PREVENTION OF G-QUADRUPLEX FOLDING BY LIGAND-MEDIATED

STABILIZATION OF THE G-TRIPLEX FOLDING INTERMEDIATE

by

Harrison Bracey, B.S.

A thesis submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Science
with a Major in Biochemistry
December 2018

Committee Members:

Sean Kerwin, Chair

Karen Lewis

Wendi David

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

G-quadruplex DNA is an extensively studied non-canonical DNA structure formed in guanine rich areas of the human genome. It is proposed that G-quadruplex structures play a role in telomere assembly, replication, and transcription. Experimental evidence has shown that an intermediate in the G-quadruplex folding pathway may exist as G-triplex DNA. The structure of G-triplex DNA is similar to that of G-quadruplex DNA. However, the G-triplex structure is unique in that it has a major groove. The goals of this project are to characterize different G-triplex structure topologies and to find small molecular weight ligands that can selectively bind to the major groove and stabilize the G-triplex intermediates to prevent G-quadruplex folding. To characterize G-triplex topology and stabilization, CD spectra and CD-derived melting temperatures (Tm) were compared to UV thermal difference spectra (TDS) and UV derived Tm of a collection of model G-triplex forming oligonucleotides. Potential ligands were identified by an *in silico* screen between truncated thrombin binding aptamer (T1) and 1.2 million compounds using molecular docking software, Autodock Vina, and a database of commercially available ligands, ZINC12. To determine if potential ligands bind to and stabilize TBA13, UV-vis spectroscopy melting assay and a FRET bound Tm assay employing a dual-labeled T1 oligonucleotide was performed.

I. INTRODUCTION

Non-canonical DNA

Genomic deoxyribonucleic acid (DNA) is a double stranded duplex connected by Watson-Crick base pairing, where adenine binds to thymine and guanine binds to cytosine in equal amounts via hydrogen bonding (31). Since the discovery of the Watson-Crick duplex structure of B-DNA, there have also been other discoveries of irregular or non-canonical DNA structures. There are more than 10 types of non-canonical DNA (30). These non-canonical DNA forms include cruciform, hairpin, H-DNA, G-DNA, and Z-DNA structures, among others (3, 12, 16, 19). Non-canonical DNA structures have been proposed to play a role in DNA transcription and replication, the regulation of various oncogenes, and in genomic instability leading to chromosomal breakage and other genetic defects (2, 3, 12, 16, 28, 34). The distribution of non-B-DNA forming sequences in the human genome is not random (29). In fact, many of these sequences are found in "hot-spot" areas of genomic instability (30, 34, 35).

G-Quadruplex DNA Structure

G-quadruplex DNA is an extensively studied non-canonical DNA structure formed by guanine-rich DNA sequences. The G-quadruplex is composed of two or more G-tetrad units, which are formed by four guanine bases in plane with each other and connected through Hoogsteen base pairing. G-tetrads stack upon one another (12, 16, 25). These structures have also been shown to be stabilized by monovalent cations such as K^+ and Na^+ (12, 25). G-quadruplexes can adopt many different topologies depending on their specific sequence and local conditions (27). These different topologies can be classified

