoechst 33342 dye. NOESY spectrum in 90% H_2O/10%D_2O using 1-1 spin echo solvent suppression (pH 6.9), at 277K. The

zoomed in spectra on the left shows the 11-14ppm region where peaks would be expected for stacked base pairs.

3.2.3 DNA Aptamers and the t*Bu* Hoechst dye

3.2.3.1 Original DNA Aptamer 1D Proton NMR

As discussed above, 1D proton NMR experiments are relatively affordable and fast while providing useful initial insights into a ligand aptamer interaction. A titration of the original DNA aptamer with the t*Bu* Hoechst dye was observed using 1D proton NMR. As the concentration of t*Bu* Hoechst dye was increased, the peaks in the spectra increased in number and became sharper (Figure 3.10). Based on the final 1D spectrum of 1:1 dye to DNA, it appeared that the original DNA aptamer was binding to the dye in a single conformation. The DNA appears to be fully bound to the dye at a 1:1 ratio indicating a single stranded structure bound to one ligand instead of a multi-stranded structure.[30,154] Therefore, the NMR spectrum is also not indicative of a multi-stranded structure and does not support the native gel results. The differences in results could be due to the dye not migrating into the gel or due to the lower concentration of dye in gel compared to an NMR sample. The presence of the dye could force the aptamer into a single stranded conformation. Typically if a G-quadruplex is present there will be well-defined peaks in the 10.5-12ppm region of a 1D proton NMR spectrum.[93,108,155] It is important to note that peaks within the 10.5-12ppm region of the spectrum can be a result of other structures as well, therefore it is not guaranteed the peaks are a result of a quadruplex but it does support the prediction.[93,108,120,121,138,155] In the spectrum of the original DNA aptamer bound to t*Bu* Hoechst dye there are resolved peaks within that region. This evidence could support the presence of a G-quadruplex.
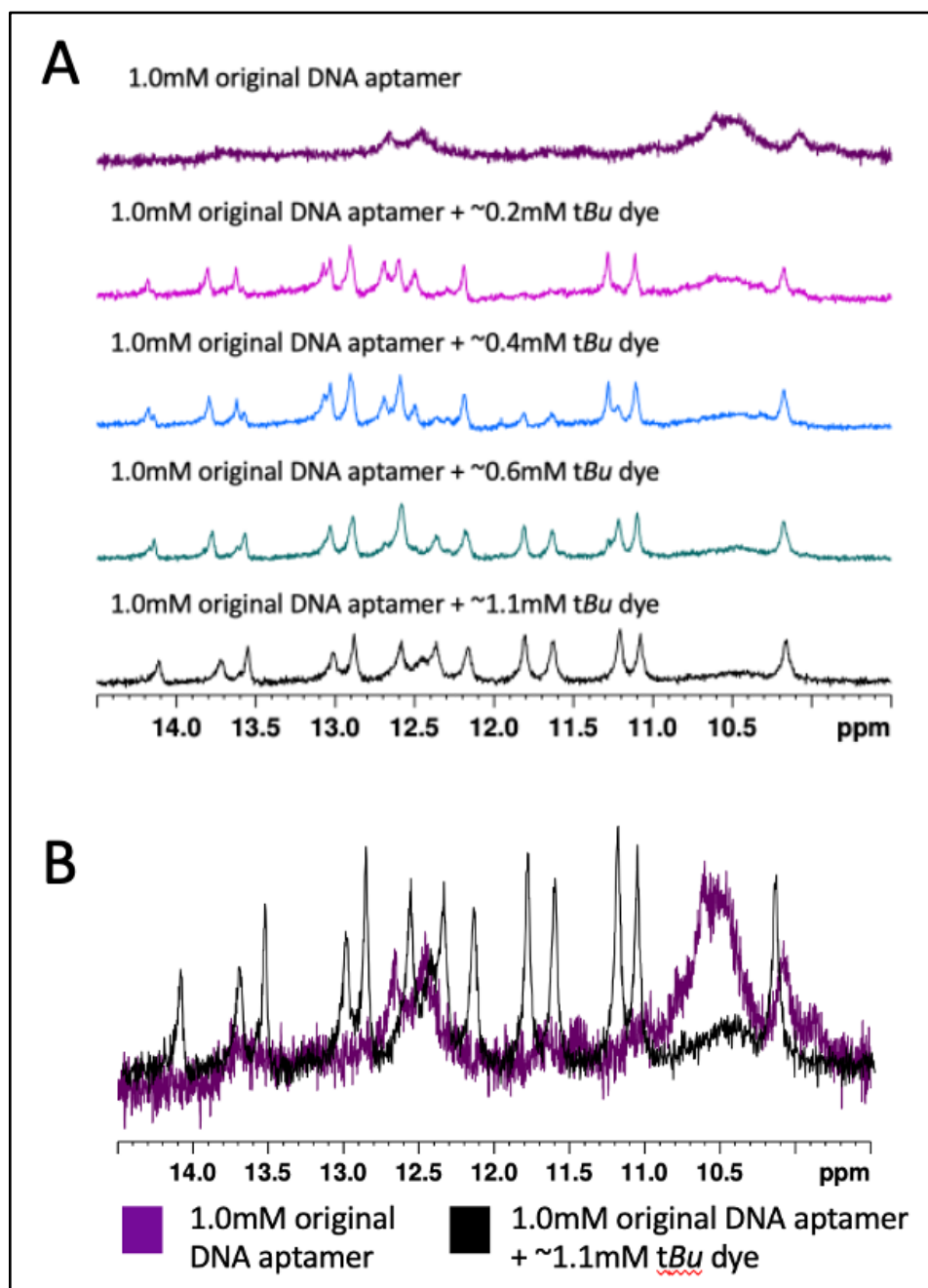
Figure 3.10 1D proton NMR of original DNA aptamer with t*Bu* Hoechst dye. 1D spectra in 90% H_2O/10%D_2O using 1-1 spin echo solvent suppression (pH 6.9), at 277K. A) Stepwise titration of the original DNA aptamer with t*Bu* Hoechst dye. The dark purple sample is 1.0mM original DNA aptamer, the light purple sample is 1.0mM original DNA aptamer and 0.2mM t*Bu* Hoechst dye, the blue sample is 1.0mM original DNA aptamer and 0.4mM t*Bu* Hoechst dye, the teal sample is 1.0mM original DNA aptamer and 0.6mM t*Bu* Hoechst dye, and the blue sample is 1.0mM original DNA aptamer and 1.1mM t*Bu* Hoechst dye. B) Overlay of the original DNA aptamer with no dye (dark purple) and the original DNA aptamer fully bound to the t*Bu* Hoechst dye (black).

There was a clear difference between the spectrum of the original DNA aptamer bound to the Hoechst 33342 versus the spectrum of the original DNA aptamer bound to the *tBu* Hoechst dye (Figure 3.11). The narrow and well-defined peaks in the *tBu* Hoechst dye spectrum indicated the binding of the dye resulting in unique tertiary structure, whereas the Hoechst 33342 spectra are very broad and noisy. It was clear that no assignment would be possible for the original DNA aptamer bound to Hoechst 33342, but structural determination of the original DNA aptamer and *tBu* Hoechst dye complex seems feasible.
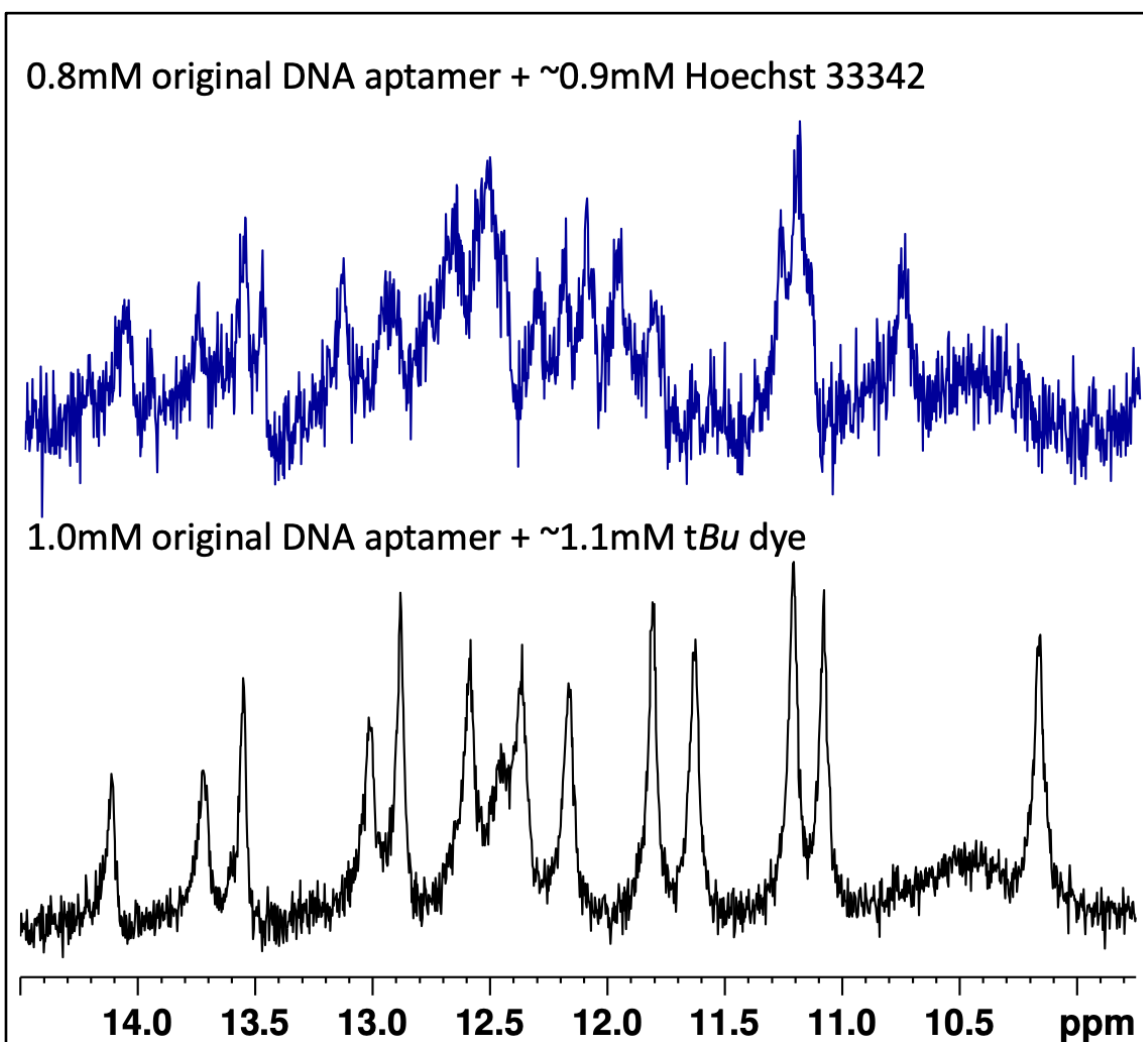
Figure 3.11 Comparison of original aptamer with Hoechst 33342 and tBu Hoechst dye. 1D spectra in 90% H_2O/10%D_2O using 1-1 spin echo solvent suppression (pH 6.9), at 277K. When the original DNA aptamer is bound to Hoechst 33342 the peaks are clearly much broader, and the

spectrum is overall noisier. The original DNA aptamer with the t*Bu* Hoechst dye shows more defined peaks and it appears that the aptamer is forming a single well-defined structure.

3.2.3.2 Extended Stem Aptamer 1D Proton NMR

The extended stem aptamer was also initially tested with a 1D proton NMR experiment. The sample was prepared for NMR by adding NMR buffer and 10%D_2O before storing the sample at ~8°C in the refrigerator overnight. At the low temperature of ~8°C the sample (similarly to Figure 3.12), therefore, to improve the solubility of the aptamer additional salt was added to a concentration of 80mM KCl and the NMR experiments were run at 283K instead of 277K (Figure 3.13). A 1:1 ratio of dye and DNA was added prior to performing any NMR, in attempt to allow the aptamer to form its desired structure with the dye, therefore it was not possible to perform a titration of the extended stem aptamer with dye. The spectrum of the extended stem aptamer with t*Bu* Hoechst dye showed a reasonable number of peaks in the 10-15ppm region of the spectrum, originating from the imino protons, based on the length of the sequence.[120,121] As previously mentioned, if a quadruplex is present, resolved peaks in the 10-12ppm region would be expected. Within the 10-12ppm region of the extended stem aptamer spectrum there are fewer peaks, but the present of a G-quadruplex cannot be completely eliminated.

Figure 3.12 Precipitation of I₂₅ aptamer in NMR tube. The white precipitate formed at the bottom of the tube after storing in the refrigerator overnight. Similar white precipitates were observed in the I$_{16}$ DNA aptamer, extended stem aptamer and the extended stem I$_{16}$ aptamer.
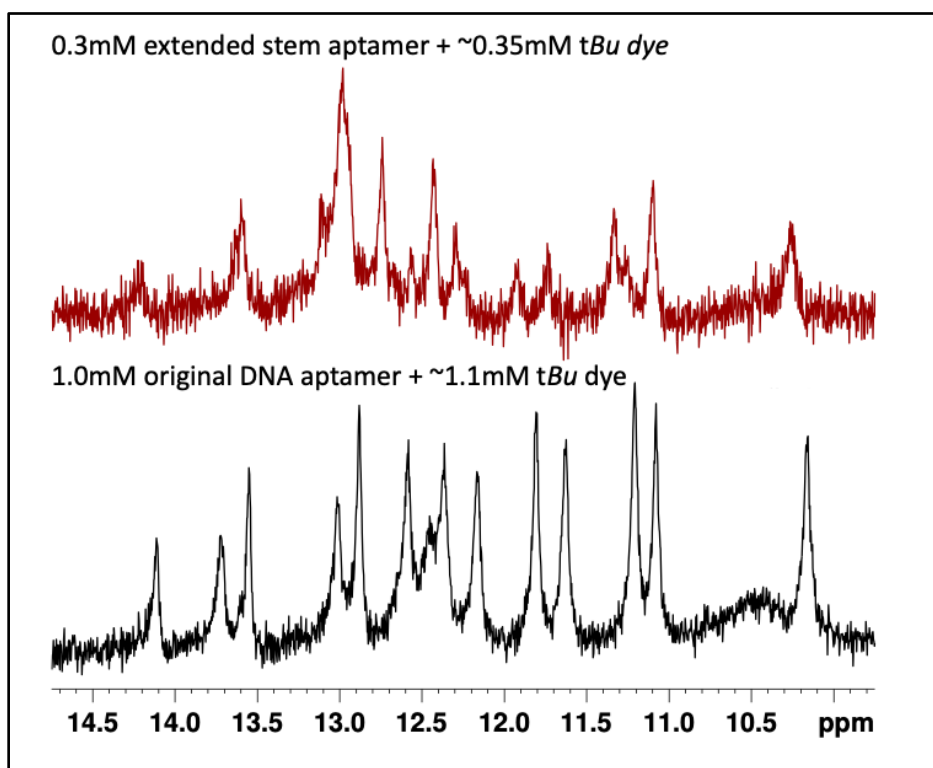
Figure 3.13 1D proton NMR of extended stem aptamer with t*Bu* Hoechst dye. 1D spectrum in 90% H$_2$O/10%D$_2$O using 1-1 spin echo solvent suppression (pH 6.9), at 283K. The peaks in the

extended stem aptamer spectrum are broader compared to the original DNA aptamer. The original DNA aptamer has more peaks present that indicate the presence of a unique tertiary structure compared to the extended stem aptamer.

3.2.3.3 t*Bu* Hoechst Dye NMR

Performing NMR experiments on a sample containing only ligand without any nucleic acid present can be beneficial to indicate what peaks are resulting from the ligand and to observe any shift that may occur upon binding. Two experiments were performed on a 5mM sample of the t*Bu* Hoechst dye: 1) a 1D proton NMR in 90% H_2O/10%D_2O (Figure 3.14), and 2) a COSY in 100% D_2O (Figure 3.15). In 1990 Searle and Embry, and later in 1993 Embry *et al*., published papers that studied the NMR spectra of Hoechst 33258 bound to a DNA duplex.[58,64] Aside from the addition of the t*Bu* groups and PEG tail on the t*Bu* Hoechst dye, the rest of the structure of dye remains the same as other Hoechst dyes, therefore the published data were used to provide direction when assigning the peaks in the spectra of the t*Bu* Hoechst dye. Both papers assigned the DH6'-DH7' and the DH6''-DH7'' peaks in approximately the 6.8-8.0ppm region of their spectra, with the DH6'-DH7' peaks appearing at a higher chemical shift value. Additionally, the papers reported the DH2'''-DH6''' and DH3'''-DH5''' peaks around 3.0-4.0ppm. It would be expected that the DH2'''-DH6''', DH3'''-DH5''', DH6'-DH7', and the DH6''-DH7'' would all produce through bond COSY or TOSCY peaks. The peaks between 3.0-4.0ppm were assigned to the PEG tail. Based on the literature values the peaks were tentatively assigned in the COSY spectrum (Figure 3.15) and the corresponding chemical shift values are listed in Table 3.1.
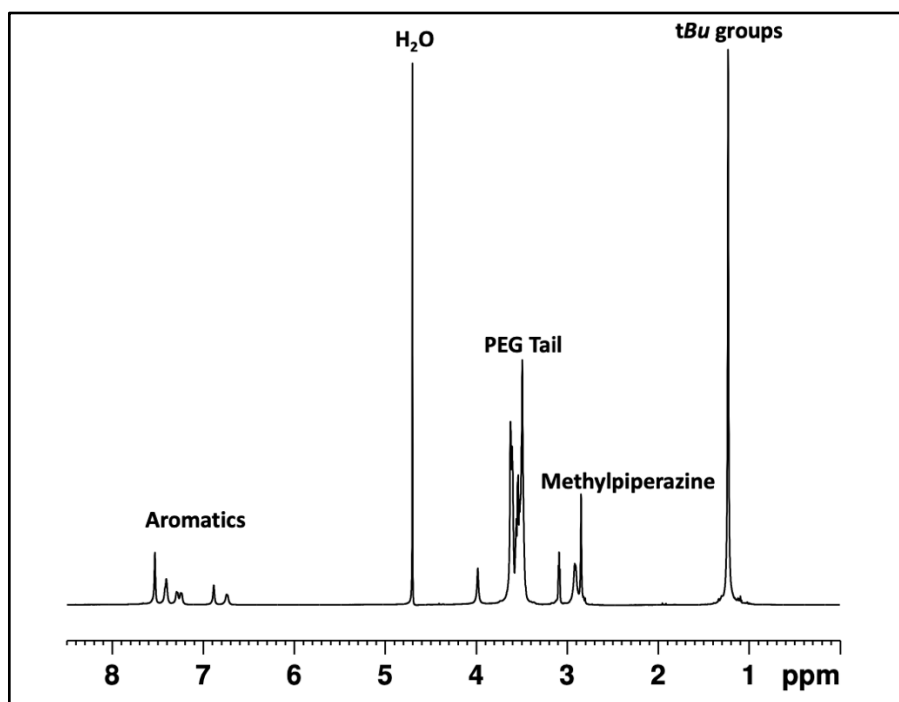
Figure 3.14 1D of t*Bu* Hoechst dye with tentative assignments. 1D spectra in 90% H_2O/10%D_2O of 5.0mM t*Bu* Hoechst dye using presaturation (pH 6.9), at 277K.
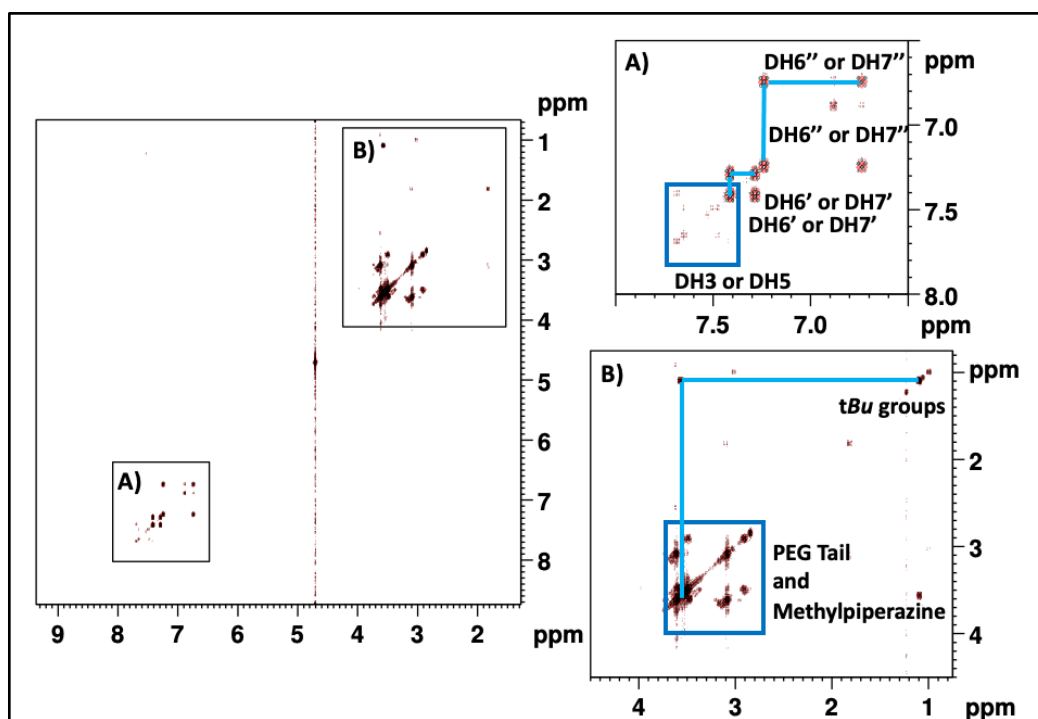
Figure 3.15 2D COSY of t*Bu* Hoechst dye. COSY spectrum in 100% D_2O of 5.0mM t*Bu* Hoechst dye sample (pH 6.9), at 298K. Blue boxes indicate approximate regions for the PEG tail and methylpiperazine groups.

Table 3.1 Corresponding Chemical Shifts of Dye Peaks

Proton	Approximate Chemical Shift of Assigned Peaks
t*Bu* Group	1.10ppm
DH6'' – DH7''	6.7/7.2ppm
DH6' – DH7'	7.3/7.4ppm

3.2.3.4 2D NMR Experiments

Once the peaks in the COSY of the t*Bu* Hoechst dye were tentatively assigned those assignments could be transferred to the TOCSY peaks of the original DNA aptamer and the extended stem aptamer, each bound to the t*Bu* Hoechst dye. This is done by overlaying the spectra (Figure 3.16). Both aptamers contained similar peaks in the 6.8-8.0ppm region of their respective spectra, that were tentatively assigned as DH6'-DH7' and DH6''-DH7'', having shifted from the unbound dye. The extended stem aptamer has an additional set of peaks that could be unbound dye, evidence of two dye molecules binding to the aptamer, or binding in more than one conformation. Each aptamer has additional TOCSY peaks in the 5.0-6.5ppm/6.5-7.5ppm region of their spectra that can be attributed to H5-H6 of the cytosines.[120,121]
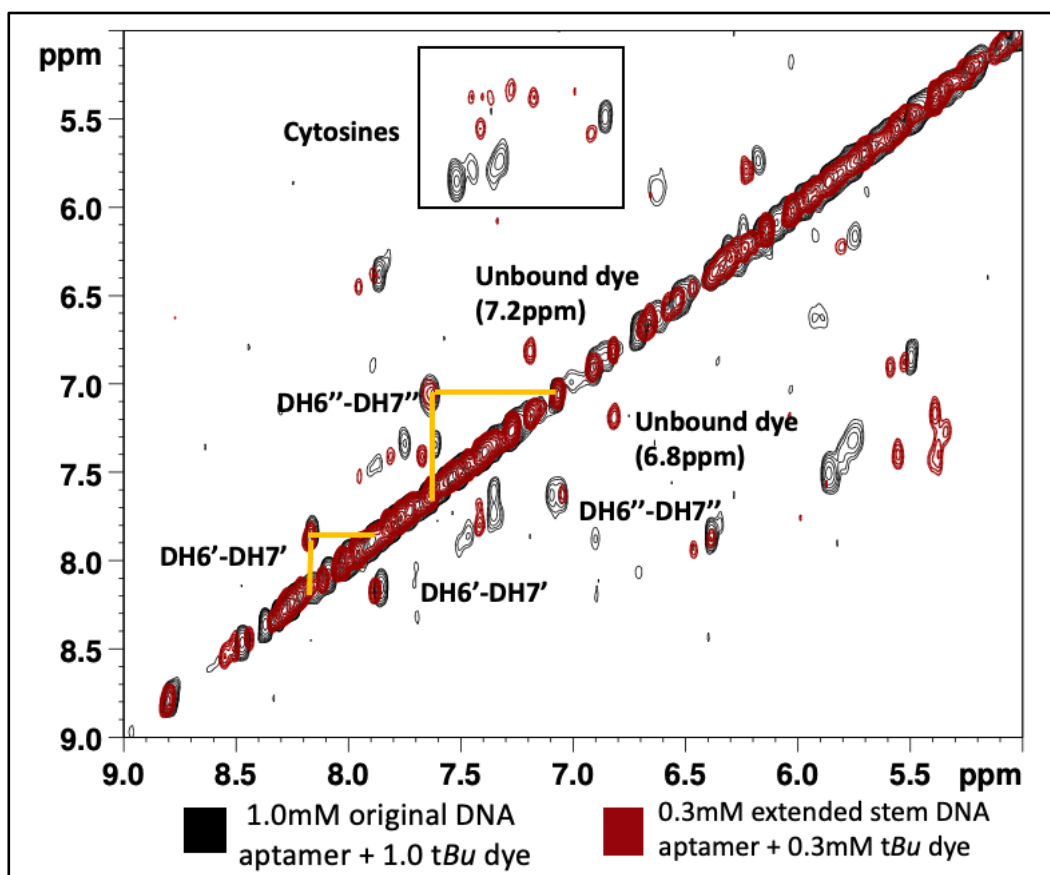
Figure 3.16 TOCSY of extended stem aptamer and original DNA Aptamer to compare and assign dye peaks. TOCSY spectrum acquired in 100% D_2O, with potassium phosphate buffer (pH 6.9), at 298K. The original DNA aptamer contains four cytosine peaks as expected.

The next step in exploring the Hoechst DNA structure was to perform a D_2O NOESY and overlay the TOCSY assignments (Figure 3.17). Although the overlaid spectra provided new information, the spectra were also useful for further supporting the dye assignments discussed above. The peak at 0.3ppm in the spectra was tentatively assigned as the t*Bu* groups. Based on the structure of the dye, one could predict expected NOE cross peaks based on distance to the t*Bu* groups. For example, the DH3 and DH5 are very close to the t*Bu* group so a cross peak would be expected, whereas the DH3''' and DH5''' are very far therefore no cross peak would be expected. Through this logic the t*Bu* peaks were used to further confirm the COSY based assignments of the dye. The assigned DH2'''-DH6''' and DH3'''-DH7''' did not have NOE cross peaks to the t*Bu* as

78

expected. The DH6''-DH7'' and DH6'-DH7' are also far from the *tBu* groups and did not produce cross peaks to the groups. In addition to confirming previous dye assignments, the D_2O NOESY also allowed for an additional tentative dye assignment. Two peaks at ~8.5ppm and ~8.8ppm did not have TOCSY peaks but did have NOEs to the *tBu* groups. These two peaks were tentatively assigned DH3-DH5 due to their proximity to the *tBu* groups and their expected lack of through bond signals. The two peaks were very intense therefore it was hypothesized that the peaks could be exchange peaks, which would later be investigated through a ROESY experiment.
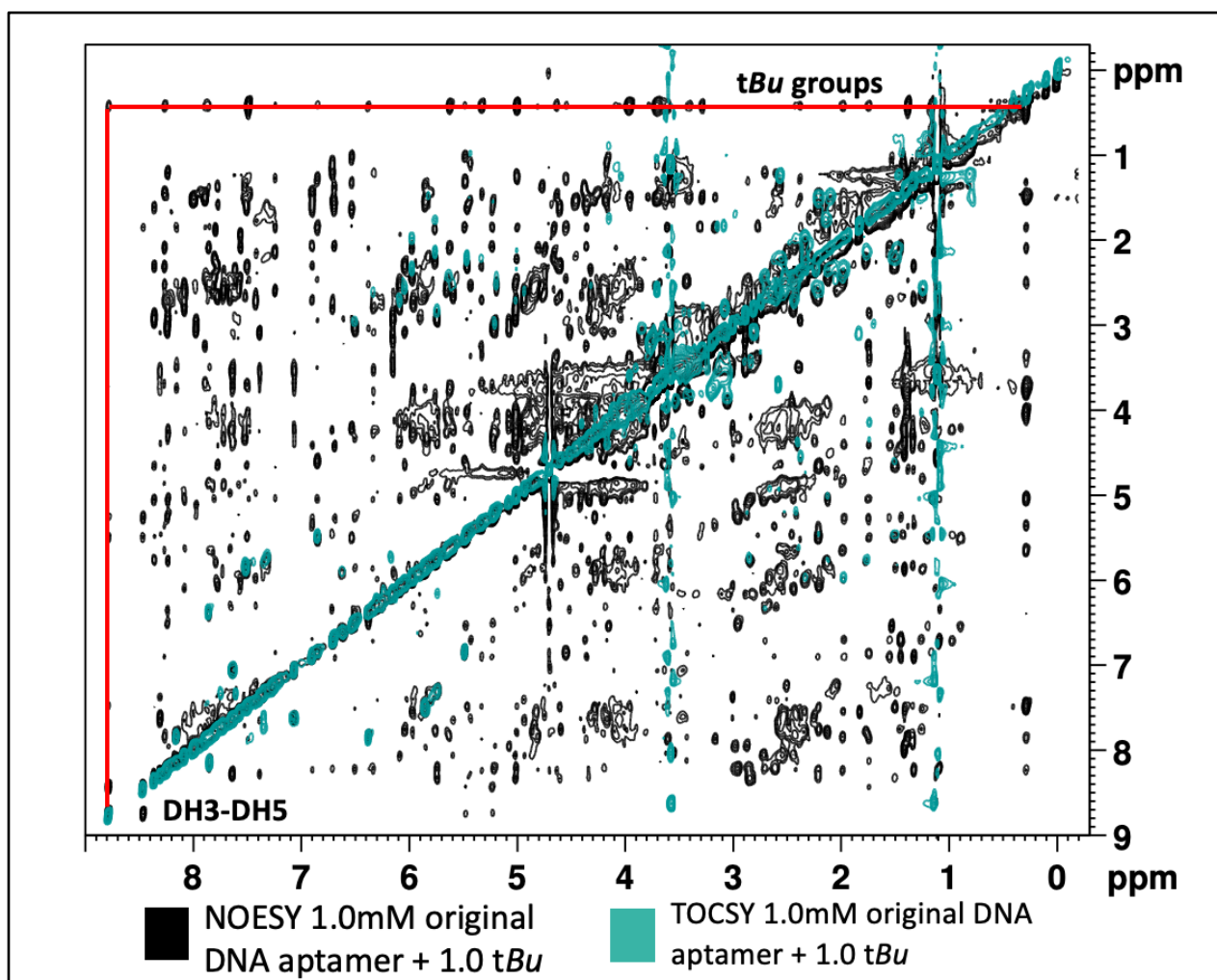
Figure 3.17 Overlay of non-exchangeable protons and TOCSY of Original DNA Aptamer.
1.0mM original DNA aptamer in 100% D_2O, with potassium phosphate buffer (pH 6.9), at 298K.

To further investigate the potential exchange peaks between the DH3-DH5 a ROESY experiment was utilized (Figure 3.18). If a rotating frame Overhauser effect (ROE) peak is a result of chemical exchange it would be expected that the peak would be positive, whereas NOE's would be negative.[138] Around ~8.5ppm there are no positive peaks present, therefore it is unlikely that the DH3-DH5 peaks seen in Figure 3.17 are exchange peaks. At approximately 0.3ppm and approximately 0.4ppm there are positive cross peaks indicating exchange peaks in this region of the spectrum. The exchange peaks are likely originating from exchange between the free and bound t*Bu* groups.
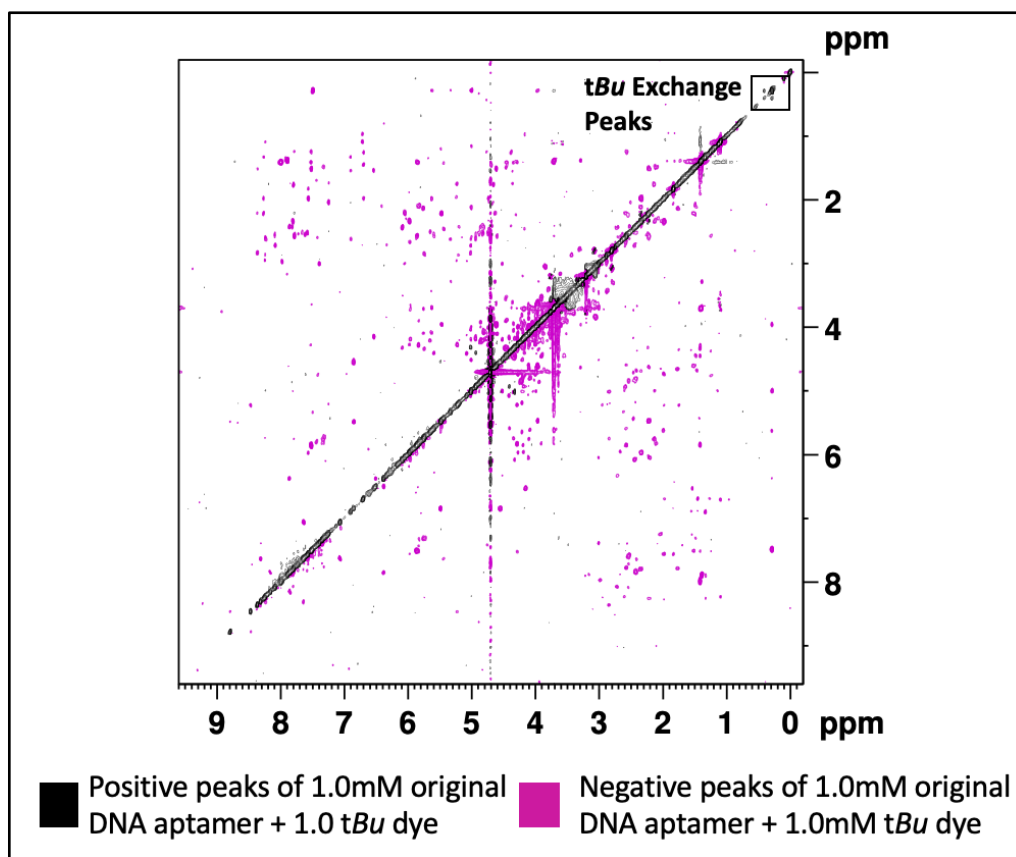
Figure 3.18 ROESY Experiment. ROESY spectrum in 100% D_2O, with potassium phosphate (pH 6.9), at 298K.

90% H_2O/10%D_2O NOESYs were also performed for the original DNA aptamer and the extended stem aptamer. Due to the precipitation issue the extend stem aptamer was run at 283K.

Within the 11-14ppm region of the spectra is where peaks from imino groups of stacked base pairs are expected. Overlaying the extended stem aptamer and original DNA aptamer show many similar peaks but a few peaks have been altered (Figure 3.19). Additionally there are fewer peaks in the extended stem aptamer. The changes in the spectrum suggests that extending the stem of the quadruplex does alter the structure.
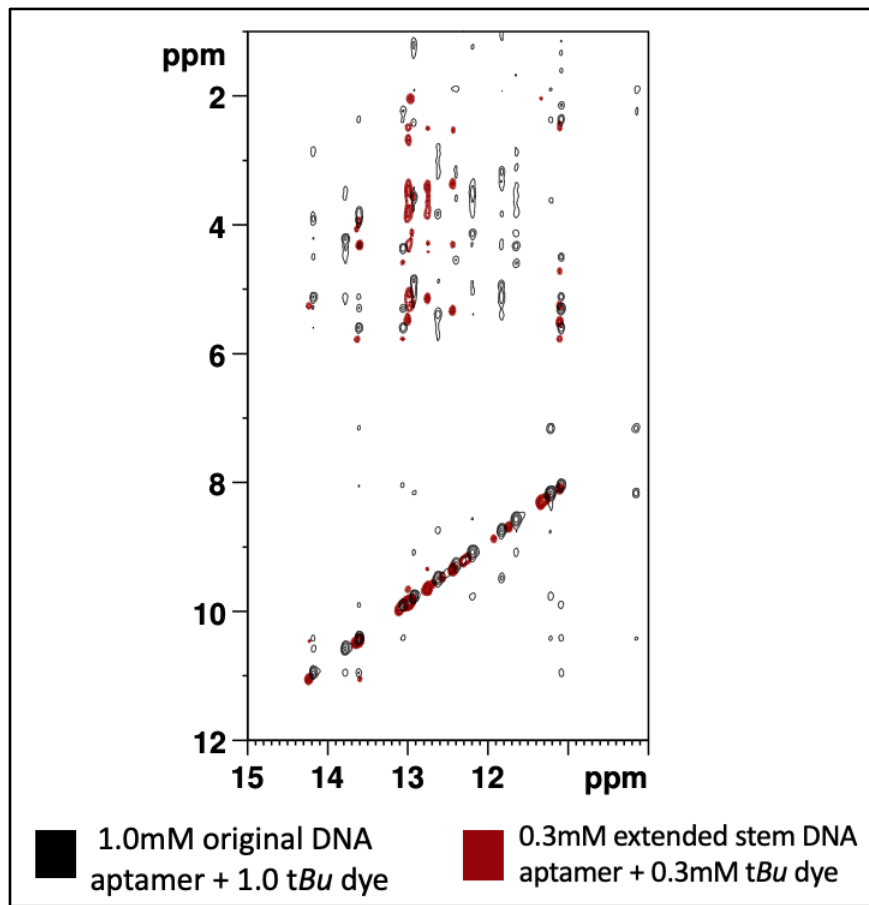
Figure 3.19 Comparison of exchangeable protons in the original DNA Aptamer and the extended stem aptamer. NOESY spectra in $90\%H_2O/10\%D_2O$ using 1-1 spin echo solvent suppression, potassium phosphate buffer (pH 6.9), at with extended stem aptamer at 283K and original DNA aptamer at 277K.

3.2.3.5 Inosine Substitution

Two versions of the original DNA aptamer with inosine substitution were obtained, I_{16} DNA aptamer and I_{25} DNA aptamer. After the samples were prepared for 1D NMR and stored in

the fridge overnight, it was observed that a precipitate was forming in both samples. The salt concentration was increased in attempt to reduce precipitate formation. At a concentration of 80mM KCl and 10mM phosphate buffer I_{16} DNA aptamer stayed in solution. For I_{25} DNA aptamer the salt concentration was increased to 150mM KCl but the sample continued to precipitate therefore no NMR experiments could be performed. Since I_{16} DNA aptamer remained in solution at an increased salt concentration it was possible to perform NMR experiments with the sample.

As with previous samples a 1D 1H NMR of the I_{16} DNA aptamer with t*Bu* Hoechst dye and a 2D 90% H_2O/10%D_2O NOESY was aquired. These spectra were overlaid with the original DNA aptamer for comparison. Both the 1D proton NMR and the 2D 90% H_2O/10%D_2O NOESY contained broad peaks (Figure 3.20). The broad peaks and the precipitate forming further support the potential formation of a multi stranded structure, such as a dimer or a tetramer. There are differences between the 1D spectra indicating the I_{16} substitution is causing alterations to the structure. In the 2D NOESY there are peaks in the 7-9ppm/10-14ppm region that correspond to three adenine thymine base pairs in a stem. Within the 7-9ppm/10-14ppm region of the I_{16} spectrum, peaks from the H2 inosine to the H8 on the guanine would be expected.[73] Also within the 7-9ppm/10-14ppm region of the spectrum there would be a very intense and sharp peak from the inosine H2 to the imino of the inosine, if the imino is involved in a hydrogen bonded base pair.[73,156] There are no addition peaks in this region that could correspond to the inosine-guanine or to intra inosine peaks. The lack of these peaks could indicate that the I_{16} is not involved in the quadruplex, but it could also indicate that the inosine is disrupting the structure. Additionally, the I_{16} spectrum was obtained at a lower concentration compared to the original DNA aptamer, which could also impact the sensitivity of the spectrum. In both the I_{16} DNA aptamer and the original

DNA aptamer the adenine thymine base pair peaks are present therefore it does not appear that the

structure is completely disrupted by the inosine substitution (Figure 3.21).
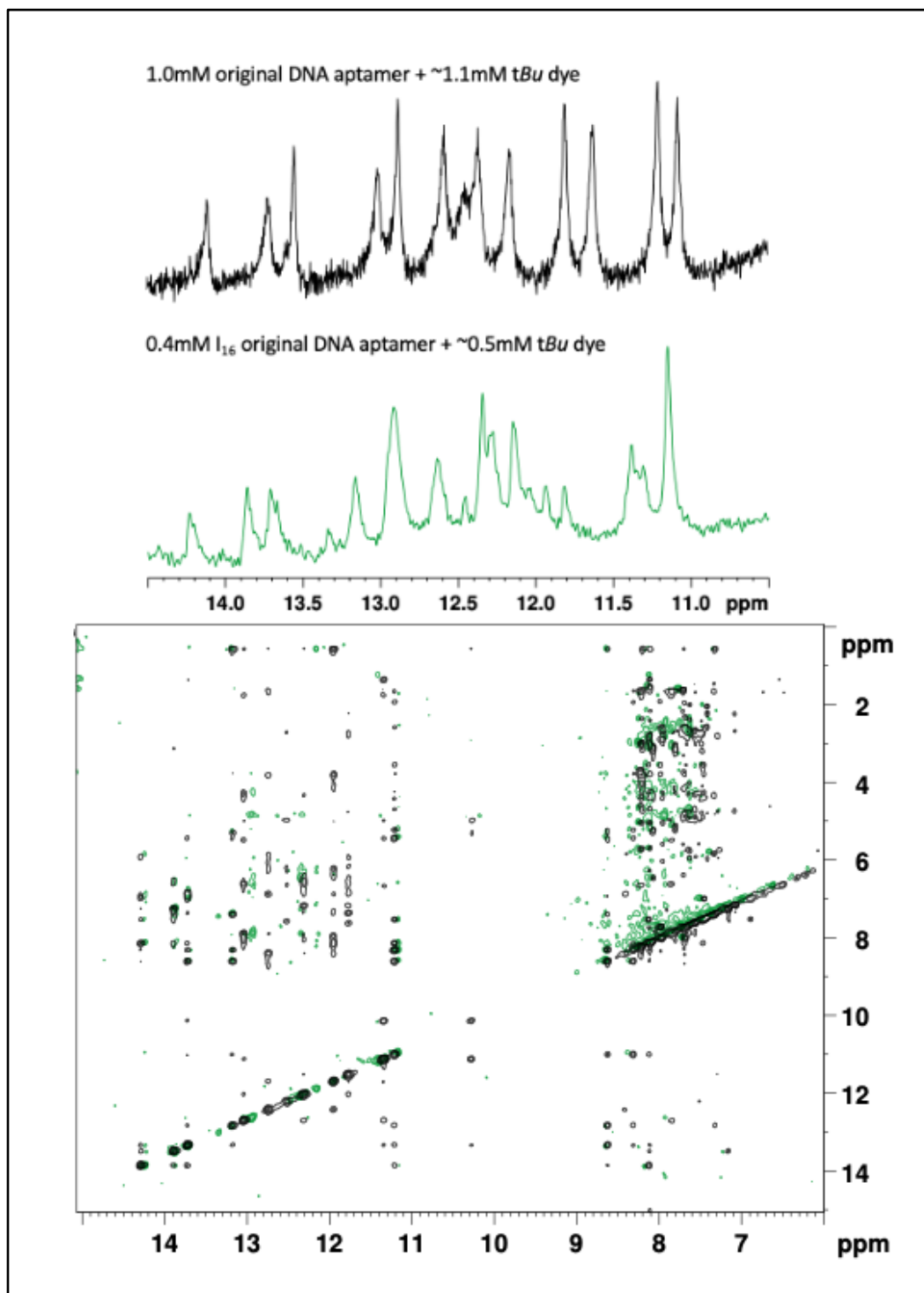
Figure 3.20 I$_{16}$ DNA Aptamer compared to original DNA Aptamer. Original DNA aptamer
1D in 90% H$_2$O/10%D$_2$O using 1-1 spin echo solvent suppression, potassium phosphate buffer

(pH 6.9), at 277K. I_{16} DNA aptamer 1D in 90% H_2O/10%D_2O using 1-1 spin echo solvent suppression, pH 6.9, at 283K. Black spectra are the original DNA aptamer. Green spectra are the I_{16} DNA aptamer. NOESY spectra in 90%H_2O/10%D_2O using 1-1 spin echo solvent suppression, potassium phosphate buffer (pH 6.9), with original DNA aptamer at 277K and I_{16} DNA aptamer at 283K.
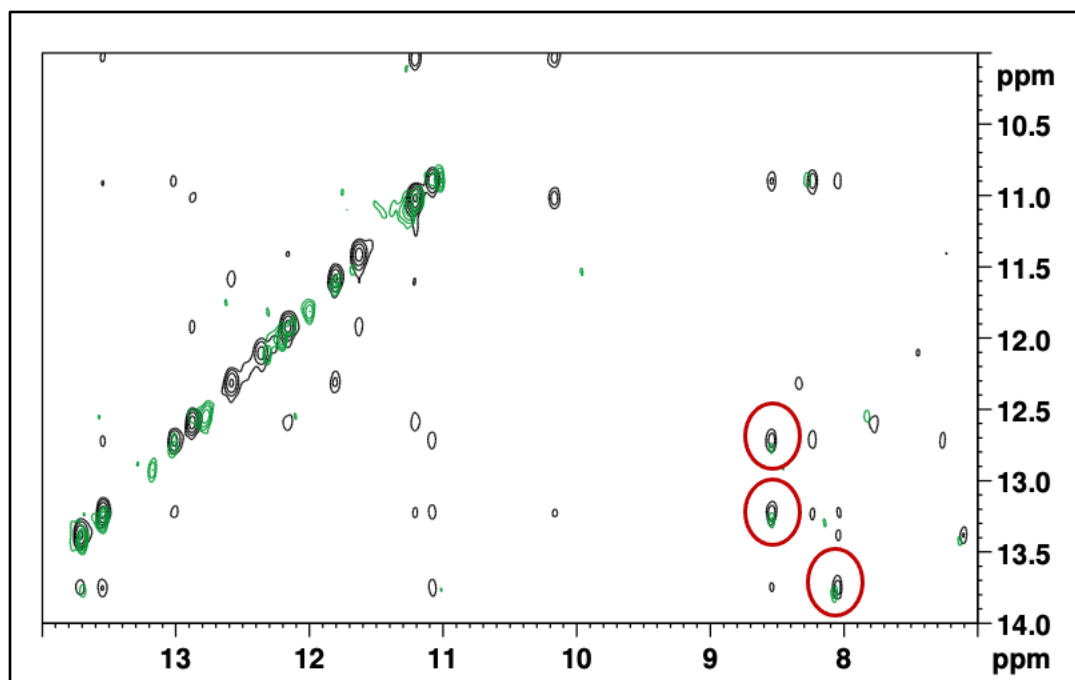
Figure 3.21 Zoomed in H_2O NOESY. The peaks circled in red are tentatively assigned as the H2 peaks from bases paired adenines.