1

into different groups, such as parallel, antiparallel or hybrids of both (4). G-quadruplexes can be intermolecular or intramolecular, but it is assumed that biologically relevant G-quadruplexes are intramolecular (3, 27). Intramolecular G-quadruplex structures are predicted to form at specific G-rich regions *in vivo*.
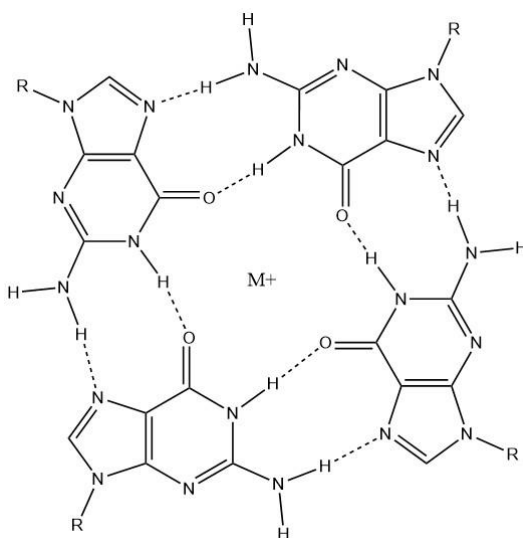
Figure 1. Example of a G-tetrad subunit exhibiting Hoogsteen base pairing between guanine bases and stabilized by a monovalent cation.
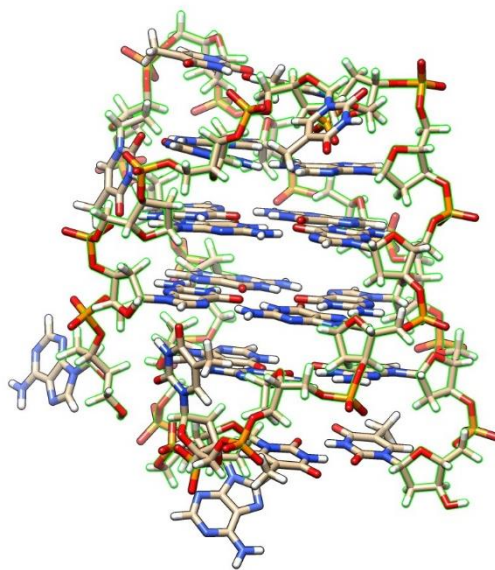
Figure 2. NMR structure of an intramolecular telomeric G-quadruplex structure.

Effects of G-Quadruplex DNA

The G-quadruplex forming sequence motif, $G_{2+}N_{1-7}G_{2+}N_{1-7}G_{2+}N_{1-7}G_{2+}$, is present in over 370,000 sites in the human genome (2, 3, 12, 25) G-quadruplex sequence motifs were found to be more evolutionarily conserved than expected by chance which may be due to the biological roles they play in telomere assembly, replication, and transcription (3, 27, 29). G-quadruplex forming sequences have also been found within 500 base pairs of mitotic and meiotic double strand break (DSB) sites and are abundant in "hot-spot" areas of the human genome associated with genomic instability (29, 35).

G-Quadruplex in the Telomere

Human telomerase is an enzyme that lengthens the G-strand of the telomere and is inactive in most somatic cells. Human telomerase is overexpressed in 80% of all cancers and is thought to promote the lifespan of malignant cells (25). G-quadruplex DNA has been shown to play an important role in telomeric stability (29). Intramolecular anti-parallel G-quadruplex structures have been shown to inhibit telomerase activity whereas intermolecular parallel G-quadruplex structures do not (21, 33). Stabilization of telomeric G-quadruplex by small molecular weight ligands has been shown to inhibit telomerase activity *in vitro* (14, 23). However, further research is needed to determine if G-quadruplex stabilizing ligands are effective *in vivo* and if they have any deleterious effects on non-telomeric G-quadruplexes (3).

G-Quadruplex in Replication

Opportunities for G-quadruplex formation can arise during DNA replication. The discontinuous nature of the lagging strand leaves it transiently single-stranded which

provides opportunities for G-quadruplex formation during DNA replication. G-quadruplexes that are formed during DNA replication or are pre-existing must first be unfolded for DNA replication to be completed (3). DNA helicases have been shown to have G-quadruplex unwinding activity *in vitro* (3). Loss of G-quadruplex unwinding by human helicases is associated with human diseases caused by genomic instability (9, 20, 24). FANCJ and Pif1 are helicases that can unwind G-quadruplex DNA (18, 22). Evidence that loss that G-quadruplex unwinding is associated with human disease has been observed when cell lines of human patients with Fanconi anemia carrying FANCJ mutations showed deletions that overlap G-rich regions with the potential to form G-quadruplex structures (18). It has been observed that in the absence of Pif1, DNA replication slows and DSBs occur at many of the G-quadruplex sequence motifs (22). It has also been observed in the absence of Pif1, when G-quadruplex sequence motifs are present in the template of the leading strand during replication, G-quadruplex structure formation can occur. This can result in deletions and rearrangement events (29).