Similarly to previous samples, the I_{16} extended stem aptamer precipitated in the refrigerator around ~8°C therefore the same steps were taken as the extended stem aptamer: running the NMR at 283K instead of 277K, adding *tBu* Hoechst dye to 1:1 ratio before cooling instead of doing a titration, and increasing the salt concentration to 80mM KCl instead of 100mM KCl. Once the I_{16} extended stem aptamer was dissolved and remained in solution, a 1D proton NMR and a 2D 90% H_2O/10%D_2O NOESY were run (Figure 3.22). Similarly to the I_{16} DNA aptamer, the I_{16} extended stem aptamer did not contain additional peaks in the 7-9ppm/10-14ppm which could be due to a disruption in the quadruplex or it could be that I_{16} is not involved in the quadruplex. The extended stem aptamer and I_{16} extended stem aptamer 1Ds are relatively similar, indicating that the I_{16}

extended stem modification does not alter the structure as drastically as the I_{16} DNA aptamer substitution does.
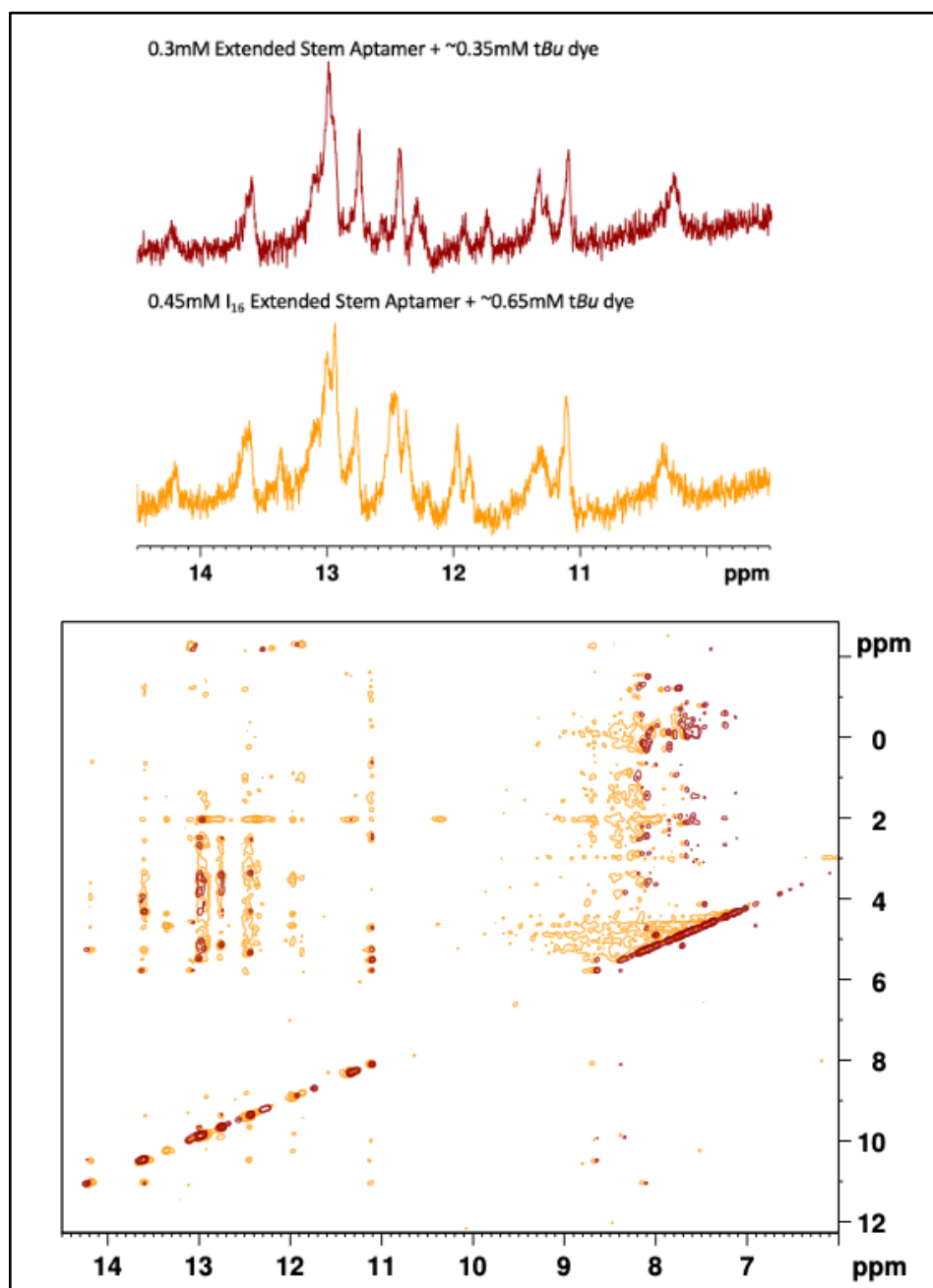
Figure 3.22 I_{16} Extend Stem Aptamer compared to Extended Stem Aptamer. 1D in 90% $H_2O/10\%D_2O$ using 1-1 spin echo solvent suppression, potassium phosphate buffer (pH 6.9), at 283K. Red spectra are the extended stem aptamer. Orange spectra are the I_{16} extended stem aptamer. NOESY spectra in 90%$H_2O/10\%D_2O$ using 1-1 spin echo solvent suppression, potassium phosphate buffer (pH 6.9), at with extended stem aptamer and I_{16} extended stem aptamer at 283K.

3.3 Conclusions

The first NMR experiments presented in this chapter, aimed to characterize the original DNA aptamer with Hoechst 33342. However, the NMR spectra were noisy, and the peaks were broad, especially when compared to the original DNA aptamer with t*Bu* Hoechst dye. The reason for the poor spectra is two-fold. First, the Hoechst 33342 was developed to bind to dsDNA and is not specific for the original DNA aptamer. Secondly, the original DNA aptamer was selected for the t*Bu* Hoechst dye, therefore Hoechst 33342 was not the intended target for the aptamer. Although the experiments of the original DNA aptamer and Hoechst 33342 did not show much promise, the experiments with original DNA aptamer and t*Bu* Hoechst dye did.

By overlaying the COSY of the t*Bu* Hoechst dye with the TOCSY of the extended stem aptamer and the original DNA aptamer, it was possible to tentatively assign the t*Bu* Hoechst dye peaks. Based on the computational models of the hypothesized G-quadruplexes presented in Chapter 2, inosine substituted samples were tested. Unfortunately, the I_{25} DNA aptamer continually precipitated out of solution therefore no NMR could be performed. The resulting NMR spectrum of I_{16} DNA aptamer had many similarities compared to the non-inosine substituted samples, indicating that the I_{16} substitution was not completely altering the structure of the aptamer.

The extended stem aptamer shows some differences when compared to the original DNA aptamer. The original DNA aptamer produced more well defined signals in the 10.5-12ppm region that are more consistent with a quadruplex structure compared to the extended stem aptamer. Overall, the NMR spectra presented in this chapter are consistent with the presence of a G-quadruplex structure due to the presence of peaks in the 10.5-12ppm region of the spectra. Based on the NMR spectra presented in this chapter, elucidation of a structure should be achievable for

the original DNA aptamer bound to the t*Bu* Hoechst dye. Well assignments were attempted, further

research is required for concrete assignments.

Chapter 4   Summary and Future Work

When the Hoechst DNA aptamer was first selected, a secondary structure consisting of a stem and loop was hypothesised. The stem was believed to provide a binding site and the loop was believed to provide adequate space for the bulky *tBu* groups on the dye to bind, but no formal structural studies were ever performed on the aptamer. This project aimed to better characterize the Hoechst DNA aptamer through biochemical and structural methods and lay the groundwork for detailed structural studies by NMR spectroscopy. Early on it was noted that the sequence was rich in guanines, providing the potential of a G-quadruplex. The hypothesis of a G-quadruplex focused on a single stranded structure and in Chapter 2 computational modeling was employed to test the feasibility of a quadruplex in the structure. A combination of online tools and modeling software supported the hypothesis of a single-stranded G-quadruplex and inosine substitutions were performed based on those models. Interestingly, the native gels presented in Chapter 2 indicated that the original DNA aptamer was forming a multi-stranded structure. As discussed previously, it is not possible to determine how many strands of DNA are present in the structure in the native gel, as migration is a result of multiple factors. The CD experiments in Chapter 2 suggested the presence of a parallel G-quadruplex. The native gel and CD provided the groundwork for the conclusion that the main structure the Hoechst DNA aptamer is forming a multi-stranded G-quadruplex structure under certain conditions.

Chapter 3 focused on gaining structural information through NMR experiments. There was a clear difference between the spectra of the original DNA aptamer with the Hoechst 33342 compared to the *tBu* Hoechst dye, with Hoechst 33342 producing much broader peaks and nosier spectra. When comparing all the variations of the Hoechst aptamer, it was noted that the NMR spectra for the original DNA aptamer with t*Bu* Hoechst dye produced the clearest peaks, indicating

a well-defined tertiary structure was forming, therefore assignment should be possible. While the NMR spectra support the presence of a G-quadruplex, they are not indicative of a multi-stranded structure. The reasoning for the presence of a single stranded structure in NMR is due to the aptamer appearing fully bound at a 1:1 ratio of DNA to dye. This is contradictory to the native gel results obtained in Chapter 2. Although native gels are not intended to denature DNA, the gels do heat up which could result in the separation of the DNA from the dye. The Hoechst dye is not negatively charged, and it would not move through the gel, therefore the structure observed in the native gel may not be representative of a fully bound aptamer-ligand complex. The evidence provided in the NMR experiments suggests that the aptamer equilibrates to a monomer at higher ligand concentrations. The K_d value of 4.7μM calculated by a fluorescence titration may support separation of the dye from the aptamer within the gel and therefor future experiments could study increased dye concentrations with the native gel sample.[157,158] Through a COSY it was possible to assign the t*Bu* Hoechst dye, and to transfer the dye assignments to spectra of the original DNA aptamer and the extended stem aptamer. Furthermore, inosine substituted samples were observed to contain many altered peaks for both the original DNA aptamer and the extended stem aptamer.

While the work presented was able to provide insight into the previously undetected G-quadruplex, there is still potential for additional research. As discussed in Chapter 1, potassium or sodium is required to help stabilize G-quadruplex structures. In the future, experiments to observe the difference between CD spectra with sodium vs potassium could provide additional information regarding the quadruplex present.[89] Similarly comparing 1D proton NMR spectra with sodium, potassium, and low salt could also provide additional context regarding the G-quadruplex.[89,91] It could also be interesting to expand the result to include UV melting curves with different salts present.[93–95] Additionally, the inosine samples did not appear to be involved in a G-quadruplex but

running CD on the sample could indicate whether a quadruplex is present. There is still much work to be done characterizing this aptamer and ultimately, the end goal for future research would be to elucidate a solved structure for the original DNA aptamer bound to t*Bu* Hoechst dye.

Reference

(1)   F. H. Crick. On Protein Synthesis. *Symp. Soc. Exp. Biol.* **1958**, *12*, 138–163. https://doi.org/10.1038/227561a0.

(2)   Crick, F. Central Dogma of Molecular Biology. *Nature* **1970**, *227*, 561–563. https://doi.org/10.1007/978-1-4020-6754-9_2672.

(3)   Lodish, Harvey; Berk. Arnold; Matsudaira, Paul; Kaiser, Chris; Krieger, Monty; Scott, Matthew; Zipursky, Lawerence; Darnell, J. *Molecular Cell Biology*, 5th edition; Freeman and Company: New York: W.H., 2003. p 40, 101-105, 499, 518.

(4)   Batey, R. T.; Rambo, R. P.; Doudna, J. A. Tertiary Motifs in RNA Structure and Folding. *Angew. Chem. Int. Ed.* **1999**, *38*, 2326–2343. https://doi.org/10.1002/(SICI)1521-3773(19990816)38:16<2326::AID-ANIE2326>3.0.CO;2-3.

(5)   Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag New York Inc: New York, 1984. https://doi.org/10.1007/978-1-4612-5190-3. p 13-24, 220-244.

(6)   SantaLucia, J.; Hicks, D. The Thermodynamics of DNA Structural Motifs. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415–440. https://doi.org/10.1146/annurev.biophys.32.110601.141800.

(7)   Neidle, S. *Oxford Handbook of Nucleic Acid Structure*; Oxford University Press: New York, 1998.

(8)   Ussery, D. W. DNA Structure: A-, B- and Z-DNA Helix Families. *Encycl. Life Sci.* **2002**. https://doi.org/10.1038/npg.els.0003122.

(9)   Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Structure of a B-DNA Dodecamer: Conformation and Dynamics. *Proc. Natl. Acad. Sci. USA* **1981**, *78* (4), 2179–2183. https://doi.org/10.1073/pnas.78.4.2179.

(10) Klostermen, P. S.; Shah, S. A.; Steitz, T. A. Crystal Structures of Two Plasmid Copy Control Related RNA Duplexes : An 18 Base Pair Duplex at 1 . 20 Å Resolution and a 19 Base Pair Duplex at 1 . 55 Å Resolution. *Biochemistry* **1999**, *38*, 14784–14792. https://doi.org/10.1021/bi9912793.

(11) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J.Comput.Chem.* **2004**, *25* (13), 1605–1612. https://doi.org/10.1002/jcc.20084.

(12) Cech, T. R. RNA as an Enzyme. *Biochem. Int.* **1986**, *18* (5), 64–75. https://doi.org/10.1016/s0009-9120(03)00064-x.

(13) Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. The RNA Moiety of Ribonuclease P Is the Catalytic Subunit of the Enzyme. *Cell* **1983**, *35*, 849–857. https://doi.org/10.1016/0092-8674(83)90117-4.

(14) Chen, Y.; Qi, F.; Gao, F.; Cao, H.; Xu, D.; Salehi-Ashtiani, K.; Kapranov, P. Hovlinc Is a Recently Evolved Class of Ribozyme Found in Human LncRNA. *Nat. Chem. Biol.* **2021**, *17*, 601–607. https://doi.org/10.1038/s41589-021-00763-0.

(15) Tuerk, C.; Gold, L. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science.* **1990**, *249* (4968), 505–510. https://doi.org/10.1126/science.2200121.

(16) Ellington, A. D.; Szostak, J. W. In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990**, *346* (6287), 818–822. https://doi.org/10.1038/346818a0.

(17) Cowperthwaite, M. C.; Ellington, A. D. Bioinformatic Analysis of the Contribution of Primer Sequences to Aptamer Structures. *J. Mol. Evol.* **2008**, *67* (1), 95–102. https://doi.org/10.1007/s00239-008-9130-4.

(18)   Nonaka, Y.; Sode, K.; Ikebukuro, K. Screening and Improvement of an Anti-VEGF DNA Aptamer Www.Mdpi.Com/Journal/ Molecules. *Molecules* **2010**, *15* (1), 215–225. https://doi.org/10.3390/molecules15010215.

(19)   Jeon, S. H.; Kayhan, B.; Ben-Yedidia, T.; Arnon, R. A DNA Aptamer Prevents Influenza Infection by Blocking the Receptor Binding Region of the Viral Hemagglutinin. *J. Biol. Chem.* **2004**, *279* (46), 48410–48419. https://doi.org/10.1074/jbc.M409059200.

(20)   Sando, S.; Narita, A.; Aoyama, Y. Light-up Hoechst-DNA Aptamer Pair: Generation of an Aptamer-Selective Fluorophore from a Conventional DNA-Staining Dye. *ChemBioChem* **2007**, *8* (15), 1795–1803. https://doi.org/10.1002/cbic.200700325.

(21)   Sando, S.; Narita, A.; Hayami, M.; Aoyama, Y. Transcription Monitoring Using Fused RNA with a Dye-Binding Light-up Aptamer as a Tag: A Blue Fluorescent RNA. *Chem.Comm.* **2008**, No. 33, 3858–3860. https://doi.org/10.1039/b808449a.

(22)   Liss, M.; Petersen, B.; Wolf, H.; Prohaska, E. An Aptamer-Based Quartz Crystal Protein Biosensor. *Anal.Chem.* **2002**, *74* (17), 4488–4495. https://doi.org/10.1021/ac011294p.

(23)   Stoltenburg, R.; Schubert, T.; Strehlitz, B. In Vitro Selection and Interaction Studies of a DNA Aptamer Targeting Protein A. *PLoS One* **2015**, *10* (7), e0134403. https://doi.org/10.1371/journal.pone.0134403.

(24)   Long, F.; Gao, C.; Shi, H. C.; He, M.; Zhu, A. N.; Klibanov, A. M.; Gu, A. Z. Reusable Evanescent Wave DNA Biosensor for Rapid, Highly Sensitive, and Selective Detection of Mercury Ions. *Biosens. Bioelectron.* **2011**, *26* (10), 4018–4023. https://doi.org/10.1016/j.bios.2011.03.022.

(25)   Dolgosheina, E. V.; Jeng, S. C. Y.; Panchapakesan, S. S. S.; Cojocaru, R.; Chen, P. S. K.; Wilson, P. D.; Hawkins, N.; Wiggins, P. A.; Unrau, P. J. RNA Mango Aptamer-

Fluorophore: A Bright, High-Affinity Complex for RNA Labeling and Tracking. *ACS Chem Biol.* **2014**, *9* (10), 2412–2420. https://doi.org/10.1021/cb500499x.

(26)   Grate, D.; Wilson, C. Laser-Mediated, Site-Specific Inactivation of RNA Transcripts. *Proc. Natl. Acad. Sci. USA* **1999**, *96* (11), 6131–6136. https://doi.org/10.1073/pnas.96.11.6131.

(27)   Sassanfar, M.; Szostak, J. W. An RNA Motif That Binds ATP. *Nature* **1993**, *364*, 550–553. https://doi.org/10.1038/364550a0.

(28)   Huizenga, D. E.; Szostak, J. W. A DNA Aptamer That Binds Adenosine and ATP. *Biochemistry* **1995**, *34* (2), 656–665. https://doi.org/10.1021/bi00002a033.

(29)   Dieckmann, T.; Suzuki, E.; Nakamura, G. K.; Feigon, J. Solution Structure of an ATP-Binding RNA Aptamer Reveals a Novel Fold. *RNA* **1996**, *2*, 628–640.

(30)   Lin, C. H.; Patel, D. J. Structural Basis of DNA Folding and Recognition in an AMP-DNA Aptamer Complex: Distinct Architectures but Common Recognition Motifs for DNA and RNA Aptamers Complexed to AMP. *Chem. Biol.* **1997**, *4* (11), 817–832. https://doi.org/10.1016/S1074-5521(97)90115-0.

(31)   Trachman, R. J.; Autour, A.; Jeng, S. C. Y.; Abdolahzadeh, A.; Andreoni, A.; Cojocaru, R.; Garipov, R.; Dolgosheina, E. V.; Knutson, J. R.; Ryckelynck, M.; Unrau, P. J.; Ferré-D'Amaré, A. R. Structure and Functional Reselection of the Mango-III Fluorogenic RNA Aptamer. *Nat.Chem.Biol.* **2019**, *15* (5), 472–479. https://doi.org/10.1038/s41589-019-0267-9.

(32)   Da Costa, J. B.; Dieckmann, T. Entropy and Mg2+ Control Ligand Affinity and Specificity in the Malachite Green Binding RNA Aptamer. *Mol.BioSyst.* **2011**, *7*, 2156–2163. https://doi.org/10.1039/c1mb05075c.

(33)   Duchardt-Ferner, E.; Juen, M.; Bourgeois, B.; Madl, T.; Kreutz, C.; Ohlenschläger, O.;

Wöhnert, J. Structure of an RNA Aptamer in Complex with the Fluorophore Tetramethylrhodamine. *Nucleic Acids Res.* **2020**, *48* (2), 949–961. https://doi.org/10.1093/nar/gkz1113.

(34) Patel, D. J.; Suri, A. K.; Jiang, F.; Jiang, L.; Kumar, P. F. R. A.; Nonin, S. Structure, Recognition and Adaptive Binding in RNA Aptamer Complexes. *J.Mol. Biol.* **1997**, *272*, 645–664. https://doi.org/10.1006/jmbi.1997.1281.

(35) Harmann, T.; Dinshaw, P. Adaptive Recognition by Nucleic Acid Aptamers. *Mater. Sci. Eng. Collect.* **2000**, *287* (5454), 820–825.

(36) Costa, J. B. Da; Andreiev, A. I.; Dieckmann, T. Thermodynamics and Kinetics of Adaptive Binding in the Malachite Green RNA Aptamer. *Biochemistry* **2013**, *52*, 6575–6583. https://doi.org/10.1021/bi400549s.

(37) Gilbert, S. D.; Reyes, F. E.; Edwards, A. L.; Batey, R. T. Adaptive Ligand Binding by the Purine Riboswitch in the Recognition of Guanine and Adenine Analogs. *Structure* **2009**, *17* (6), 857–868. https://doi.org/10.1016/j.str.2009.04.009.

(38) Paige, J.; Wu, K.; Jaffrey, S. RNA Mimics of Green Fluorescent Protein. *Science (80-.).* **2011**, *333* (6042), 642–646. https://doi.org/10.1126/science.1207339.

(39) Warner, K. D.; Chen, M. C.; Song, W.; Strack, R. L.; Thorn, A.; Jaffrey, S. R.; Amaré, A. R. F. Structural Basis for Activity of Highly Efficient RNA Mimics of Green Fluorescent Protein. *Nat. Struct. Mol. Biol.* **2014**, *21* (8), 658–665. https://doi.org/10.1038/nsmb.2865.

(40) Fernandez-millan, P.; Autour, A.; Ennifar, E.; Westhof, E.; Ryckelynck, M. Crystal Structure and Fluorescence Properties of the ISpinach Aptamer in Complex with DFHBI. *RNA* **2017**, *23* (12), 1788–1795. https://doi.org/10.1261/rna.063008.117.

(41) Flinders, J.; DeFina, S. C.; Brackett, D. M.; Baugh, C.; Wilson, C.; Dieckmann, T.

Recognition of Planar and Nonplanar Ligands in the Malachite Green - RNA Aptamer Complex. *ChemBioChem* **2004**, *5* (1), 62–72. https://doi.org/10.1002/cbic.200300701.

(42) Baugh, C.; Grate, D.; Wilson, C. 2.8 Å Crystal Structure of the Malachite Green Aptamer. *J. Mol. Biol.* **2000**, *301* (1), 117–128. https://doi.org/10.1006/jmbi.2000.3951.

(43) Warner, K. D.; Sjekloa, L.; Song, W.; Filonov, G. S.; Jaffrey, S. R.; Ferré-D'Amaré, A. R. A Homodimer Interface without Base Pairs in an RNA Mimic of Red Fluorescent Protein. *Nat. Chem. Biol.* **2017**, *13* (11), 1195–1201. https://doi.org/10.1038/nchembio.2475.

(44) Sjekloća, L.; Ferré-D'Amaré, A. R. Binding between G Quadruplexes at the Homodimer Interface of the Corn RNA Aptamer Strongly Activates Thioflavin T Fluorescence. *Cell Chem. Biol.* **2019**, *26* (8), 1159–1168. https://doi.org/10.1016/j.chembiol.2019.04.012.

(45) *Fluorescence Sensors and Biosensors*, 1st Edition; Thompson, R. B., Ed.; CRC Press: Boca Raton, 2005. p 1-4.

(46) Bonel, L.; Vidal, J. C.; Duato, P.; Castillo, J. R. An Electrochemical Competitive Biosensor for Ochratoxin A Based on a DNA Biotinylated Aptamer. *Biosens.Bioelectron.* **2011**, *26* (7), 3254–3259. https://doi.org/10.1016/j.bios.2010.12.036.

(47) Gu, H.; Duan, N.; Wu, S.; Hao, L.; Xia, Y.; Ma, X.; Wang, Z. Graphene Oxide-Assisted Non-Immobilized SELEX of Okdaic Acid Aptamer and the Analytical Application of Aptasensor. *Sci.Rep.* **2016**, *6*, 1–9. https://doi.org/10.1038/srep21665.

(48) Wang, R.; Xiang, Y.; Zhou, X.; Liu, L. hua; Shi, H. A Reusable Aptamer-Based Evanescent Wave All-Fiber Biosensor for Highly Sensitive Detection of Ochratoxin A. *Biosens.Bioelectron.* **2015**, *66*, 11–18. https://doi.org/10.1016/j.bios.2014.10.079.

(49) Radi, A. E.; Sánchez, J. L. A.; Baldrich, E.; O'Sullivan, C. K. Reusable Impedimetric Aptasensor. *Anal.Chem.* **2005**, *77* (19), 6320–6323. https://doi.org/10.1021/ac0505775.

(50) McCauley, T. G.; Hamaguchi, N.; Stanton, M. Aptamer-Based Biosensor Arrays for Detection and Quantification of Biological Macromolecules. *Anal.Biochem.* **2003**, *319* (2), 244–250. https://doi.org/10.1016/S0003-2697(03)00297-5.

(51) Kubik, M. F.; Bell, C.; Fitzwater, T.; Watson, S. R.; Tasset, D. M. Isolation and Characterization of 2′-Fluoro-, 2′-Amino-, and 2′-Fluoro-/Amino-Modified RNA Ligands to Human IFN-γ That Inhibit Receptor Binding. *J. Immunol.* **1997**, *159* (1), 259–267.

(52) Kujau, M. J.; Wölfl, S. Intramolecular Derivatization of 2'-Amino-Pyrimidine Modified RNA with Functional Groups That Is Compatible with Re-Amplification. *Nucleic Acids Res.* **1998**, *26* (7), 1851–1853. https://doi.org/10.1093/nar/26.7.1851.

(53) Stojanovic, M. N.; Kolpashchikov, D. M. Modular Aptameric Sensors. *J Am Chem Soc* **2004**, *126* (30), 9266–9270. https://doi.org/10.1021/ja032013t.

(54) Xu, W.; Lu, Y. A Label-Free Fluorescent Aptamer Sensor Based on Regulation of Malachite Green Fluorescence. *Anal.Chem.* **2010**, *82* (2), 574–586. https://doi.org/10.1021/ac9018473.

(55) Kellenberger, C. A.; Hallberg, Z. F.; Hammond, M. C. *Live Cell Imaging Using Riboswitch-Spinach TRNA Fusions as Metabolite-Sensing Fluorescent Biosensors*; Springer Science+Business Media New York: New York, 2015; Vol. 1316. https://doi.org/10.1007/978-1-4939-2730-2_8.

(56) Latt, S. A.; Stetten, G. Spectral Studies on 33258 Hoechst and Related Bisbenzimidazole Dyes Useful for Fluorescent Detection Deoxyribonucleic Acid Synthesis. *J.Histochem. Cytochem.* **1976**, *24* (1), 24–33. https://doi.org/10.1177/24.1.943439.

(57) Bucevičius, J.; Lukinavičius, G.; Gerasimaite, R. The Use of Hoechst Dyes for DNA Staining and Beyond. *Chemosensors* **2018**, *6* (2), 18.

https://doi.org/10.3390/chemosensors6020018.

(58) Searle, M. S.; Embrey, K. J. Sequence-Specific Interaction of Hoescht 33258 with the Minor Groove of an Adenine-Tract DNA Duplex Studied in Solution by 1H NMR Spectroscopy. *Nucleic Acids Res.* **1990**, *18* (13), 3753–3762. https://doi.org/10.1093/nar/18.13.3753.

(59) Sandhu, L. C.; Warters, R. L.; Dethlefsen, L. A. Fluorescence Studies of Hoechst 33342 with Supercoiled and Relaxed Plasmid PBR322 DNA. *Cytometry* **1985**, *6* (3), 191–194. https://doi.org/10.1002/cyto.990060304.

(60) Portugal, J.; Waring, M. J. Assignment of DNA Binding Sites for 4′,6-Diamidine-2-Phenylindole and Bisbenzimide (Hoechst 33258). A Comparative Footprinting Study. *Biochim. Biophys. Acta* **1988**, *949* (2), 158–168. https://doi.org/10.1016/0167-4781(88)90079-6.

(61) Harshman, K. D.; Dervan, P. B. Molecular Recognition of B-DNA by Hoechst 33258. *Nucleic Acids Res.* **1985**, *13* (13), 4825–4835. https://doi.org/10.1093/nar/13.13.4825.

(62) Sriram, M.; van der Marel, G. A.; Roelen, H. L.; van Boom, J. H.; Wang, A. H. Conformation of B-DNA Containing O6-Ethyl-G-C Base Pairs Stabilized by Minor Groove Binding Drugs: Molecular Structure of d(CGC[E6G]AATTCGCG Complexed with Hoechst 33258 or Hoechst 33342. *EMBO J.* **1992**, *11* (1), 225–232. https://doi.org/10.1002/j.1460-2075.1992.tb05045.x.

(63) Berdalet, E.; Dortch, Q. New Double-Staining Technique for RNA and DNA Measurement in Marine Phytoplankton. *Mar.Ecol.Prog.Ser.* **1991**, *73* (1), 295–305. https://doi.org/10.3354/meps073295.

(64) Embrey, K. J.; Searle, M. S.; Craik, D. J. Interaction of Hoechst 33258 with the Minor Groove of the A T-Rich DNA Duplex d(GGTAATTACC)2 Studied in Solution by NMR

Spectroscopy +. *Eur. J. Biochem.* **1993**, *211* (3), 437–447. https://doi.org/10.1111/j.1432-1033.1993.tb17569.x.

(65)   Invitrogen. Hoechst Stains https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FLSG%2Fmanuals%2Fmp21486.pdf (accessed Nov 8, 2021).

(66)   Zhu, Z.; Yang, C.; Zhou, X.; Qin, J. Label-Free Aptamer-Based Sensors for L -Argininamide by Using Nucleic Acid Minor Groove Binding Dyes. *Chem. Commun.* **2011**, *47*, 3192–3194. https://doi.org/10.1039/c0cc04844e.

(67)   Sarpong, K.; Datta, B. Nucleic-Acid-Binding Chromophores as Efficient Indicators of Aptamer-Target Interactions. *J. Nucleic Acids* **2012**, *2012(4968)*, 1–7. https://doi.org/10.1155/2012/247280.

(68)   Le, H.; Jiang, X.; Zhang, M.; Ye, B. Label-Free Fluorescent Assay of ATP Based on an Aptamer-Assisted Light-up of Hoechst Dyes. *Anal. Methods* **2014**, *6*, 2028–2030. https://doi.org/10.1039/c3ay42187b.

(69)   Reuter, J. S.; Mathews, D. H. RNAstructure: Software for RNA Secondary Structure Prediction and Analysis. *BMC Bioinformatics.* **2010**, *11* (129). https://doi.org/10.1186/1471-2105-11-129.

(70)   Gellert, M.; Lipsett, M. N.; Davies, D. R. Helix Formation by Guanylic Acid. *Proc. Natl. Acad. Sci. USA* **1962**, *48*, 2013–2018. https://doi.org/10.1073/pnas.48.12.2013.

(71)   Patel, D. J.; Phan, A. T.; Kuryavyi, V. Human Telomere , Oncogenic Promoter and 5 ' -UTR G-Quadruplexes : Diverse Higher Order DNA and RNA Targets for Cancer Therapeutics. *Nucleic Acids Res* **2007**, *35* (22), 7429–7455. https://doi.org/10.1093/nar/gkm711.

(72) Phan, A. T.; Kuryavyi, V.; Patel, D. J. DNA Architecture : From G to Z. *Curr.Opin.Struct.Biol.* **2006**, *16*, 288–298. https://doi.org/10.1016/j.sbi.2006.05.011.

(73) Smith, F. W.; Feigon, J. Strand Orientation in the DNA Quadruplex Formed from the Oxytricha Telomere Repeat Oligonucleotide d(G4T4G4) in Solution. *Biochemistry* **1993**, *32* (32), 8682–8692. https://doi.org/10.1021/bi00084a040.

(74) Zhang, N.; Phan, T. A.; Patel, D. J. (3 + 1) Assembly of Three Human Telomeric Repeats into an Asymmetric Dimeric G-Quadruplex. *J.Am.Chem.Soc.* **2005**, *127* (49), 17277–17285. https://doi.org/10.1021/ja0543090.

(75) Norseen, J.; Johnson, F. B.; Lieberman, P. M. Role for G-Quadruplex RNA Binding by Epstein-Barr Virus Nuclear Antigen 1 in DNA Replication and Metaphase Chromosome Attachment. *J. Virol.* **2009**, *83* (20), 10336–10346. https://doi.org/10.1128/JVI.00747-09.

(76) Phan, A. T.; Modi, Y. S.; Patel, D. J. Propeller-Type Parallel-Stranded G-Quadruplexes in the Human c-Myc Promoter. *J.Am.Chem.Soc.* **2004**, *126* (28), 8710–8716. https://doi.org/10.1021/ja048805k.

(77) Rankin, S.; Reszka, A. P.; Huppert, J.; Zloh, M.; Parkinson, G. N.; Todd, A. K.; Ladame, S.; Balasubramanian, S.; Neidle, S. Putative DNA Quadruplex Formation within the Human C-Kit Oncogene. *J.Am.Chem.Soc.* **2005**, *127* (30), 10584–10589. https://doi.org/10.1021/ja050823u.

(78) Wanrooij, P. H.; Uhler, J. P.; Simonsson, T.; Falkenberg, M.; Gustafsson, C. M. G-Quadruplex Structures in RNA Stimulate Mitochondrial Transcription Termination and Primer Formation. *Proc. Natl. Acad. Sci. USA* **2010**, *107* (37), 16072–16077. https://doi.org/10.1073/pnas.1006026107.

(79) Bhattacharyya, D.; Arachchilage, G. M.; Basu, S. Metal Cations in G-Quadruplex Folding

and Stability. *Front. Chem.* **2016**, *4* (38), 1–14. https://doi.org/10.3389/fchem.2016.00038.

(80) Kolesnikova, S.; Curtis, E. A. Structure and Function of Multimeric G-Quadruplexes. *Molecules.* **2019**, *24* (3074), 1–20. https://doi.org/10.3390/molecules24173074.

(81) Do, N. Q.; Lim, K. W.; Teo, M. H.; Heddi, B.; Phan, A. T. Stacking of G-Quadruplexes : NMR Structure of a G-Rich Oligonucleotide with Potential Anti-HIV and Anticancer Activity. *Nucleic Acids Res.* **2011**, *39* (21), 9448–9457. https://doi.org/10.1093/nar/gkr539.

(82) Krishnan-Ghosh, Y.; Liu, D.; Balasubramanian, S. Formation of an Interlocked Quadruplex Dimer by d(GGGT). *J. Am. Chem. Soc.* **2004**, *129* (35), 11009–11016. https://doi.org/10.1021/ja049259y.

(83) Cogoi, S.; Paramasivam, M.; Spolaore, B.; Xodo, L. E. Structural Polymorphism within a Regulatory Element of the Human KRAS Promoter: Formation of G4-DNA Recognized by Nuclear Proteins. *Nucleic Acids Res.* **2008**, *36* (11), 3765–3780. https://doi.org/10.1093/nar/gkn120.

(84) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA : Sequence , Topology and Structure. *Nucleic Acids Res.* **2006**, *34* (19), 5402–5415. https://doi.org/10.1093/nar/gkl655.

(85) Do, N. Q.; Lim, K. W.; Teo, M. H.; Heddi, B.; Phan, A. T. Stacking of G-Quadruplexes : NMR Structure of a G-Rich Oligonucleotide with Potential Anti-HIV and Anticancer Activity Y. *Nucleic Acids Res.* **2011**, *39* (21), 9448–9457. https://doi.org/10.1093/nar/gkr539.

(86) Sen, D.; Gilbert, W. A Sodium-Potassium Switch in the Formation of Four-Stranded G4-DNA. *Nature* **1990**, *344* (6265), 410–414. https://doi.org/10.1038/344410a0.

(87) Williamson, J. R.; Raghuraman, M. K.; Cech, T. R. Monovalent Cation-Induced Structure

of Telomeric DNA : The G -Quartet Model. *Cell.* **1989**, *59*, 871–880. https://doi.org/10.1016/0092-8674(89)90610-7.

(88) Risitano, A.; Fox, K. R. Inosine Substitutions Demonstrate That Intramolecular DNA Quadruplexes Adopt Different Conformations in the Presence of Sodium and Potassium. *Bioorg. Med. Chem.Lett.* **2005**, *15*, 2047–2050. https://doi.org/10.1016/j.bmcl.2005.02.050.

(89) Saran, R.; Piccolo, K. A.; He, Y.; Kang, Y.; Huang, P. J. J.; Wei, C.; Chen, D.; Dieckmann, T.; Liu, J. Thioflavin T Fluorescence and Nmr Spectroscopy Suggesting a Non-g-Quadruplex Structure for a Sodium Binding Aptamer Embedded in DNAzymes. *Can. J. Chem.* **2021**, *99* (11), 860–866. https://doi.org/10.1139/cjc-2021-0024.

(90) Saintomé, C.; Amrane, S.; Mergny, J. L.; Alberti, P. The Exception That Confirms the Rule: A Higher-Order Telomeric G-Quadruplex Structure More Stable in Sodium than in Potassium. *Nucleic Acids Res.* **2016**, *44* (6), 2926–2935. https://doi.org/10.1093/nar/gkw003.

(91) Hud, N. V.; Smith, F. W.; Anet, F. A. L.; Feigon, J. The Selectivity for K+ versus Na+ in DNA Quadruplexes Is Dominated by Relative Free Energies of Hydration: A Thermodynamic Analysis by 1H NMR. *Biochemistry* **1996**, *35* (48), 15383–15390. https://doi.org/10.1021/bi9620565.

(92) Roxo, C.; Kotkowiak, W.; Pasternak, A. G-Quadruplex-Forming Aptamers—Characteristics, Applications, and Perspectives. *Molecules* **2019**, *24* (3781). https://doi.org/10.3390/molecules24203781.

(93) Amrane, S.; Adrian, M.; Heddi, B.; Serero, A.; Nicolas, A.; Mergny, J.; Tua, A. Formation of Pearl-Necklace Monomorphic G-Quadruplexes in The. *J. Am. Chem. Soc* **2012**, No. 134, 5807–5816. https://doi.org/dx.doi.org/10.1021/ja208993r | J.

(94)   Majhi, P. R.; Shafer, R. H. Characterization of an Unusual Folding Pattern in a Catalytically Active Gunaine Quadruplex Structure. *Biopolymers* **2006**, *82*, 558–569. https://doi.org/10.1002/bip.

(95)   Liu, W.; Zhu, H.; Zheng, B.; Cheng, S.; Fu, Y.; Li, W.; Lau, T. C.; Liang, H. Kinetics and Mechanism of G-Quadruplex Formation and Conformational Switch in a G-Quadruplex of PS2.M Induced by Pb2+. *Nucleic Acids Res.* **2012**, *40* (9), 4229–4236. https://doi.org/10.1093/nar/gkr1310.

(96)   Wu, Y.; Midinov, B.; White, R. J. Electrochemical Aptamer-Based Sensor for Real-Time Monitoring of Insulin. *ACS Sensors* **2019**, *4* (2), 498–503. https://doi.org/10.1021/acssensors.8b01573.

(97)   Macaya, R. F.; Schultze, P.; Smith, F. W.; Roe, J. A.; Feigon, J. Thrombin-Binding DNA Aptamer Forms a Unimolecular Quadruplex Structure in Solution. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3745–3749. https://doi.org/10.1073/pnas.90.8.3745.

(98)   Dailey, M. M.; Clarke Miller, M.; Bates, P. J.; Lane, A. N.; Trent, J. O. Resolution and Characterization of the Structural Polymorphism of a Single Quadruplex-Forming Sequence. *Nucleic Acids Res.* **2010**, *38* (14), 4877–4888. https://doi.org/10.1093/nar/gkq166.

(99)   Paborsky, L. R.; McCurdy, S. N.; Griffin, L. C.; Toole, J. J.; Leung, L. L. K. The Single-Stranded DNA Aptamer-Binding Site of Human Thrombin. *J. Biol. Chem.* **1993**, *268* (28), 20808–20811. https://doi.org/10.1016/s0021-9258(19)36856-5.

(100)   Pica, A.; Russo Krauss, I.; Merlino, A.; Nagatoishi, S.; Sugimoto, N.; Sica, F. Dissecting the Contribution of Thrombin Exosite i in the Recognition of Thrombin Binding Aptamer. *FEBS J.* **2013**, *280* (24), 6581–6588. https://doi.org/10.1111/febs.12561.

(101) Benzaghou, I.; Bougie, I.; Picard-Jean, F.; Bisaillon, M. Energetics of RNA Binding by the West Nile Virus RNA Triphosphatase. *FEBS Lett.* **2006**, *580*, 867–877. https://doi.org/10.1016/j.febslet.2006.01.006.

(102) Eddy, J.; Maizels, N. Gene Function Correlates with Potential for G4 DNA Formation in the Human Genome. *Nucleic Acids Res.* **2006**, *34* (14), 3887–3896. https://doi.org/10.1093/nar/gkl529.

(103) Garant, J.; Perreault, J.; Scott, M. S. Motif Independent Identification of Potential RNA G-Quadruplexes by G4RNA Screener. *Bioinformatics* **2017**, *33* (22), 3532–3537. https://doi.org/10.1093/bioinformatics/btx498.

(104) Kikin, O.; Antonio, L. D.; Bagga, P. S. QGRS Mapper: A Web-Based Server for Predicting G-Quadruplexes in Nucleotide Sequences. *Nucleic Acids Res.* **2006**, *34*, 676–682. https://doi.org/10.1093/nar/gkl253.

(105) Varizhuk, A.; Ischenko, D.; Tsvetkov, V.; Novikov, R.; Kulemin, N.; Kaluzhny, D.; Vlasenok, M.; Naumov, V.; Smirnov, I.; Pozmogova, G. The Expanding Repertoire of G4 DNA Structures. *Biochimie* **2017**, *135*, 54–62. https://doi.org/10.1016/j.biochi.2017.01.003.

(106) Brunger, A. T.; Adams, P. D.; Clore, G. M.; Delano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Cryst.* **1998**, *54*, 905–921. https://doi.org/10.1107/S0907444998003254.

(107) Brunger, A. T. Version 1 . 2 of the Crystallography and NMR System. *Nat. Protoc.* **2007**, *2* (11), 2728–2733. https://doi.org/10.1038/nprot.2007.406.

(108) Wang, Y.; Patel, D. J. Solution Structure of the Tetrahymena Telomeric Repeat d(T2G4)4 G-Tetraplex. *Structure* **1994**, *2* (12), 1141–1156. https://doi.org/10.1016/s0969-2126(94)00117-0.

(109) Jiang, F.; Patel, D. J.; Zhang, X.; Zhao, H.; Jones, R. A. Specific Labeling Approaches to Guanine and Adenine Imino and Amino Proton Assignments in the AMP – RNA Aptamer Complex. *J. Biomol. NMR* **1997**, *9*, 55–62. https://doi.org/10.1023/a:1018623601946.

(110) Kolesnikova, S.; Hubálek, M.; Bednárová, L.; Cvačka, J.; Curtis, E. A. Multimerization Rules for G-Quadruplexes. *Nucleic Acids Res.* **2017**, *45* (15), 8684–8696. https://doi.org/10.1093/nar/gkx637.

(111) Mukundan, V. T.; Do, N. Q.; Phan, A. T. HIV-1 Integrase Inhibitor T30177 Forms a Stacked Dimeric G-Quadruplex Structure Containing Bulges. *Nucleic Acids Res.* **2011**, *39* (20), 8984–8991. https://doi.org/10.1093/nar/gkr540.

(112) Fried, M. G. Measurement of Protein-DNA Interaction Parameters by Electrophoresis Mobility Shift Assay. *Electrophoresis* **1989**, *10* (5–6), 366–376. https://doi.org/10.1002/elps.1150100515.

(113) Wassarman, K. M. *Native Gel Electrophoresis to Study the Binding and Release of RNA Polymerase by 6S RNA. In Keiler K. (Eds) Bacterial Regulatory RNA. Methods in Molecular Biology (Methods and Protocols)*, Humana Pre.; Totowa, NJ, 2012; Vol. 905. p 259-272

(114) Lane, D.; Prentki, P.; Chandler, M. Use of Gel Retardation to Analyze Protein-Nucleic Acid Interactions. *Microbiol. Rev.* **1992**, *56* (4), 509–528. https://doi.org/10.1128/mmbr.56.4.509-528.1992.

(115) Kypr, J.; Kejnovská, I.; Renčiuk, D.; Vorlíčková, M. Circular Dichroism and

Conformational Polymorphism of DNA. *Nucleic Acids Res.* **2009**, *37* (6), 1713–1725. https://doi.org/10.1093/nar/gkp026.

(116) Engman, K. C.; Sandin, P.; Osborne, S.; Brown, T.; Billeter, M.; Lincoln, P.; Nordén, B.; Albinsson, B.; Wilhelmsson, L. M. DNA Adopts Normal B-Form upon Incorporation of Highly Flourescent DNA Base Analogue TC: NMR Structure and UV-Vis Spectroscopy Characterization. *Nucleic Acids Res.* **2004**, *32* (17), 5087–5095. https://doi.org/10.1093/nar/gkh844.

(117) Paramasivan, S.; Rujan, I.; Bolton, P. H. Circular Dichroism of Quadruplex DNAs: Applications to Structure, Cation Effects and Ligand Binding. *Methods* **2007**, *43* (4), 324–331. https://doi.org/10.1016/j.ymeth.2007.02.009.

(118) del Villar-Guerra, R.; Trent, J. O.; Chaires, J. B. G-Quadruplex Secondary Structure from Circular Dichroism Spectroscopy. *Angew. Chem. Int. Ed. Engl.* **2018**, *57* (24), 7171–7175. https://doi.org/10.1002/anie.201709184.G-quadruplex.

(119) Keeler, J. *Understanding NMR Spectroscopy*, Second.; John Wiley and Sons, Inc: Chichester, U.K., 2010.

(120) Dieckmann, T.; Piazza, M.; Bonneau, E.; Legault, P. Biomolecular NMR Spectroscopy of Ribonucleic Acids. *eLS* **2014**, 1–11. https://doi.org/10.1002/9780470015902.a0021033.pub2.

(121) Flinders, J.; Dieckmann, T. NMR Spectroscopy of Ribonucleic Acids. *Prog.Nucl. Magn. Reson. Spectrosc.* **2006**, *48* (2–3), 137–159. https://doi.org/10.1016/j.pnmrs.2006.03.001.

(122) Dai, J.; Punchihewa, C.; Ambrus, A.; Chen, D.; Jones, R. A.; Yang, D. Structure of the Intramolecular Human Telomeric G-Quadruplex in Potassium Solution : A Novel Adenine Triple Formation. *Nucleic Acids Res.* **2007**, *35* (7), 2440–2450.

https://doi.org/10.1093/nar/gkm009.

(123) Sambrook, J.; Russell, D. W. *Molecular Cloning. A Labortory Manual*, Third Edit.; Cold Spring Harbor Laboratory Press: New York, 2001. Chapter 5.26-5.28, 5.40-5.46, 5.51-5.54.

(124) Li, Y.; Chen, S.; Liu, N.; Ma, L.; Wang, T.; Veedu, R. N.; Li, T.; Zhang, F.; Zhou, H.; Cheng, X.; Jing, X. A Systematic Investigation of Key Factors of Nucleic Acid Precipitation toward Optimized DNA/RNA Isolation. *Biotechniques* **2020**, *68* (4), 191–199. https://doi.org/10.2144/BTN-2019-0109.

(125) Razzell, W. E. The Precipitation of Polyribonucleotides With Magnesium Salts and Ethanol. *J. Biol. Chem.* **1963**, *238* (9), 3053–3057. https://doi.org/10.1016/S0021-9258(18)51864-0.

(126) Thermo Fisher. User Guide Invitrogen SYBR ™ Safe DNA Gel Stain https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2Fsybr_safe_dna_gel_stain_man.pdf (accessed Oct 7, 2021).

(127) Savitzky, A.; Golay, M. J. E. Smoothing and Differentiation. *Anal. Chem* **1964**, *36* (8), 1627–1639.