G-Quadruplex in Transcription and Regulation

High concentrations of G-quadruplex sequence motifs are found near promoter regions in the human genome. This suggests a potential function of G-quadruplexes in gene regulation at the transcriptional level (10). Promoters of human oncogenes and regulatory genes have been shown to contain more G-quadruplex motifs than the promoters of housekeeping and tumor suppressor genes (8). A well-studied system for the role of G-quadruplex structures in transcription involves the *c-MYC* locus. Increased levels of *c-MYC* expression promotes tumorigenesis and is observed in 80% of human cancer cells. Nuclease hypersensitive element III (NHE III) is found upstream of the *c-*

MYC promoter and controls over 85-90% of *c-MYC* transcription. This element contains

a G-quadruplex sequence motif and when mutated so as to abrogate G-quadruplex

formation, transcription of *c-MYC* increases (26).

G-Triplex DNA

It has been hypothesized and experimental evidence has shown evidence that an

intermediate in the G-quadruplex folding pathway may exist as G-triplex DNA (5, 13,

16). The structure of the G-triplex is similar to the G-quadruplex, as it consists of G-triad

planes stabilized by Hoogsteen base pairing and is able to coordinate a metal cation at the

center (16, 27) The topology of G-triplex folding intermediates is similar to that of the G-

quadruplex product (27). However, unique to the G-triplex structure is a major groove,

into which a fourth strand of DNA can bind to form a more stable G-quadruplex

structure.
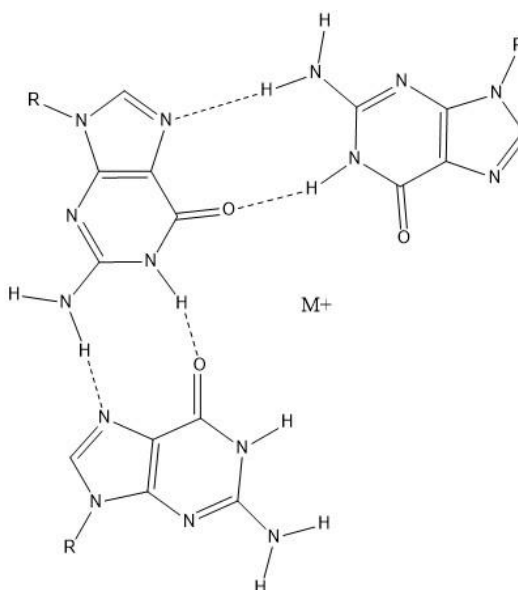
Figure 3. Example of a G-triad subunit with Hoogsteen base pairing between guanine bases and stabilized by a metal cation.
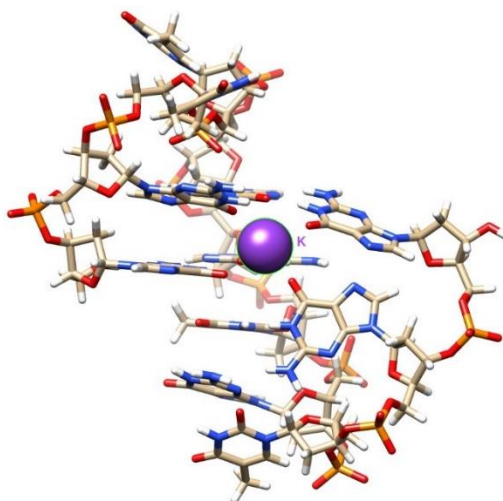
Figure 4. NMR structure of the G-triplex TBA13 stabilized with a potassium ion at the center.

Studying the G-triplex Intermediate

There are several methods currently being utilized to study the structural features of G-quadruplex. Such methods in use are atomic force microscopy (AFM), circular dichroism (CD), X-ray crystallography, Förster resonance energy transfer (FRET), molecular tweezers, nuclear magnetic resonance (NMR), surface plasmon resonance (SPR), and UV-vis spectroscopy (4, 13, 15, 19). Using these methods researchers have been able to identify the presence of G-quadruplex DNA as well as structural features. The topology of G-quadruplex has usually been assigned by using advanced NMR, X-ray crystallography, or through CD (4, 13).