(128) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; Walt, S. J. van der; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimr, R.; Mulbregt, P. van. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature* **2020**, *17* (3), 261–272.

(129) Graphpad Software. Graphpad Prism v6.00 for Windows. La Jolla: California USA 2020.

(130) Poon, K.; MacGregor, R. B. Formation and Structural Determinants of Multi-Stranded Guanine-Rich DNA Complexes. *Biophys. Chem.* **2000**, *84* (3), 205–216. https://doi.org/10.1016/S0301-4622(00)00106-X.

(131) Friebolin, H. *Basic One- and Two-Dimensional NMR Spectroscopy*, 5th Editio.; WILEY-VCH: Weinheim, 2011. p 1-21, 230-232, 281-282, 297-311.

(132) Doucleff, M.; Hatcher-Skeers, M.; Crane, N. *Pocket Guide to Biomolecular NMR*; Springer: Berlin, 2011. p 8-13.

(133) Varani, G.; Aboul-ela, F.; Allain, F. H. NMR Investigation of RNA Structure. *Prog.Nucl. Magn. Reson. Spectrosc.* **1996**, No. 29, 51–127. https://doi.org/10.1016/0079-6565(96)01028-x.

(134) Furtig, B.; Richter, C.; Wohnert, J.; Schwalbe, H. NMR Spectroscopy of RNA. *ChemBioChem* **2003**, *4*, 936–962. https://doi.org/10.1002/cbic.200300700.

(135) Kotar, A.; Foley, H. N.; Baughman, K. M.; Keane, S. C. Advanced Approaches for Elucidating Structures of Large RNAs Using NMR Spectroscopy and Complementary Methods. *Methods* **2020**, *183*, 93–107. https://doi.org/10.1016/j.ymeth.2020.01.009.

(136) Dieckmann, T.; Feigon, J. Assignment Methodology for Larger RNA Oligonucleotides: Application to an ATP-Binding RNA Aptamer. *J. Biomol. NMR* **1997**, *9* (3), 259–272. https://doi.org/10.1023/A:1018622708674.

(137) Nikonowicz, E. P.; Sirr, A.; Legault, P.; Jucker, F. M.; Baer, L. M.; Pardi, A. Preparation of 13C and 15N Labelled RNAs for Heteronuclear Multi-Dimensional NMR Studies. *Nucleic Acids Res.* **1992**, *20* (17), 4507–4513. https://doi.org/10.1093/nar/20.17.4507.

(138) Evans, J. N. S. *Biomolecular NMR Spectroscopy*; Oxford University Press: Oxford, 1995. p 21-22,74,75, 343-387, 432-434.

(139) Katahira, M.; Mashima, T. Nucleic Acid NMR – Introduction. In: Roberts G.C.K. (Eds) Encyclopedia of Biophysics. *Encycl. Biophy.*; Springer, Berlin, Heidelberg, 2013; pp 1763–1772. https://doi.org/10.1007/978-3-642-16712-6_302.

(140) Patel, D. J.; Kozlowski, S. A.; Nordheim, A.; Rich, A. Right-Handed and Left-Handed DNA : Studies of B- and Z-DNA by Using Proton Nuclear Overhauser Effect and P NMR. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 1413–1417.

(141) Lim, K. W.; Lacroix, L.; Yue, D. J. E.; Lim, J. K. C.; Lim, J. M. W.; Phan, A. T. Coexistence of Two Distinct G-Quadruplex Conformations in the HTERT Promoter. *J.Am.Chem.Soc.* **2010**, *132*, 12331–12342. https://doi.org/10.1021/ja101252n.

(142) Collie, G. W.; Promontorio, R.; Hampel, S. M.; Micco, M.; Neidle, S.; Parkinson, G. N. Structural Basis for Telomeric G-Quadruplex Targeting by Naphthalene Diimide Ligands. *J. Am. Chem. Soc.* **2012**, *134* (5), 2723–2731. https://doi.org/10.1021/ja2102423.

(143) Mao, Bi.; Gu, Z.; Gorin, A.; Chen, J.; Hingerty, B.; Amin, S.; Broyde, S.; Geacintov, N. E.; Patel, D. J. Solution Conformation of the (+)-Trans-Antibenzo[g]Chrysene-DA Adduct Opposite DT in a DNA Duplex. *J. Mol. Biol.* **1999**, *38*, 10831–10842. https://doi.org/10.1021/bi991212f.

(144) Hermann, T.; Patel, D. J. Adaptive Recognition by Nucleic Acid Aptamers. *Science (80-.).* **2000**, *287* (5454), 820–825. https://doi.org/10.1126/science.287.5454.820.

(145) Yue, D. J. E.; Lim, K. W.; Phan, A. T. Formation of (3+1) G-Quadruplexes with a Long Loop by Human Telomeric DNA Spanning Five or More Repeats. *J.Am.Chem.Soc.* **2011**, *133*, 11462–11465. https://doi.org/10.1021/ja204197d.

(146) Hu, L.; Lim, K. W.; Bouaziz, S.; Tua, A. Giardia Telomeric Sequence d(TAGGG)4 Forms Two Intramolecular G-Quadruplexes in K + Solution : Effect of Loop Length and Sequence

on the Folding Topology. *J.Am.Chem.Soc.* **2009**, No. 131, 16824–16831. https://doi.org/10.1021/ja905611c. p 477-480.

(147) Webba da Silva, M. NMR Methods for Studying Quadruplex Nucleic Acids. *Methods* **2007**, *43*, 264–277. https://doi.org/10.1016/j.ymeth.2007.05.007.

(148) Hoult, D. I. Solvent Peak Saturation with Single Phase and Quadrature Fourier Transformation. *J.Magn.Reson.* **1976**, *21* (2), 337–347. https://doi.org/10.1016/0022-2364(76)90081-0.

(149) Briand, J.; Ernst, R. R. Computer-Optimized Homonuclear TOCSY Experiments with Suppression of Cross Relaxation. *Chem.Phys.Lett.* **1991**, *185* (3.4), 276–285. https://doi.org/10.1016/S0009-2614(91)85060-A.

(150) Braun, S.; Kalinowski, H.-O.; Berger, S. *150 and More Basic NMR Experiments*; Wiley-VCH: Weinheim, 1998.

(151) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. Structure Determination of a Tetrasaccharide: Transiet Nuclear Overhauser Effects in the Rotating Frame. *J. Am. Chem. Soc* **1984**, *106*, 811–813.

(152) Sklenář, V.; Bax, A. Spin-Echo Water Suppression for the Generation of Pure-Phase Two-Dimensional NMR Spectra. *J.Magn.Reson.* **1987**, *74* (3), 469–479. https://doi.org/10.1016/0022-2364(87)90269-1.

(153) Kumar, A.; Ernst, R. R.; Wüthrich, K. A Two-Dimensional Nuclear Overhauser Enhancement (2D NOE) Experiment for the Elucidation of Complete Proton-Proton Cross-Relaxation Networks in Biological Macromolecules. *Biochem. Biophys.Res.Commun.* **1980**, *95* (1), 1–6. https://doi.org/10.1016/0006-291X(80)90695-6.

(154) Nakatani, K.; Hagihara, S.; Goto, Y.; Kobori, A.; Hagihara, M.; Hayashi, G.; Kyo, M.;

Nomura, M.; Mishima, M.; Kojima, C. Small-Molecule Ligand Induces Nucleotide Flipping in (Cag)n Trinucleotide Repeats. *Nat. Chem. Biol.* **2005**, *1* (1), 39–43. https://doi.org/10.1038/nchembio708.

(155) Phan, A. T.; Modi, Y. S.; Patel, D. J. Two-Repeat Tetrahymena Telomeric d(TGGGGTTGGGGT) Sequence Interconverts between Asymmetric Dimeric G-Quadruplexes in Solution. *J. Mol. Biol.* **2004**, *338* (1), 93–102. https://doi.org/10.1016/j.jmb.2004.02.042.

(156) Kuryavyi, V.; Majumdar, A.; Shallop, A.; Chernichenko, N.; Skripkin, E.; Jones, R.; Patel, D. J. A Double Chain Reversal Loop and Two Diagonal Loops Define the Architecture of a Unimolecular DNA Quadruplex Containing a Pair of Stacked G (Syn) Á G (Syn) Á G (Anti) Á G (Anti) Tetrads Flanked by a G Á (T-T) Triad and a T Á T Á T Triple. *J. Mol. Biol* **2001**, *310*, 181–194. https://doi.org/10.1006/jmbi.2001.4759.

(157) Tahiri-Alaoui, A.; Frigotto, L.; Manville, N.; Ibrahim, J.; Romby, P.; Jamesa, W. High Affinity Nucleic Acid Aptamers for Streptavidin Incorporated into Bi-Specific Capture Ligands. *Nucleic Acids Res* **2002**, *30* (10). https://doi.org/10.1093/nar/30.10.e45.

(158) Jing, M.; Bowser, M. T. A Review of Methods for Measuring Aptamer-ProteinEquilibria. *Anal Chim Acta* **2011**, *686* (1), 9–18. https://doi.org/10.1016/j.aca.2010.10.032.

(159) Karsisiotis, A. I.; Webba da Silva, M. Structural Probes in Quadruplex Nucleic Acid Structure Determination by NMR. *molecules* **2012**, *17*, 13073–13086. https://doi.org/10.3390/molecules171113073.

Appendix A: Supplementary Figures
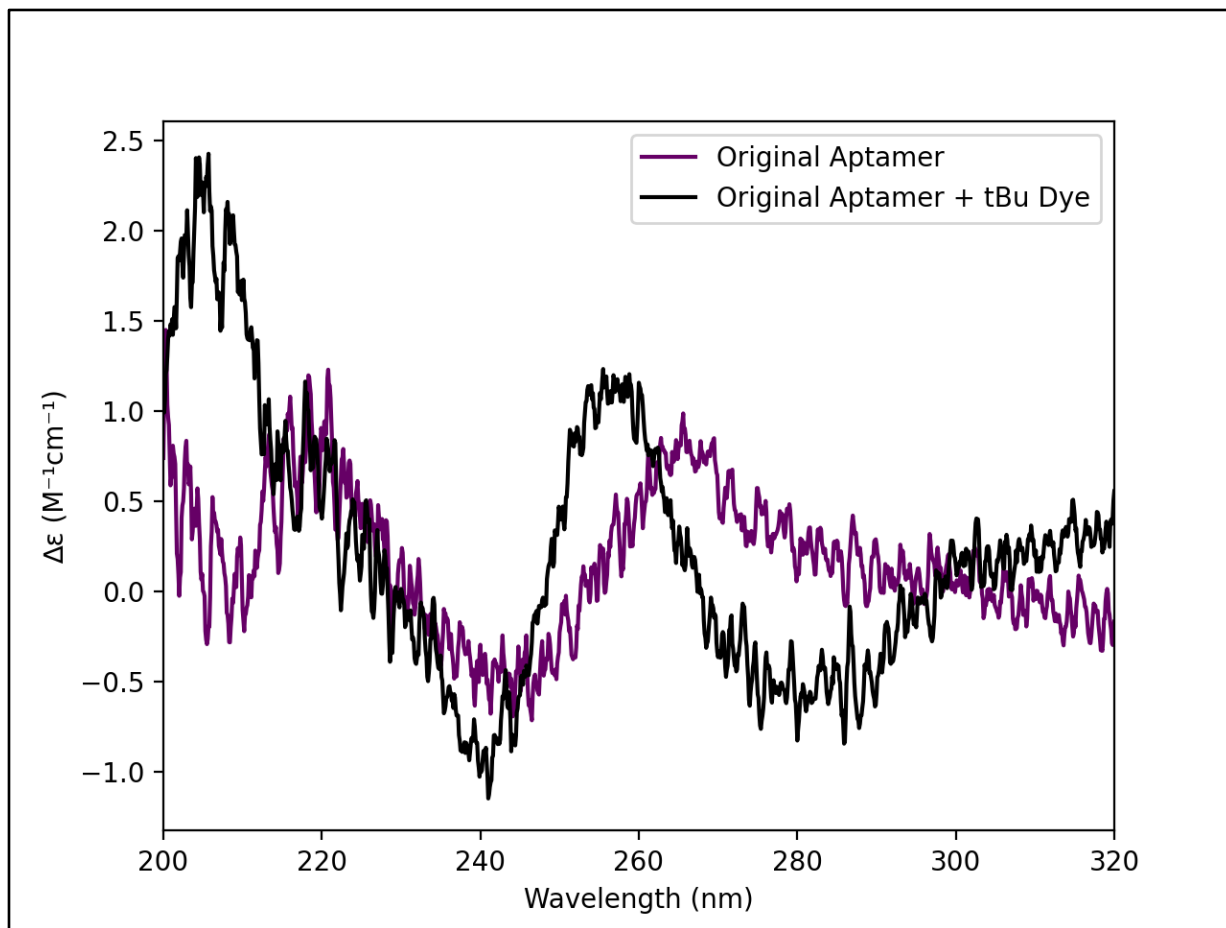
Figure A.1 CD Scan of original DNA aptamer and t*Bu* Hoechst dye without Savizky-Golay Smoothing. Samples contain 40μM of DNA and were applicable 50μM t*Bu* Hoechst dye. The scan was taken from 320nm to 200nm, with a data interval of 0.5nm, response time of 1 second, band width of 0.5nm, and a scanning speed of 30nm minute[-1]. Four scans were accumulated. The baselines of either buffer or buffer and t*Bu* Hoechst dye were then subtracted from their corresponding data.
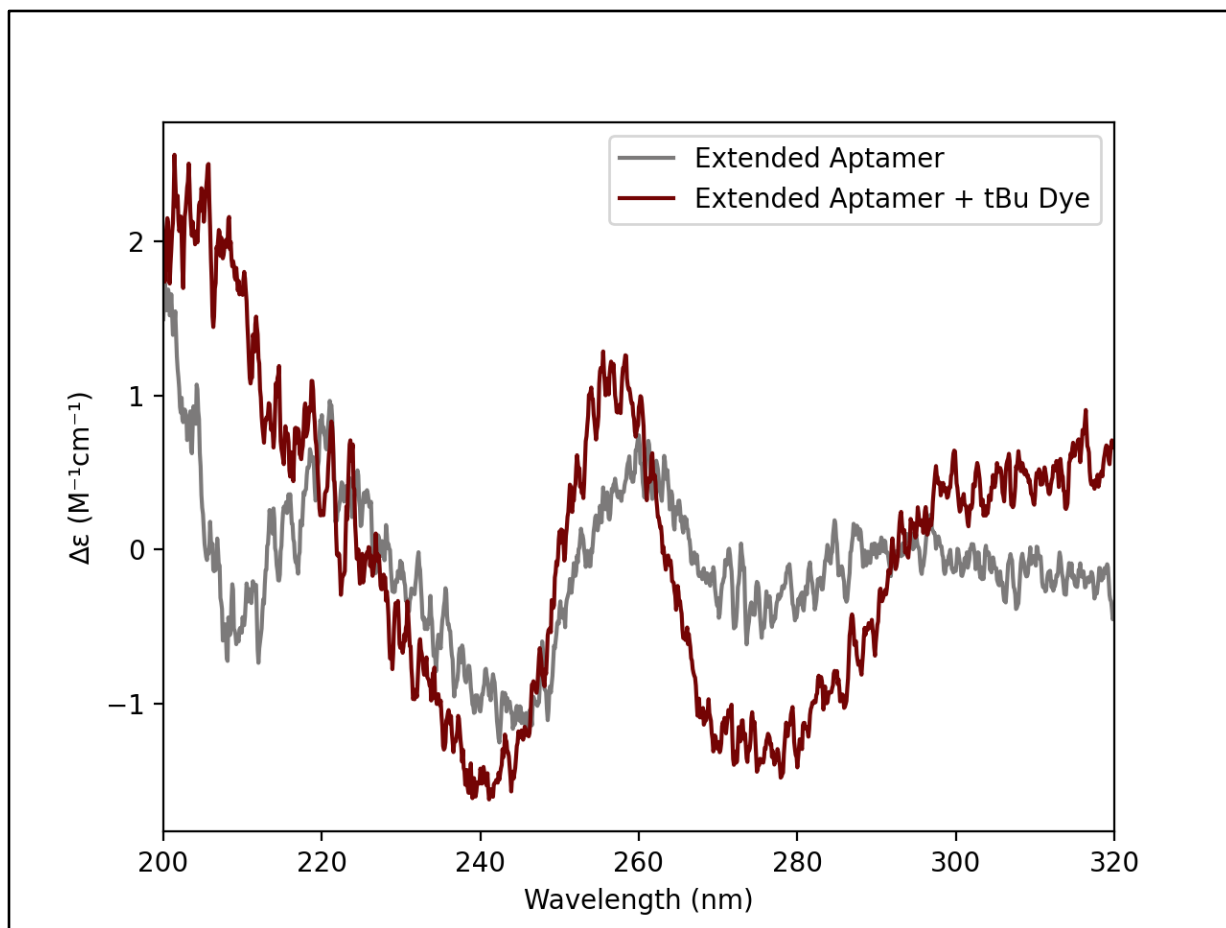
Figure A.2 CD Scan of extended stem aptamer and t*Bu* Hoechst dye without Savizky-Golay Smoothing. Samples contain 40μM of DNA and were applicable 50μM t*Bu* Hoechst dye. The scan was taken from 320nm to 200nm, with a data interval of 0.5nm, response time of 1 second, band width of 0.5nm, and a scanning speed of 30nm minute[-1]. Four scans were accumulated. The baselines of either buffer or buffer and t*Bu* Hoechst dye were then subtracted from their corresponding data.
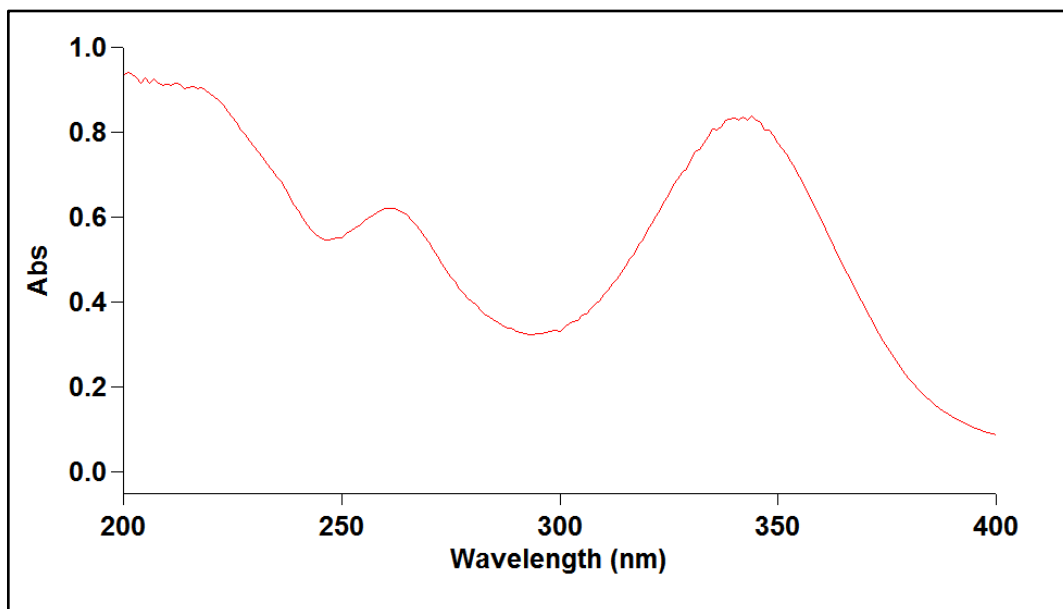
Figure A.3 Absorbance spectrum for t*Bu* Hoechst dye. The maximum of the peaks are at 344 nm, 261 nm, and around 205 nm.

Appendix B: NMR Pulse Programs

1D 1H with 11-Spin Echo

Solvent Suppression

```
;zggpjrse
;avance-version (13/08/01)
;1D sequence
;
;V. Sklenar & A. Bax, J. Magn. Reson.
74, 469 (1987)

#include <Avance.incl>
#include <Grad.incl>

1 ze
2 30m
 d1
 50u UNBLKGRAD
 p1 ph1
 d19
 p1 ph2
 4u
 p16:gp1
 d16
 p1 ph3
 d19*2
 p1 ph4
 p16:gp1
 d16
 4u BLKGRAD
 go=2 ph31
 30m mc #0 to 2 F0(zd)
exit

ph1=0 2
ph2=2 0
ph3=0 0 1 1 2 2 3 3
ph4=2 2 3 3 0 0 1 1
ph31=0 2 2 0

;pl1 : f1 channel - power level for pulse
(default)
```

```
;p1 : f1 channel -  90 degree high power
pulse
;p16: homospoil/gradient pulse
;d1 : relaxation delay; 1-5 * T1
;d16: delay  for  homospoil/gradient
recovery
;d19:  delay  for  binomial  water
suppression
;   d19 = (1/(2*d)), d = distance of next
null (in Hz)
;ns: 8 * n, total number of scans: NS *
TD0
;ds: 4

;for z-only gradients:
;gpz1: 20%

;use gradient files:
;gpnam1: SMSQ10.100

;$Id:   zggpjrse,v   1.1   2013/08/30
09:42:08 ber Exp $
```

1D 1H with Presaturation

```
;SD_1Dpresat
;avance-version
;1D sequence with f1 presaturation

#include <Avance.incl>

"d12=20u"
"d13=3u"

1 ze
2 d12 pl9:f1
 p18*0.6 ph28
 d13
 p18*0.4 ph29
 d12 pl1:f1
 p1 ph1
 go=2 ph31
 wr #0
exit
```

ph1=0 2 2 0 1 3 3 1
ph28= 0
ph29= 1
ph31=0 2 2 0 1 3 3 1

;pl1 : f1 channel - power level for pulse (default)
;pl9 : f1 channel - power level for presaturation
;p1 : f1 channel - 90 degree high power pulse
;d1 : relaxation delay; 1-5 * T1
;d12: delay for power switching [20 usec]
;d13: short delay   [3 usec]

COSY

;cosydfph
;avance-version (12/01/11)
;2D homonuclear shift correlation
;phase sensitive
;with double quantum filter
;
;phasecycle: A. Derome & M. Williamson, J. Magn. Reson. 88,
;   177 - 185 (1990)
;
;$CLASS=HighRes
;$DIM=2D
;$TYPE=
;$SUBTYPE=
;$COMMENT=

#include <Avance.incl>

"d13=4u"

"in0=inf1"

"d0=in0/2-p1*4/3.1416"

1 ze
2 d1
3 p1 ph1

  d0
  p1 ph2
  d13
  p1 ph3
  go=2 ph31
  d1 mc #0 to 2 F1PH(calph(ph1, +90), caldel(d0, +in0))
exit

ph1=1 1 1 1 0 0 0 0
ph2=0 0 0 0 1 1 1 1
ph3=1 2 3 0 2 3 0 1
ph31=0 3 2 1 3 2 1 0

;pl1 : f1 channel - power level for pulse (default)
;p1 : f1 channel - 90 degree high power pulse
;d0 : incremented delay (2D)
;d1 : relaxation delay; 1-5 * T1
;d13: short delay   [4 usec]
;inf1: 1/SW = 2 * DW
;in0: 1/(1 * SW) = 2 * DW
;nd0: 1
;ns: 8 * n
;ds: 16
;td1: number of experiments
;FnMODE: States-TPPI, TPPI, States or QSEQ

;Processing

;PHC0(F1): 90
;PHC1(F1): -180
;FCOR(F1): 1

;$Id: cosydfph,v 1.9 2012/01/31 17:49:22 ber Exp $

2D NOESY with 11-Spin Echo

Solvent Suppression

```
;SD_noe11ezg
;2D NOE with 11echo and gradients
;phase sensitive using States-TPPI
;with presaturation during relaxation
delay and mixing time
;the shortest mixing time ~50 ms!!!!
otherwise problems with radiation
;damping
;added decoupling, 9/15/94; thorsten
;changed for DRX, 05/01/96; thorsten
;modified for Topspin 3, Feb 15 2019,
Sameer / UoGuelph

;Allows  you  to  center  spectrum
anywhere.
;Uses cnst47/48
; 1) set o2 to the desired center of
spectrum in f1
; 2) set cnst47 to the difference in Hz
between o2 and water (the second value to
water
; 3) set cnst48 to the difference in Hz
between o2 and the desired center of
spectrum in f2
; e.g. for a spectrum centered at 7.5
ppm in F1/F2 @ 600 MHz, o2 = 4500, cnst47
~ -1680, cnst48 = 0
;changed 1/21/97 - ELW

;d2 calculation added 3-22-97, td

#include <Avance.incl>
#include <Grad.incl>

"d11=30m"
"d16=200u"
"l3=td1/2"
"d0=(in0/2)-(p1*4/3.14159)"
"d8=d9-p16-d16"
"d2=1s/(cnst2*4)-(p1*0.667)"
```

```
1  ze
   d11
2  d11
3  d11 pl1:f1
4  d11
   LOCKH_OFF
   d1
   LOCKH_ON
   p1 ph1
   d0
   p1 ph2
   d8
    p16:gp1
    d16 fq=cnst47:f1
   ;Water
   p1 ph3
   d2
   p1 ph6
    p16:gp2
    d16
   p1 ph7
   d2
   d2
   p1 ph8
    p16:gp2
    d16 fq=cnst48:f1
   ;Center
   go=2 ph31
   d11 wr #0 if #0 ip1 zd
    lo to 3 times 2
    d11 id0
     lo to 4 times l3
   LOCKH_OFF
exit

ph1= 0 2 0 2  1 3 1 3  2 0 2 0  3 1 3 1
     2 0 2 0  3 1 3 1  0 2 0 2  1 3 1 3
ph2= 0 0 2 2  1 1 3 3  2 2 0 0  3 3 1 1
     2 2 0 0  3 3 1 1  0 0 2 2  1 1 3 3
ph3= 0 0 0 0  1 1 1 1  2 2 2 2  3 3 3 3
ph6= 2 2 2 2  3 3 3 3  0 0 0 0  1 1 1 1
ph7= 0 0 0 0  1 1 1 1  2 2 2 2  3 3 3 3
ph8= 2 2 2 2  3 3 3 3  0 0 0 0  1 1 1 1
ph20=0 0 0 0  0 0 0 0  2 2 2 2  2 2 2 2
ph31=0 2 2 0  1 3 3 1  2 0 0 2  3 1 1 3
```

;pl1: Proton power level
;p1 : 90 degree proton pulse
;d9 : the mixing time
;d0 : incremented delay (2D) = in0/2-(4/pi*p1)
;d2 : delay for exitation maximum, calculated based
; on cnst2 = offset of first max in Hz
;d11: delay for disk I/O [30 msec]
;gp1: Homospoil during mixing time
;gp2: Watergate gradient
;in0: 1/(SW)
;nd0: 1
;NS: 2,4 8,16 or 32 * n
;DS: 4
;td1: number of experiments
;MC2: States-TPPI

2D NOESY With Presaturation

;SD_noesypr
;avance-version
;2D homonuclear correlation via dipolar coupling
;dipolar coupling may be due to noe or chemical exchange.
;phase sensitive using States-TPPI method
;with presaturation during relaxation delay

; changed presat to 'vladi-trick' 03-24-96, td

#include <Avance.incl>
#include <Grad.incl>

"d11=30m"
"d12=20u"
"d13=5u"
"d16=200u"

"l3=(td1/2)"
"d0=((in0/2)-(p1*4/3.14159))"
"p19=d9-d12-p16-d16"

1 ze
2 d11
3 d11
4 d12 pl9:f1
 LOCKH_OFF
 p18*0.8 ph29
 d13
 p18*0.2 ph30
 LOCKH_ON
 d12 pl1:f1
 p1 ph1
 d0
 p1 ph2
 d12 pl9:f1
 p19 ph18
 p16:gp1
 d16 pl1:f1
 p1 ph3
 go=2 ph31
 d11 wr #0 if #0 ip1 zd
 lo to 3 times 2
 d11 id0
 lo to 4 times l3
 LOCKH_OFF
exit

ph1 =0 2
ph2 =0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph3 =0 0 2 2 1 1 3 3
ph18=0
ph29=0
ph30=1
ph31=0 2 2 0 1 3 3 1 2 0 0 2 3 1 1 3

;pl1 : f1 channel - power level for pulse (default)
;pl9 : f1 channel - power level for presaturation
;p1 : f1 channel - 90 degree high power pulse
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d9 : mixing time

118

;d11: delay for disk I/O [30 msec]

;d12: delay for power switching [20 usec]

;d13: short delay [3 usec]

;L3: loop for phase sensitive 2D using States-TPPI method: l3 = td1/2

;in0: 1/(1 * SW) = 2 * DW

;nd0: 1

;NS: 8 * n

;DS: 16

;td1: number of experiments

;MC2: States-TPPI

2D Dipsi-TOCSY

;dipsi2esgpph

;avance-version (19/02/08)

;homonuclear Hartman-Hahn transfer using DIPSI2 sequence

; for mixing

;phase sensitive

;water suppression using excitation sculpting with gradients

;(use parameterset DIPSI2ESGPPH)

;

;A.J. Shaka, C.J. Lee & A. Pines, J. Magn. Reson. 77, 274 (1988)

;T.-L. Hwang & A.J. Shaka, J. Magn. Reson.,

; Series A 112 275-279 (1995)

;

;$CLASS=HighRes

;$DIM=2D

;$TYPE=

;$SUBTYPE=

;$COMMENT=

prosol relations=<triple>

#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"p2=p1*2"

"d11=30m"
"d12=20u"
"d13=4u"

"in0=inf1"

"d0=in0*0.5-p1*4/3.1416"

"TAU=de+p1*2/3.1416+4u"

"FACTOR1=(d9/(p6*115.112))/2"
"l1=FACTOR1*2"

"acqt0=0"
baseopt_echo

1 ze
2 d11
3 d12 pl32:f1
  d1 cw:f1 ph29
  d13 do:f1
  d12 pl1:f1
  p1 ph1
  d0
  p1 ph2

  50u UNBLKGRAD
  p16:gp1
  d16 pl10:f1

  ;begin DIPSI2
4 p6*3.556 ph23
  p6*4.556 ph25
  p6*3.222 ph23
  p6*3.167 ph25
  p6*0.333 ph23
  p6*2.722 ph25
  p6*4.167 ph23
  p6*2.944 ph25
  p6*4.111 ph23

  p6*3.556 ph25
  p6*4.556 ph23
  p6*3.222 ph25
  p6*3.167 ph23

119

```
p6*0.333 ph25                          TAU
p6*2.722 ph23                          p16:gp4
p6*4.167 ph25                          d16 pl0:f1
p6*2.944 ph23                          (p12:sp1 ph6:r):f1
p6*4.111 ph25                          4u
                                       d12 pl1:f1
p6*3.556 ph25
p6*4.556 ph23                          p2 ph7
p6*3.222 ph25
p6*3.167 ph23                          4u
p6*0.333 ph25                          p16:gp4
p6*2.722 ph23                          d16
p6*4.167 ph25                          4u BLKGRAD
p6*2.944 ph23
p6*4.111 ph25                          go=2 ph31
                                       d11 mc #0 to 2 F1PH(calph(ph1, +90)
p6*3.556 ph23                   & calph(ph29, +90), caldel(d0, +in0))
p6*4.556 ph25                          exit
p6*3.222 ph23
p6*3.167 ph25                          ph1=0 2
p6*0.333 ph23                          ph2=0 0 0 0 2 2 2 2
p6*2.722 ph25                          ph3=0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
p6*4.167 ph23                          ph4=0 0 1 1
p6*2.944 ph25                          ph5=2 2 3 3
p6*4.111 ph23                          ph6=0 0 0 0 1 1 1 1
lo to 4 times l1                       ph7=2 2 2 2 3 3 3 3
                                       ph23=3
;end DIPSI2                             ph25=1
                                       ph29=0
4u                                     ph31=0 2 2 0 0 2 2 0 2 0 0 2 2 0 0 2
p16:gp2
d16 pl1:f1                             ;pl0 : 0W
                                       ;pl1 : f1 channel - power level for pulse
p1 ph3                          (default)
                                       ;pl10: f1 channel - power level for
p16:gp3                         TOCSY-spinlock
d16 pl0:f1                             ;pl32: f1 channel - power level for low
(p12:sp1 ph4:r):f1              power presaturation
4u                                     ;sp1 : f1 channel - shaped pulse 180
d12 pl1:f1                      degree
                                       ;p1 : f1 channel - 90 degree high power
p2 ph5                          pulse
                                       ;p2 : f1 channel - 180 degree high
4u                              power pulse
p16:gp3                                ;p6 : f1 channel - 90 degree low power
d16                             pulse
```

;p12: f1 channel - 180 degree shaped pulse (Squa100.1000)   [2 msec]
;p16: homospoil/gradient pulse
;d0 : incremented delay (2D)
;d1 : relaxation delay; 1-5 * T1
;d9 : TOCSY mixing time
;d11:   delay   for   disk   I/O [30 msec]
;d12:   delay   for   power   switching [20 usec]
;d13: short delay   [4 usec]
;d16:   delay   for   homospoil/gradient recovery
;l1:   loop   for   DIPSI   cycle: ((p6*115.112) * l1) = mixing time
;inf1: 1/SW = 2 * DW
;in0: 1/(1 * SW) = 2 * DW
;nd0: 1
;ns: 8 * n
;ds: 16
;td1: number of experiments
;FnMODE: States-TPPI, TPPI, States or QSEQ

;use gradient ratio: gp 1 : gp 2 : gp 3 : gp 4
;   1 :   3 :   31 :   11

;for z-only gradients:
;gpz1: 1%
;gpz2: 3%
;gpz3: 31%
;gpz4: 11%

;use gradient files:
;gpnam1: SMSQ10.100
;gpnam2: SMSQ10.100
;gpnam3: SMSQ10.100
;gpnam4: SMSQ10.100

;set pl32 to 0W when presaturation is not required
;   use pl1 + 75 to 80dB to reduce radiation damping

;Processing

;PHC0(F1): 90
;PHC1(F1): -180
;FCOR(F1): 1

;$Id: dipsi2esgpph,v 1.15 2019/02/11 14:04:14 ber Exp $

2D ROESY

;roesyphpr
;avance-version (12/01/11)
;2D ROESY with cw spinlock for mixing
;phase sensitive
;
;A. Bax & D.G. Davis, J. Magn. Reson 63, 207-213 (1985)
;
;$CLASS=HighRes
;$DIM=2D
;$TYPE=
;$SUBTYPE=
;$COMMENT=

#include <Avance.incl>

"d11=30m"
"d12=20u"
"d13=4u"

"in0=inf1"

"d0=in0/2-p1*2/3.1416-4u"

"acqt0=0"

1 ze
2 d11
3 d12 pl9:f1
  d1 cw:f1 ph29
  d13 do:f1
  d12 pl1:f1
  p1 ph1
  d0

121

```
    4u pl11:f1
    p15 ph2
    go=2 ph31
    d11 mc #0 to 2 F1PH(calph(ph1, +90)
& calph(ph29, +90), caldel(d0, +in0))
    exit

    ph1=0 2 2 0 1 3 3 1
    ph2=0 2 0 2 1 3 1 3
    ph29=0
    ph31=0 2 2 0 1 3 3 1
```

;pl1 : f1 channel - power level for pulse (default)

;pl9 : f1 channel - power level for presaturation

;pl11: f1 channel - power level for ROESY-spinlock

;p1 : f1 channel - 90 degree high power pulse

;p15: f1 channel - pulse for ROESY spinlock

;d0 : incremented delay (2D)

;d1 : relaxation delay; 1-5 * T1

;d11: delay for disk I/O [30 msec]

;d12: delay for power switching [20 usec]

;d13: short delay [4 usec]

;inf1: 1/SW = 2 * DW

;in0: 1/(1 * SW) = 2 * DW

;nd0: 1

;ns: 8 * n

;ds: 16

;td1: number of experiments

;FnMODE: States-TPPI, TPPI, States or QSEQ

;Processing

;PHC0(F1): 180

;PHC1(F1): -180

;FCOR(F1): 1

;$Id: roesyphpr,v 1.11 2012/01/31 17:49:28 ber Exp $

Appendix C: G-quadruplex Chemical Shifts and Restraints for Structure Calculations and Modelling

URL for CNSsolve: http://cns-online.org/v1.3/

All restraints are in CNS format. All CNS taken from the CNS package:

Brunger, A. T.; Adams, P. D.; Clore, G. M.; Delano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Cryst.* **1998**, *54*, 905–921.

Brunger, A. T. *Nat. Protoc.* **2007**, *2* (11), 2728–2733.

NOE Restraint Format:

assi (atom identifier) (atom identifier) maximum distance estimated distance minimum distance

Hydrogen Bond Restraint Format:

assign (atom identifier) (atom identifier) angle +angle -angle

Planarity Restraint Format:

selection=   ((atom identifier) or (atom identifier) or
(atom identifier) or (atom identifier) or
(atom identifier) or (atom identifier))

Quadruplex A

NOE Restraints

assi (resid 15 and name H1)(resid 16 and name H1) 3.5 2.5 2.5
assi (resid 8 and name H1)(resid 7 and name H1) 3.5 2.5 2.5
assi (resid 18 and name H1)(resid 19 and name H1) 3.5 2.5 2.5
assi (resid 25 and name H8)(resid 24 and name H6) 4.0 2.5 2.5

assi (resid 15 and name H1')(resid 16 and name H8) 4.0 2.5 2.5

assi (resid 8 and name H1')(resid 7 and name H8) 4.0 2.5 2.5

assi (resid 18 and name H1')(resid 19 and name H8) 4.0 2.5 2.5

Hydrogen Bond Restraints

! for G15/ G8 base pair

assign (resid 8 and name N1) (resid 15 and name O6)   3.00   1.0   1.0

assign (resid 8 and name H1) (resid 15 and name O6)   2.10   1.0   1.0

assign (resid 8 and name N2) (resid 15 and name N7)   2.80   1.0   1.0

assign (resid 8 and name H22) (resid 15 and name N7)   1.80   1.0   1.0

! for G19/ T24 base pair

assign (resid 19 and name O6) (resid 24 and name N3)   3.00   1.5   1.5

assign (resid 19 and name O6) (resid 24 and name H3)   2.00   1.5   1.5

assign (resid 19 and name N1) (resid 24 and name O4)   3.00   1.5   1.5

assign (resid 19 and name H1) (resid 24 and name O4)   2.10   1.5   1.5

! for G8/ G19 base pair

assign (resid 19 and name N2) (resid 8 and name O6)   3.00   1.0   1.0

assign (resid 19 and name H22) (resid 8 and name O6)   2.10   1.0   1.0

assign (resid 19 and name N2) (resid 8 and name N7)   2.80   1.0   1.0

assign (resid 19 and name H21) (resid 8 and name N7)   1.80   1.0   1.0

! for G15/ T24 base pair

assign (resid 15 and name N2) (resid 24 and name O4)   2.91   1.5   1.5

assign (resid 15 and name H22) (resid 24 and name O4)   2.91   1.5   1.5

! for G16/ G7 base pair

assign (resid 7 and name N1) (resid 16 and name O6)   3.00   1.0   1.0

assign (resid 7 and name H1) (resid 16 and name O6)   2.10   1.0   1.0

assign (resid 7 and name N2) (resid 16 and name N7)   2.80   1.0   1.0

assign (resid 7 and name H22) (resid 16 and name N7)   1.80   1.0   1.0

! for G18/ G25 base pair

assign (resid 18 and name N1) (resid 25 and name O6)   3.00   1.0   1.0

assign (resid 18 and name H1) (resid 25 and name O6)   2.10   1.0   1.0

assign (resid 18 and name N2) (resid 25 and name N7)   2.80   1.0   1.0

assign (resid 18 and name H22) (resid 25 and name N7)   1.80   1.0   1.0

! for G7/ G18 base pair

assign (resid 18 and name N1) (resid 7 and name O6)   3.00   1.5   1.5

assign (resid 18 and name H1) (resid 7 and name O6)   2.10   1.5   1.5

assign (resid 18 and name N2) (resid 7 and name N7)   2.80   1.5   1.5

assign (resid 18 and name H21) (resid 7 and name N7)   1.80   1.5   1.5

! for G25/ G16 base pair

assign (resid 25 and name N1) (resid 16 and name O6)   3.00   1.0   1.0

assign (resid 25 and name H1) (resid 16 and name O6)   2.10   1.0   1.0

assign (resid 25 and name N2) (resid 16 and name N7)   2.80   1.0   1.0

assign (resid 25 and name H22) (resid 16 and name N7)   1.80   1.0   1.0

Planarity Restraints

! G8-G15

!--

group

selection=   ((resid 8 and name N1) or (resid 8 and name N3) or

        (resid 8 and name C5) or (resid 15 and name N1) or

        (resid 15 and name N3) or (resid 15 and name C5))

weight = $pscale end

! G19-T24

!--

group

selection=   ((resid 19 and name N1) or (resid 19 and name N3) or

        (resid 19 and name C5) or (resid 24 and name N1) or

        (resid 24 and name N3) or (resid 24 and name C5))

weight = $pscale end

! G8-G19

!--

group

selection=   ((resid 8 and name N1) or (resid 8 and name N3) or

       (resid 8 and name C5) or (resid 19 and name N1) or

       (resid 19 and name N3) or (resid 19 and name C5))

weight = $pscale end

! G19-T24

!---

group

selection=   ((resid 19 and name N1) or (resid 19 and name N3) or

       (resid 19 and name C5) or (resid 24 and name N1) or

       (resid 24 and name N3) or (resid 24 and name C5))

weight = $pscale end

! G7-T16

!---

group

selection=   ((resid 7 and name N1) or (resid 7 and name N3) or

       (resid 7 and name C5) or (resid 16 and name N1) or

       (resid 16 and name N3) or (resid 16 and name C5))

weight = $pscale end

! G18-G25

!---

group

selection=   ((resid 18 and name N1) or (resid 18 and name N3) or

       (resid 18 and name C5) or (resid 25 and name N1) or

       (resid 25 and name N3) or (resid 25 and name C5))

weight = $pscale end

! G7-G18

!---

group

selection=   ((resid 7 and name N1) or (resid 7 and name N3) or

       (resid 7 and name C5) or (resid 18 and name N1) or

       (resid 18 and name N3) or (resid 18 and name C5))

weight = $pscale end

! G25-G16

!--

group

selection=   ((resid 25 and name N1) or (resid 25 and name N3) or

(resid 25 and name C5) or (resid 16 and name N1) or

(resid 16 and name N3) or (resid 16 and name C5))

weight = $pscale end

Resulting Model

Figure C.1 Quadruplex A as Modelled in CNS.[79,80,159] A) The bases with hydrogen bonds to form the predicted G-quadruplex. B) An overview with arrows to indicate the directionality of the DNA strand.

Quadruplex B

NOE Restraints

assi (resid 15 and name H1)(resid 16 and name H1) 3.5 2.5 2.5

assi (resid 9 and name H1)(resid 8 and name H1) 3.5 2.5 2.5

assi (resid 18 and name H1)(resid 19 and name H1) 3.5 2.5 2.5

assi (resid 25 and name H8)(resid 24 and name H6) 4.0 2.5 2.5

assi (resid 15 and name H1')(resid 16 and name H8) 4.0 2.5 2.5

assi (resid 9 and name H1')(resid 8 and name H8) 4.0 2.5 2.5

assi (resid 18 and name H1')(resid 19 and name H8) 4.0 2.5 2.5

Hydrogen Bond Restraints

! for G15/ G9 base pair

assign (resid 9 and name N1) (resid 15 and name O6)   3.00   1.0   1.0

assign (resid 9 and name H1) (resid 15 and name O6)   2.10   1.0   1.0

assign (resid 9 and name N2) (resid 15 and name N7)   2.80   1.0   1.0

assign (resid 9 and name H22) (resid 15 and name N7)   1.80   1.0   1.0

! for G19/ T24 base pair

assign (resid 19 and name O6) (resid 24 and name N3)   3.00   1.5   1.5

assign (resid 19 and name O6) (resid 24 and name H3)   2.00   1.5   1.5

assign (resid 19 and name N1) (resid 24 and name O4)   3.00   1.5   1.5

assign (resid 19 and name H1) (resid 24 and name O4)   2.10   1.5   1.5

! for G9/ G19 base pair

assign (resid 19 and name N2) (resid 9 and name O6)   3.00   1.0   1.0

assign (resid 19 and name H22) (resid 9 and name O6)   2.10   1.0   1.0

assign (resid 19 and name N2) (resid 9 and name N7)   2.80   1.0   1.0

assign (resid 19 and name H21) (resid 9 and name N7)   1.80   1.0   1.0

! for G15/ T24 base pair

assign (resid 15 and name N2) (resid 24 and name O4)   2.91   1.5   1.5

assign (resid 15 and name H22) (resid 24 and name O4)   2.91   1.5   1.5

! for G16/ G8 base pair

assign (resid 8 and name N1) (resid 16 and name O6)   3.00   1.0   1.0

assign (resid 8 and name H1) (resid 16 and name O6)   2.10   1.0   1.0

assign (resid 8 and name N2) (resid 16 and name N7)   2.80   1.0   1.0

assign (resid 8 and name H22) (resid 16 and name N7)   1.80   1.0   1.0

! for G18/ G25 base pair

assign (resid 18 and name N1) (resid 25 and name O6)   3.00   1.0   1.0

assign (resid 18 and name H1) (resid 25 and name O6)   2.10   1.0   1.0

assign (resid 18 and name N2) (resid 25 and name N7)   2.80   1.0   1.0

assign (resid 18 and name H22) (resid 25 and name N7)   1.80   1.0   1.0

! for G8/ G18 base pair

assign (resid 18 and name N1) (resid 8 and name O6)   3.00   1.5   1.5

assign (resid 18 and name H1) (resid 8 and name O6)   2.10   1.5   1.5

assign (resid 18 and name N2) (resid 8 and name N7)   2.80   1.5   1.5

assign (resid 18 and name H21) (resid 8 and name N7)   1.80   1.5   1.5

! for G25/ G16 base pair

assign (resid 25 and name N1) (resid 16 and name O6)   3.00   1.0   1.0

assign (resid 25 and name H1) (resid 16 and name O6)   2.10   1.0   1.0

assign (resid 25 and name N2) (resid 16 and name N7)   2.80   1.0   1.0

assign (resid 25 and name H22) (resid 16 and name N7)   1.80   1.0   1.0

Planarity Restraints

! G9-G15

!---

group

selection=   ((resid 9 and name N1) or (resid 9 and name N3) or

        (resid 9 and name C5) or (resid 15 and name N1) or

        (resid 15 and name N3) or (resid 15 and name C5))

weight = $pscale end

! G19-T24

!---

group

selection=   ((resid 19 and name N1) or (resid 19 and name N3) or

        (resid 19 and name C5) or (resid 24 and name N1) or

        (resid 24 and name N3) or (resid 24 and name C5))

weight = $pscale end

! G9-G19

!--

group

selection=   ((resid 9 and name N1) or (resid 9 and name N3) or

        (resid 9 and name C5) or (resid 19 and name N1) or

        (resid 19 and name N3) or (resid 19 and name C5))

weight = $pscale end

! G15-T24

!--

group

selection=   ((resid 15 and name N1) or (resid 15 and name N3) or

        (resid 15 and name C5) or (resid 24 and name N1) or

        (resid 24 and name N3) or (resid 24 and name C5))

weight = $pscale end

! G8-T16

!--

group

selection=   ((resid 8 and name N1) or (resid 8 and name N3) or

        (resid 8 and name C5) or (resid 16 and name N1) or

        (resid 16 and name N3) or (resid 16 and name C5))

weight = $pscale end

! G18-G25

!--

group

selection=   ((resid 18 and name N1) or (resid 18 and name N3) or

        (resid 18 and name C5) or (resid 25 and name N1) or

        (resid 25 and name N3) or (resid 25 and name C5))

weight = $pscale end

! G8-G18

!--

group

selection=   ((resid 8 and name N1) or (resid 8 and name N3) or

(resid 8 and name C5) or (resid 18 and name N1) or

(resid 18 and name N3) or (resid 18 and name C5))

weight = $pscale end

! G25-G16

!---

group

selection=   ((resid 25 and name N1) or (resid 25 and name N3) or

(resid 25 and name C5) or (resid 16 and name N1) or

(resid 16 and name N3) or (resid 16 and name C5))

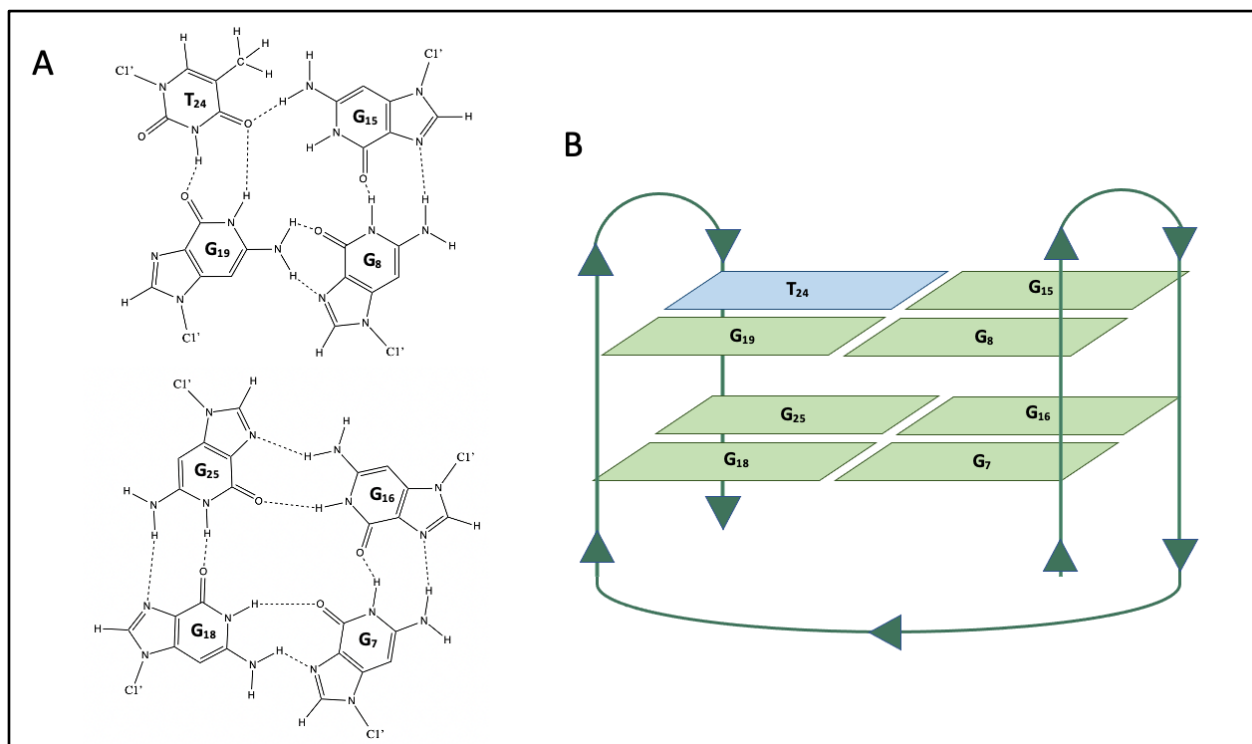weight = $pscale end

Resulting Model

Figure C.2 Quadruplex B as Modelled in CNS.[79,80,159] A) The bases with hydrogen bonds to form the predicted G-quadruplex. B) An overview with arrows to indicate the directionality of the DNA strand.

Quadruplex C

NOE Restraints

assi (resid 7 and name H1)(resid 8 and name H1) 3.5 2.5 2.5

assi (resid 15 and name H1)(resid 16 and name H1) 3.5 2.5 2.5

assi (resid 19 and name H1)(resid 18 and name H1) 3.5 2.5 2.5

assi (resid 25 and name H8)(resid 24 and name H6) 4.0 2.5 2.5

assi (resid 7 and name H1')(resid 8 and name H8) 4.0 2.5 2.5

assi (resid 15 and name H1')(resid 16 and name H8) 4.0 2.5 2.5

assi (resid 19 and name H1')(resid 18 and name H8) 4.0 2.5 2.5

Hydrogen Bond Restraints

! for G7/ G15 base pair

assign (resid 15 and name N1) (resid 7 and name O6)   3.00   1.0   1.0

assign (resid 15 and name H1) (resid 7 and name O6)   2.10   1.0   1.0

assign (resid 15 and name N2) (resid 7 and name N7)   2.80   1.0   1.0

assign (resid 15 and name H22) (resid 7 and name N7)   1.80   1.0   1.0

! for G18/ T24 base pair

assign (resid 18 and name O6) (resid 24 and name N3)   3.00   1.5   1.5

assign (resid 18 and name O6) (resid 24 and name H3)   2.00   1.5   1.5

assign (resid 18 and name N1) (resid 24 and name O4)   3.00   1.5   1.5

assign (resid 18 and name H1) (resid 24 and name O4)   2.10   1.5   1.5

! for G15/ G18 base pair

assign (resid 18 and name N2) (resid 15 and name O6)   3.00   1.0   1.0

assign (resid 18 and name H22) (resid 15 and name O6)   2.10   1.0   1.0

assign (resid 18 and name N2) (resid 15 and name N7)   2.80   1.0   1.0

assign (resid 18 and name H21) (resid 15 and name N7)   1.80   1.0   1.0

! for G7/ T24 base pair

assign (resid 7 and name N2) (resid 24 and name O4)   2.91   1.5   1.5

assign (resid 7 and name H22) (resid 24 and name O4)   2.91   1.5   1.5

! for G8/ G16 base pair

assign (resid 16 and name N1) (resid 8 and name O6)   3.00   1.0   1.0

assign (resid 16 and name H1) (resid 8 and name O6)   2.10   1.0   1.0

assign (resid 16 and name N2) (resid 8 and name N7)   2.80   1.0   1.0

assign (resid 16 and name H22) (resid 8 and name N7)   1.80   1.0   1.0

! for G19/ G25 base pair

assign (resid 19 and name N1) (resid 25 and name O6)   3.00   1.0   1.0

assign (resid 19 and name H1) (resid 25 and name O6)   2.10   1.0   1.0

assign (resid 19 and name N2) (resid 25 and name N7)   2.80   1.0   1.0

assign (resid 19 and name H22) (resid 25 and name N7)   1.80   1.0   1.0

! for G16/ G19 base pair

assign (resid 19 and name N1) (resid 16 and name O6)   3.00   1.5   1.5

assign (resid 19 and name H1) (resid 16 and name O6)   2.10   1.5   1.5

assign (resid 19 and name N2) (resid 16 and name N7)   2.80   1.5   1.5

assign (resid 19 and name H21) (resid 16 and name N7)   1.80   1.5   1.5

! for G25/ G8 base pair

assign (resid 25 and name N1) (resid 8 and name O6)   3.00   1.0   1.0

assign (resid 25 and name H1) (resid 8 and name O6)   2.10   1.0   1.0

assign (resid 25 and name N2) (resid 8 and name N7)   2.80   1.0   1.0

assign (resid 25 and name H22) (resid 8 and name N7)   1.80   1.0   1.0

Planarity Restraints

! G7-G15

!--

group

selection=   ((resid 7 and name N1) or (resid 7 and name N3) or

        (resid 7 and name C5) or (resid 15 and name N1) or

        (resid 15 and name N3) or (resid 15 and name C5))

weight = $pscale end

! G18-T24

```
!-------------------------------------------------------------------
group
selection=    ((resid  18 and name N1) or (resid  18 and name N3) or
              (resid  18 and name C5) or (resid 24 and name N1) or
              (resid 24 and name N3) or (resid 24 and name C5))
weight = $pscale end
! G15-G18
!-------------------------------------------------------------------
group
selection=    ((resid  15 and name N1) or (resid  15 and name N3) or
              (resid  15 and name C5) or (resid 18 and name N1) or
              (resid 18 and name N3) or (resid 18 and name C5))
weight = $pscale end
! G7-T24
!-------------------------------------------------------------------
group
selection=    ((resid  7 and name N1) or (resid  7 and name N3) or
              (resid  7 and name C5) or (resid 24 and name N1) or
              (resid 24 and name N3) or (resid 24 and name C5))
weight = $pscale end
! G8-T16
!-------------------------------------------------------------------
group
selection=    ((resid  8 and name N1) or (resid  8 and name N3) or
              (resid  8 and name C5) or (resid 16 and name N1) or
              (resid 16 and name N3) or (resid 16 and name C5))
weight = $pscale end
! G19-G25
!-------------------------------------------------------------------
group
selection=    ((resid  19 and name N1) or (resid  19 and name N3) or
```

(resid 19 and name C5) or (resid 25 and name N1) or

(resid 25 and name N3) or (resid 25 and name C5))

weight = $pscale end

! G16-G19

!---

group

selection=   ((resid 16 and name N1) or (resid 16 and name N3) or

(resid 16 and name C5) or (resid 19 and name N1) or

(resid 19 and name N3) or (resid 19 and name C5))

weight = $pscale end

! G25-G8

!---

group

selection=   ((resid 25 and name N1) or (resid 25 and name N3) or

(resid 25 and name C5) or (resid 8 and name N1) or

(resid 8 and name N3) or (resid 8 and name C5))

weight = $pscale end
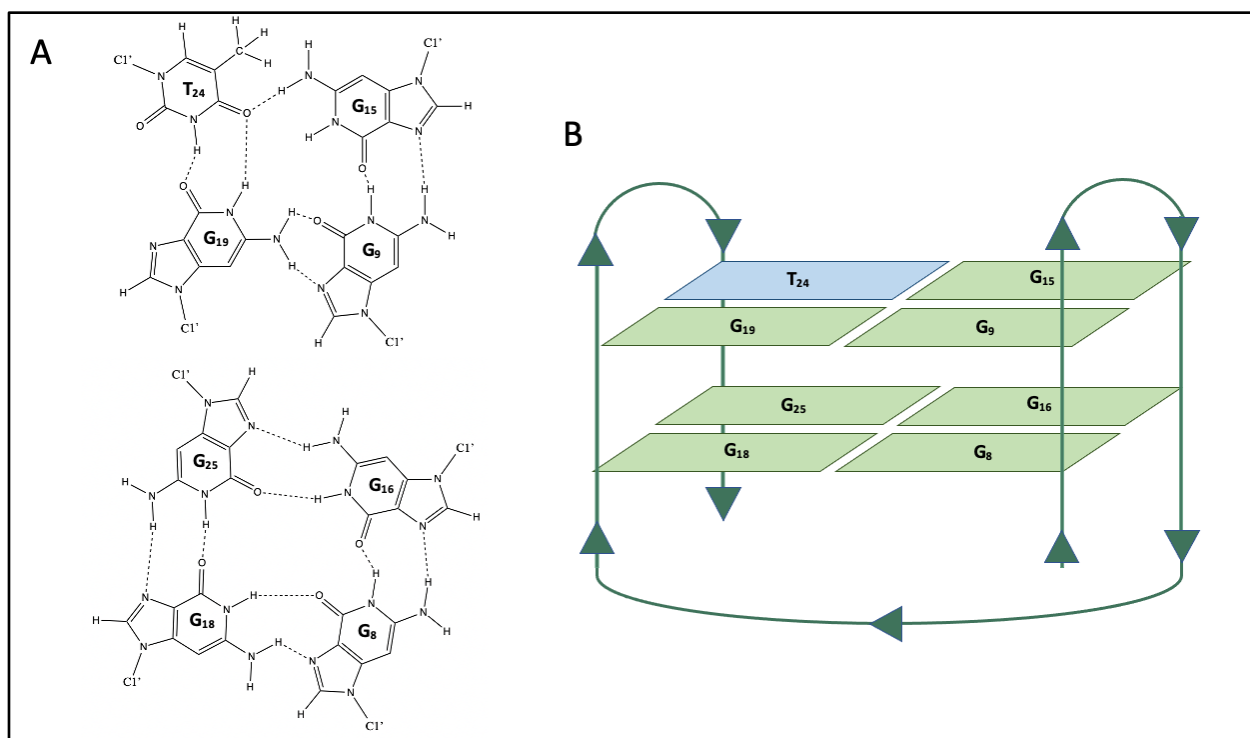
Resulting Model

Figure C.3 Quadruplex C as Modelled in CNS.[79,80,159] A) The bases with hydrogen bonds to form the predicted G-quadruplex. B) An overview with arrows to indicate the directionality of the DNA strand.

Quadruplex D

NOE Restraints

assi (resid 8 and name H1)(resid 9 and name H1) 3.5 2.5 2.5

assi (resid 15 and name H1)(resid 16 and name H1) 3.5 2.5 2.5

assi (resid 19 and name H1)(resid 18 and name H1) 3.5 2.5 2.5

assi (resid 25 and name H8)(resid 24 and name H6) 4.0 2.5 2.5

assi (resid 8 and name H1')(resid 9 and name H8) 4.0 2.5 2.5

assi (resid 15 and name H1')(resid 16 and name H8) 4.0 2.5 2.5

assi (resid 19 and name H1')(resid 18 and name H8) 4.0 2.5 2.5

Hydrogen Bond Restraints

! for G8/ G15 base pair

assign (resid 15 and name N1) (resid 8 and name O6)   3.00   1.0   1.0

assign (resid 15 and name H1) (resid 8 and name O6)   2.10   1.0   1.0

assign (resid 15 and name N2) (resid 8 and name N7)   2.80   1.0   1.0

assign (resid 15 and name H22) (resid 8 and name N7)   1.80   1.0   1.0

! for G18/ T24 base pair

assign (resid 18 and name O6) (resid 24 and name N3)   3.00   1.5   1.5

assign (resid 18 and name O6) (resid 24 and name H3)   2.00   1.5   1.5

assign (resid 18 and name N1) (resid 24 and name O4)   3.00   1.5   1.5

assign (resid 18 and name H1) (resid 24 and name O4)   2.10   1.5   1.5

! for G15/ G18 base pair

assign (resid 18 and name N2) (resid 15 and name O6)   3.00   1.0   1.0

assign (resid 18 and name H22) (resid 15 and name O6)   2.10   1.0   1.0

assign (resid 18 and name N2) (resid 15 and name N7)   2.80   1.0   1.0

assign (resid 18 and name H21) (resid 15 and name N7)   1.80   1.0   1.0

! for G8/ T24 base pair

assign (resid 8 and name N2) (resid 24 and name O4)   2.91   1.5   1.5

assign (resid 8 and name H22) (resid 24 and name O4)   2.91   1.5   1.5

! for G9/ G16 base pair

assign (resid 16 and name N1) (resid 9 and name O6)   3.00   1.0   1.0

assign (resid 16 and name H1) (resid 9 and name O6)   2.10   1.0   1.0

assign (resid 16 and name N2) (resid 9 and name N7)   2.80   1.0   1.0

assign (resid 16 and name H22) (resid 9 and name N7)   1.80   1.0   1.0

! for G19/ G25 base pair

assign (resid 19 and name N1) (resid 25 and name O6)   3.00   1.0   1.0

assign (resid 19 and name H1) (resid 25 and name O6)   2.10   1.0   1.0

assign (resid 19 and name N2) (resid 25 and name N7)   2.80   1.0   1.0

assign (resid 19 and name H22) (resid 25 and name N7)   1.80   1.0   1.0

! for G16/ G19 base pair

assign (resid 19 and name N1) (resid 16 and name O6)   3.00   1.5   1.5

assign (resid 19 and name H1) (resid 16 and name O6)   2.10   1.5   1.5

assign (resid 19 and name N2) (resid 16 and name N7)   2.80   1.5   1.5

assign (resid 19 and name H21) (resid 16 and name N7)   1.80   1.5   1.5

! for G25/ G9 base pair

assign (resid 25 and name N1) (resid 9 and name O6)   3.00   1.0   1.0

assign (resid 25 and name H1) (resid 9 and name O6)   2.10   1.0   1.0

assign (resid 25 and name N2) (resid 9 and name N7)   2.80   1.0   1.0

assign (resid 25 and name H22) (resid 9 and name N7)   1.80   1.0   1.0

Planarity Restraints

! G8-G15

!--

group

selection=   ((resid 8 and name N1) or (resid 8 and name N3) or

          (resid 8 and name C5) or (resid 15 and name N1) or

          (resid 15 and name N3) or (resid 15 and name C5))

weight = $pscale end

! G18-T24

!--

group

selection=   ((resid 18 and name N1) or (resid 18 and name N3) or

          (resid 18 and name C5) or (resid 24 and name N1) or

          (resid 24 and name N3) or (resid 24 and name C5))

weight = $pscale end

! G15-G18

!--

group

selection=   ((resid 15 and name N1) or (resid 15 and name N3) or

          (resid 15 and name C5) or (resid 18 and name N1) or

          (resid 18 and name N3) or (resid 18 and name C5))

weight = $pscale end

! G8-T24

!---

group

selection=   ((resid 8 and name N1) or (resid 8 and name N3) or

          (resid 8 and name C5) or (resid 24 and name N1) or

          (resid 24 and name N3) or (resid 24 and name C5))

weight = $pscale end

! G9-T16

!---

group

selection=   ((resid 9 and name N1) or (resid 9 and name N3) or

          (resid 9 and name C5) or (resid 16 and name N1) or

          (resid 16 and name N3) or (resid 16 and name C5))

weight = $pscale end

! G19-G25

!---

group

selection=   ((resid 19 and name N1) or (resid 19 and name N3) or

          (resid 19 and name C5) or (resid 25 and name N1) or

          (resid 25 and name N3) or (resid 25 and name C5))

weight = $pscale end

! G16-G19

!---

group

selection=   ((resid 16 and name N1) or (resid 16 and name N3) or

          (resid 16 and name C5) or (resid 19 and name N1) or

          (resid 19 and name N3) or (resid 19 and name C5))

weight = $pscale end

! G25-G9

!---

group

selection=   ((resid 25 and name N1) or (resid 25 and name N3) or

(resid 25 and name C5) or (resid 9 and name N1) or

(resid 9 and name N3) or (resid 9 and name C5))
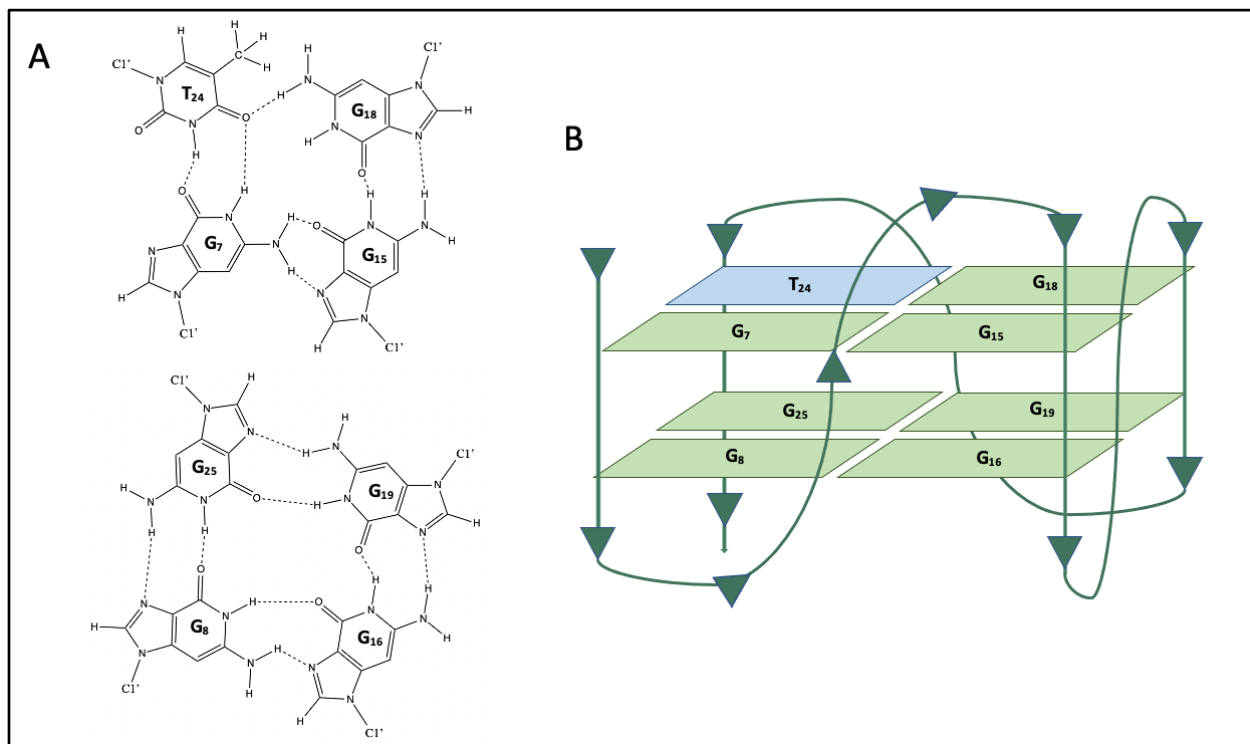
weight = $pscale end

Resulting Model

Figure C.4 Quadruplex D as Modelled in CNS.[79,80,159] A) The bases with hydrogen bonds to form the predicted G-quadruplex. B) An overview with arrows to indicate the directionality of the DNA strand.

Quadruplex E

NOE Restraints

assi (resid 7 and name H1)(resid 8 and name H1) 3.5 2.5 2.5

assi (resid 16 and name H1)(resid 15 and name H1) 3.5 2.5 2.5

assi (resid 18 and name H1)(resid 19 and name H1) 3.5 2.5 2.5

assi (resid 25 and name H8)(resid 24 and name H6) 4.0 2.5 2.5

assi (resid 7 and name H1')(resid 8 and name H8) 4.0 2.5 2.5

assi (resid 16 and name H1')(resid 15 and name H8) 4.0 2.5 2.5

assi (resid 18 and name H1')(resid 19 and name H8) 4.0 2.5 2.5

Hydrogen Bond Restraints

! for G8/ G15 base pair

assign (resid 15 and name N1) (resid 8 and name O6)   3.00   1.0   1.0

assign (resid 15 and name H1) (resid 8 and name O6)   2.10   1.0   1.0

assign (resid 15 and name N2) (resid 8 and name N7)   2.80   1.0   1.0

assign (resid 15 and name H22) (resid 8 and name N7)   1.80   1.0   1.0

! for G19/ T24 base pair

assign (resid 19 and name O6) (resid 24 and name N3)   3.00   1.5   1.5

assign (resid 19 and name O6) (resid 24 and name H3)   2.00   1.5   1.5

assign (resid 19 and name N1) (resid 24 and name O4)   3.00   1.5   1.5

assign (resid 19 and name H1) (resid 24 and name O4)   2.10   1.5   1.5

! for G15/ G19 base pair

assign (resid 19 and name N2) (resid 15 and name O6)   3.00   1.0   1.0

assign (resid 19 and name H22) (resid 15 and name O6)   2.10   1.0   1.0

assign (resid 19 and name N2) (resid 15 and name N7)   2.80   1.0   1.0

assign (resid 19 and name H21) (resid 15 and name N7)   1.80   1.0   1.0

! for G8/ T24 base pair

assign (resid 8 and name N2) (resid 24 and name O4)   2.91   1.5   1.5

assign (resid 8 and name H22) (resid 24 and name O4)   2.91   1.5   1.5

! for G7/ G16 base pair

assign (resid 16 and name N1) (resid 7 and name O6)   3.00   1.0   1.0

assign (resid 16 and name H1) (resid 7 and name O6)   2.10   1.0   1.0

assign (resid 16 and name N2) (resid 7 and name N7)   2.80   1.0   1.0

assign (resid 16 and name H22) (resid 7 and name N7)   1.80   1.0   1.0

! for G18/ G25 base pair

assign (resid 18 and name N1) (resid 25 and name O6)   3.00   1.0   1.0

assign (resid 18 and name H1) (resid 25 and name O6)   2.10   1.0   1.0

assign (resid 18 and name N2) (resid 25 and name N7)   2.80   1.0   1.0

assign (resid 18 and name H22) (resid 25 and name N7)   1.80   1.0   1.0

! for G16/ G18 base pair

assign (resid 18 and name N1) (resid 16 and name O6)   3.00   1.5   1.5

assign (resid 18 and name H1) (resid 16 and name O6)   2.10   1.5   1.5

assign (resid 18 and name N2) (resid 16 and name N7)   2.80   1.5   1.5

assign (resid 18 and name H21) (resid 16 and name N7)   1.80   1.5   1.5

! for G25/ G7 base pair

assign (resid 25 and name N1) (resid 7 and name O6)   3.00   1.0   1.0

assign (resid 25 and name H1) (resid 7 and name O6)   2.10   1.0   1.0

assign (resid 25 and name N2) (resid 7 and name N7)   2.80   1.0   1.0

assign (resid 25 and name H22) (resid 7 and name N7)   1.80   1.0   1.0

Planarity Restraints

! G8-G15

!---

group

selection=   ((resid 8 and name N1) or (resid 8 and name N3) or

        (resid 8 and name C5) or (resid 15 and name N1) or

        (resid 15 and name N3) or (resid 15 and name C5))

weight = $pscale end

! G19-T24

!---

group

selection=   ((resid 19 and name N1) or (resid 19 and name N3) or

        (resid 19 and name C5) or (resid 24 and name N1) or

        (resid 24 and name N3) or (resid 24 and name C5))

weight = $pscale end

! G16-G19

!---

group

selection=   ((resid 15 and name N1) or (resid 15 and name N3) or

(resid 15 and name C5) or (resid 19 and name N1) or

(resid 19 and name N3) or (resid 19 and name C5))

weight = $pscale end

! G8-T24

!---

group

selection=   ((resid 8 and name N1) or (resid 8 and name N3) or

(resid 8 and name C5) or (resid 24 and name N1) or

(resid 24 and name N3) or (resid 24 and name C5))

weight = $pscale end

! G7-T16

!---

group

selection=   ((resid 7 and name N1) or (resid 7 and name N3) or

(resid 7 and name C5) or (resid 16 and name N1) or

(resid 16 and name N3) or (resid 16 and name C5))

weight = $pscale end

! G18-G25

!---

group

selection=   ((resid 18 and name N1) or (resid 18 and name N3) or

(resid 18 and name C5) or (resid 25 and name N1) or

(resid 25 and name N3) or (resid 25 and name C5))

weight = $pscale end

! G16-G18

!---

group

selection=   ((resid 16 and name N1) or (resid 16 and name N3) or

(resid 16 and name C5) or (resid 18 and name N1) or

(resid 18 and name N3) or (resid 18 and name C5))

143

weight = $pscale end

! G25-G7

!--

group

selection=   ((resid 25 and name N1) or (resid 25 and name N3) or

(resid 25 and name C5) or (resid 7 and name N1) or

(resid 7 and name N3) or (resid 7 and name C5))
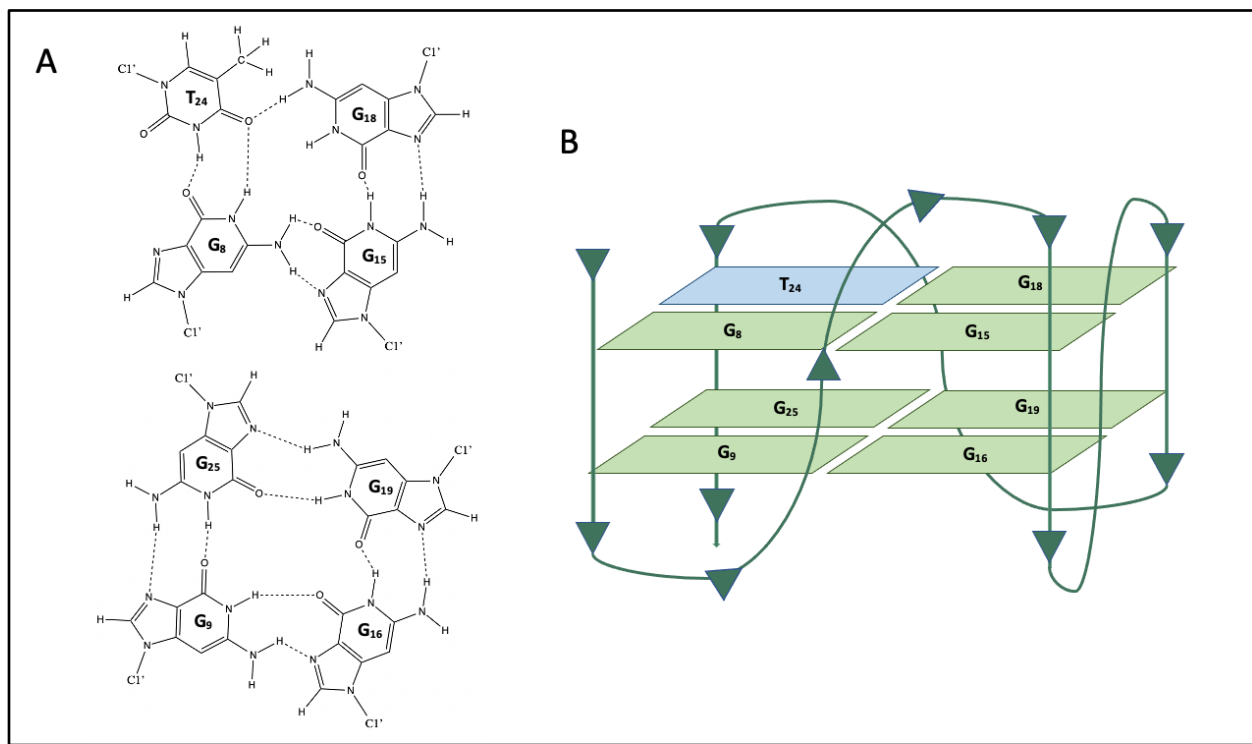
weight = $pscale end

Resulting Model
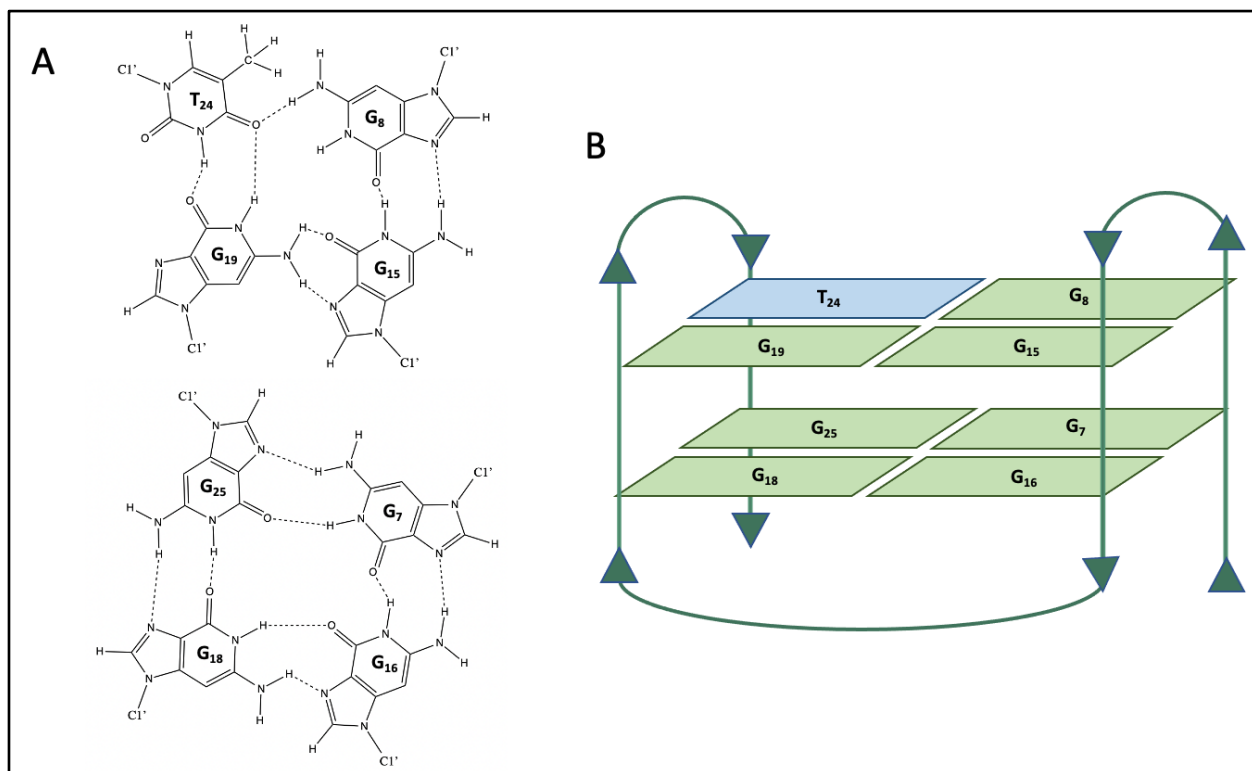
Figure C.5 Quadruplex E as Modelled in CNS.[79,80,159] A) The bases with hydrogen bonds to form the predicted G-quadruplex. B) An overview with arrows to indicate the directionality of the DNA strand.

Quadruplex F

NOE Restraints

assi (resid 8 and name H1)(resid 9 and name H1) 3.5 2.5 2.5

assi (resid 16 and name H1)(resid 15 and name H1) 3.5 2.5 2.5

assi (resid 18 and name H1)(resid 19 and name H1) 3.5 2.5 2.5

assi (resid 25 and name H8)(resid 24 and name H6) 4.0 2.5 2.5

assi (resid 8 and name H1')(resid 9 and name H8) 4.0 2.5 2.5

assi (resid 16 and name H1')(resid 15 and name H8) 4.0 2.5 2.5

assi (resid 18 and name H1')(resid 19 and name H8) 4.0 2.5 2.5

Hydrogen Bond Restraints

! for G9/ G15 base pair

assign (resid 15 and name N1) (resid 9 and name O6)   3.00   1.0   1.0

assign (resid 15 and name H1) (resid 9 and name O6)   2.10   1.0   1.0

assign (resid 15 and name N2) (resid 9 and name N7)   2.80   1.0   1.0

assign (resid 15 and name H22) (resid 9 and name N7)   1.80   1.0   1.0

! for G19/ T24 base pair

assign (resid 19 and name O6) (resid 24 and name N3)   3.00   1.5   1.5

assign (resid 19 and name O6) (resid 24 and name H3)   2.00   1.5   1.5

assign (resid 19 and name N1) (resid 24 and name O4)   3.00   1.5   1.5

assign (resid 19 and name H1) (resid 24 and name O4)   2.10   1.5   1.5

! for G15/ G19 base pair

assign (resid 19 and name N2) (resid 15 and name O6)   3.00   1.0   1.0

assign (resid 19 and name H22) (resid 15 and name O6)   2.10   1.0   1.0

assign (resid 19 and name N2) (resid 15 and name N7)   2.80   1.0   1.0

assign (resid 19 and name H21) (resid 15 and name N7)   1.80   1.0   1.0

! for G9/ T24 base pair

assign (resid 9 and name N2) (resid 24 and name O4)   2.91   1.5   1.5

assign (resid 9 and name H22) (resid 24 and name O4)   2.91   1.5   1.5

! for G8/ G16 base pair

assign (resid 16 and name N1) (resid 8 and name O6)   3.00   1.0   1.0

assign (resid 16 and name H1) (resid 8 and name O6)   2.10   1.0   1.0

assign (resid 16 and name N2) (resid 8 and name N7)   2.80   1.0   1.0

assign (resid 16 and name H22) (resid 8 and name N7)   1.80   1.0   1.0

! for G18/ G25 base pair

assign (resid 18 and name N1) (resid 25 and name O6)   3.00   1.0   1.0

assign (resid 18 and name H1) (resid 25 and name O6)   2.10   1.0   1.0

assign (resid 18 and name N2) (resid 25 and name N7)   2.80   1.0   1.0

assign (resid 18 and name H22) (resid 25 and name N7)   1.80   1.0   1.0

! for G16/ G18 base pair

assign (resid 18 and name N1) (resid 16 and name O6)   3.00   1.5   1.5

assign (resid 18 and name H1) (resid 16 and name O6)   2.10   1.5   1.5

assign (resid 18 and name N2) (resid 16 and name N7)   2.80   1.5   1.5

assign (resid 18 and name H21) (resid 16 and name N7)   1.80   1.5   1.5

! for G25/ G8 base pair

assign (resid 25 and name N1) (resid 8 and name O6)   3.00   1.0   1.0

assign (resid 25 and name H1) (resid 8 and name O6)   2.10   1.0   1.0

assign (resid 25 and name N2) (resid 8 and name N7)   2.80   1.0   1.0

assign (resid 25 and name H22) (resid 8 and name N7)   1.80   1.0   1.0

Planarity Restraints

! G9-G15

!--

group

selection=   ((resid 9 and name N1) or (resid 9 and name N3) or

        (resid 9 and name C5) or (resid 15 and name N1) or

        (resid 15 and name N3) or (resid 15 and name C5))

weight = $pscale end

! G19-T24

!--

146

group

selection=   ((resid 19 and name N1) or (resid 19 and name N3) or

          (resid 19 and name C5) or (resid 24 and name N1) or

          (resid 24 and name N3) or (resid 24 and name C5))

weight = $pscale end

! G16-G19

!---

group

selection=   ((resid 15 and name N1) or (resid 15 and name N3) or

          (resid 15 and name C5) or (resid 19 and name N1) or

          (resid 19 and name N3) or (resid 19 and name C5))

weight = $pscale end

! G9-T24

!---

group

selection=   ((resid 9 and name N1) or (resid 9 and name N3) or

          (resid 9 and name C5) or (resid 24 and name N1) or

          (resid 24 and name N3) or (resid 24 and name C5))

weight = $pscale end

! G8-T16

!---

group

selection=   ((resid 8 and name N1) or (resid 8 and name N3) or

          (resid 8 and name C5) or (resid 16 and name N1) or

          (resid 16 and name N3) or (resid 16 and name C5))

weight = $pscale end

! G18-G25

!---

group

selection=   ((resid 18 and name N1) or (resid 18 and name N3) or

```
        (resid  18 and name C5) or (resid 25 and name N1) or

        (resid 25 and name N3) or (resid 25 and name C5))

weight = $pscale end

! G16-G18

!----------------------------------------------------------------

group

selection=    ((resid  16 and name N1) or (resid  16 and name N3) or

        (resid  16 and name C5) or (resid 18 and name N1) or

        (resid 18 and name N3) or (resid 18 and name C5))

weight = $pscale end

! G25-G8

!----------------------------------------------------------------

group

selection=    ((resid  25 and name N1) or (resid  25 and name N3) or

        (resid  25 and name C5) or (resid 8 and name N1) or

        (resid 8 and name N3) or (resid 8 and name C5))

weight = $pscale end
```

Resulting Model

Figure C.6 Quadruplex F as Modelled in CNS.[79,80,159] A) The bases with hydrogen bonds to form the predicted G-quadruplex. B) An overview with arrows to indicate the directionality of the DNA strand.

Appendix D: Energies From CNS

Table D.1 CNS energies for Quadruplex A

Structure	Energy Total	Energy Bond	Energy Angle	Energy Improper Dihedral	Energy van der Waals	Energy NOE
1	462.191	17.835	61.442	35.190	176.673	21.043
2	958.198	25.196	75.267	60.190	303.778	81.787
3	687.614	23.609	63.176	45.677	234.240	53.840
4	893.582	23.046	59.329	46.953	291.092	77.005
5	586.580	16.827	66.132	34.169	197.047	39.849
6	598.982	17.454	60.745	29.541	192.767	62.231
7	838.161	34.949	65.626	28.345	220.321	55.445
8	677.183	19.240	75.040	49.800	217.091	53.006
9	935.407	28.689	68.351	53.405	299.055	82.096
10	682.634	22.966	56.560	48.298	229.227	51.888
11	616.761	18.668	60.147	24.227	186.686	38.794
12	744.086	27.810	89.524	47.099	281.075	66.864
13	671.029	20.997	66.685	40.364	217.475	43.444
14	658.894	20.955	61.210	44.456	228.298	34.473
15	961.592	22.066	55.401	57.348	270.690	81.344
16	564.623	19.396	70.800	54.097	173.192	24.508
17	941.722	23.875	56.739	32.541	208.577	39.393
18	617.346	24.808	73.488	51.171	223.181	43.420
19	486.130	20.803	65.548	33.431	204.427	24.822
20	873.332	37.883	57.900	24.171	252.600	79.214

Table D.2 CNS energies for Quadruplex B

Structure	Energy Total	Energy Bond	Energy Angle	Energy Improper Dihedral	Energy van der Waals	Energy NOE
1	875.260	28.717	91.826	47.995	255.664	71.866
2	876.906	24.861	77.413	63.731	318.960	92.634
3	807.630	24.232	60.784	37.889	223.507	54.938
4	606.306	14.393	65.457	35.531	202.523	46.586
5	1174.810	34.111	68.572	27.109	257.326	48.158
6	818.002	27.592	56.176	37.629	218.900	54.562
7	792.797	30.411	54.303	21.010	210.959	56.185
8	718.243	29.389	90.264	51.641	254.841	76.034
9	687.719	25.258	74.471	52.601	254.311	61.171
10	651.954	17.660	65.481	32.695	187.895	60.331
11	1009.290	30.818	95.501	51.813	340.090	74.911
12	667.162	26.783	98.453	55.042	249.899	65.460
13	806.544	29.318	101.550	75.431	241.184	37.825
14	755.352	23.244	72.975	33.327	270.830	51.707
15	522.157	18.909	58.696	24.370	189.598	38.137
16	684.329	14.500	52.397	32.480	189.428	46.645
17	1009.110	25.300	69.445	59.253	275.678	86.964
18	605.956	17.640	62.481	38.735	219.620	32.515
19	980.704	23.102	61.393	57.177	263.072	82.718
20	865.787	44.053	115.994	45.886	286.205	72.723

Table D.3 CNS energies for Quadruplex C

Structure	Energy Total	Energy Bond	Energy Angle	Energy Improper Dihedral	Energy van der Waals	Energy NOE
1	383.861	15.883	65.783	16.394	231.704	27.652
2	318.861	12.267	49.956	17.437	169.231	23.604
3	710.260	19.038	65.302	37.983	267.247	57.529
4	340.908	15.236	64.007	13.790	204.221	26.781
5	526.506	20.816	82.350	29.638	237.982	40.948
6	382.077	18.243	76.690	16.676	210.456	28.185
7	357.942	13.956	58.849	17.504	196.166	19.971
8	566.736	15.330	57.169	42.571	196.675	27.017
9	398.802	13.780	58.924	18.703	221.871	39.557
10	430.859	20.001	61.637	36.709	189.663	37.617
11	526.932	12.691	47.379	32.842	170.887	20.122
12	324.406	14.709	55.979	13.728	196.129	23.226
13	350.390	16.479	67.247	14.992	198.604	22.295
14	459.868	15.372	63.217	18.253	224.150	40.888
15	562.159	19.568	59.440	48.302	153.036	30.644
16	899.838	28.088	87.690	47.441	307.505	72.329
17	321.602	13.675	57.421	17.509	164.252	19.950
18	785.744	25.953	94.760	50.919	260.085	58.981
19	308.462	11.536	49.661	16.185	264.448	17.380
20	631.714	21.414	69.566	41.565	158.001	42.782

Table D.4 CNS energies for Quadruplex D

Structure	Energy Total	Energy Bond	Energy Angle	Energy Improper Dihedral	Energy van der Waals	Energy NOE
1	344.652	14.036	59.422	17.189	181.209	22.013
2	456.945	12.552	58.492	31.042	156.727	24.250
3	299.502	13.796	52.655	10.821	188.599	21.949
4	294.341	11.366	46.174	14.238	177.650	22.045
5	395.151	11.090	50.892	27.815	133.604	7.720
6	335.108	13.267	54.409	15.919	185.369	17.343
7	681.814	15.980	64.467	34.594	240.195	39.175
8	777.866	21.667	74.089	56.175	200.814	58.414
9	288.941	10.281	45.601	16.234	143.904	16.813
10	301.050	10.569	44.103	16.933	156.437	17.324
11	441.795	14.779	54.016	18.431	213.639	42.133
12	339.533	11.120	48.966	17.799	181.235	22.536
13	363.970	12.785	56.824	20.274	197.156	17.594
14	633.793	17.900	54.093	42.528	212.755	46.057
15	517.475	21.468	82.384	37.369	231.796	40.146
16	367.355	10.245	50.274	25.993	112.611	8.641
17	325.582	13.784	56.267	12.779	201.533	21.035
18	317.106	12.101	50.743	18.027	167.126	20.798
19	413.417	13.852	56.762	18.803	180.423	35.344
20	864.091	18.376	53.182	20.272	171.574	9.030

Table D.5 CNS energies for Quadruplex E

Structure	Energy Total	Energy Bond	Energy Angle	Energy Improper Dihedral	Energy van der Waals	Energy NOE
1	589.983	17.487	46.556	29.862	237.294	49.659
2	282.393	10.569	42.606	15.899	145.548	18.144
3	534.379	11.086	49.035	32.314	150.659	21.399
4	642.876	16.479	50.935	37.879	237.967	41.518
5	684.944	16.927	54.643	27.198	231.287	53.003
6	277.666	9.109	39.572	16.108	142.944	17.324
7	561.996	17.956	82.690	30.249	220.852	33.749
8	843.993	17.472	60.708	49.179	306.566	73.566
9	414.317	15.573	55.068	18.554	176.249	34.127
10	504.172	14.900	57.318	27.009	190.243	32.184
11	667.866	16.014	60.189	61.779	157.405	33.985
12	378.511	14.732	62.132	20.515	198.399	28.258
13	558.144	22.676	84.281	26.792	256.985	89.078
14	293.191	9.784	40.538	16.611	157.721	17.196
15	713.701	18.203	57.350	41.287	262.496	51.573
16	456.406	16.182	59.061	19.382	215.770	48.387
17	473.403	10.958	46.017	26.638	129.766	14.667
18	676.599	19.506	59.124	34.418	175.299	29.631
19	264.293	8.935	41.331	17.275	125.524	16.734
20	411.640	13.734	48.888	23.813	202.979	54.575

Table D.6 CNS energies for Quadruplex F

Structure	Energy Total	Energy Bond	Energy Angle	Energy Improper Dihedral	Energy van der Waals	Energy NOE
1	336.582	13.415	58.438	18.683	170.912	21.074
2	683.608	17.607	47.270	38.696	216.952	59.832
3	376.931	11.570	48.090	14.575	153.355	28.191
4	369.346	13.893	56.323	20.889	185.698	26.214
5	307.746	9.718	43.912	16.287	159.463	20.214
6	732.673	18.824	42.134	19.352	128.410	17.735
7	596.466	18.670	64.593	53.872	239.616	55.777
8	440.210	17.477	62.883	20.184	195.231	32.475
9	324.840	14.880	57.969	13.401	186.044	30.730
10	272.837	11.673	46.073	10.899	167.477	20.814
11	413.317	17.548	53.324	18.036	223.247	54.422
12	375.631	14.218	48.925	16.659	159.860	35.918
13	382.561	12.880	46.733	19.847	162.646	34.676
14	350.248	16.089	63.853	16.690	198.960	20.722
15	412.421	16.039	53.335	17.817	191.762	44.311
16	303.461	10.639	43.153	17.739	150.269	15.058
17	318.736	11.566	48.660	17.774	157.148	13.955
18	306.733	10.251	40.743	17.251	162.260	19.783
19	640.167	17.453	53.187	33.717	222.939	43.125
20	280.337	13.597	51.832	11.404	165.132	21.612