G-triplex DNA structure has only been studied using a handful of methods such as NMR and CD, the most common of which is CD (5, 13). Circular dichroism is a technique in which circularly polarized light that can be used to identify the secondary

structure of different biological molecules such as proteins and DNA (32). This method

has been used to assign the topology of G-quadruplex DNA as well as G-triplex (13).
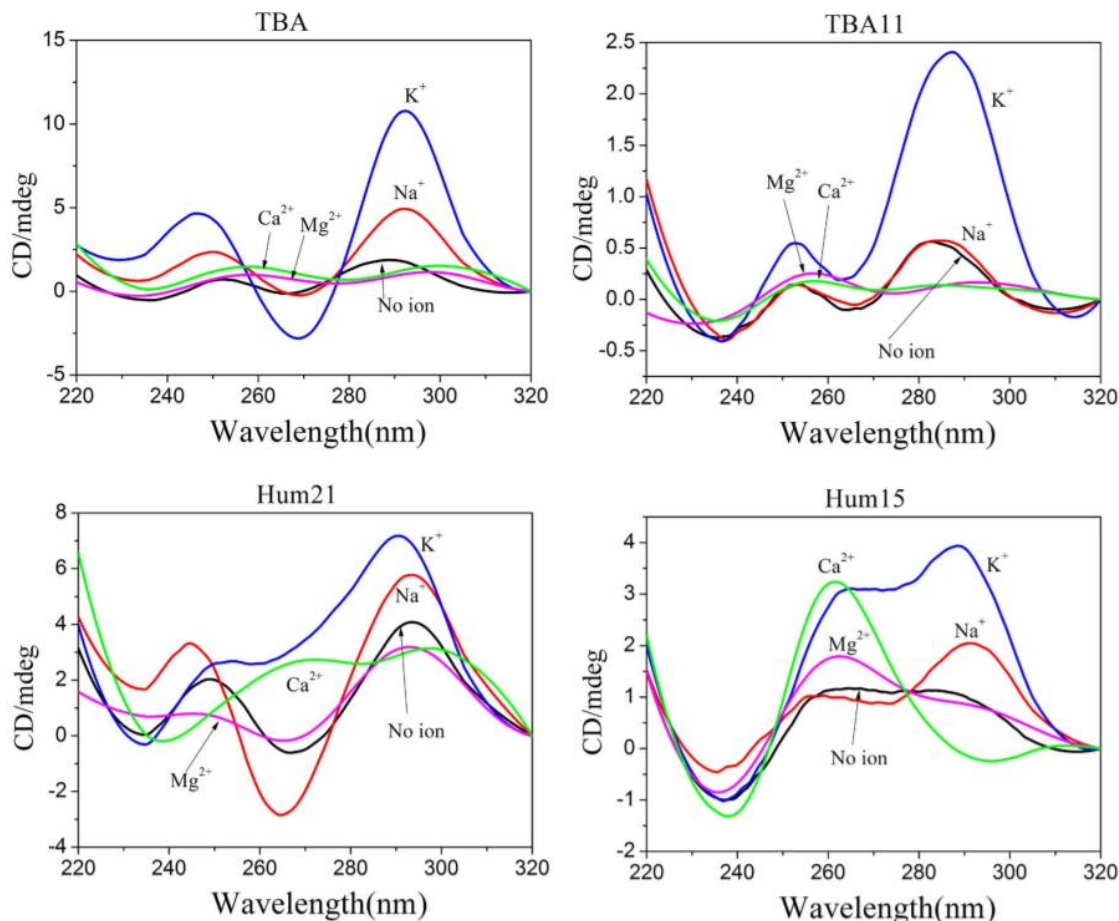
Figure 5. CD spectra of G-quadruplex and G-triplex DNA. TBA and Hum21 are examples of G-quadruplex DNA while TBA11 and Hum15 are examples of G-triplex DNA (13).

As seen in Figure 5, the CD spectra of some G-quadruplex (TBA) and the

corresponding truncated sequence G-triplex (TBA11) may look very similar or they may

look very different (Hum21 vs Hum15). An anti-parallel conformation can be seen with

the G-quadruplex TBA in the presence of Na^+ or K^+ ions. This defined as the CD spectra

showing a strong positive peak at 295 nm and a negative peak at 270 nm. The G-triplex

TBA11 in the presence of no ions, Na^+, or K^+ showed similar spectra with positive peaks around 290 nm and negative peaks around 265 nm. In both cases, Ca^{2+} and Mg^{2+} do not appear to be promoting the formation of stable secondary structures. The G-quadruplex Hum21 shows anti-parallel topologies for no ions or Na^+ ions and appears to be a hybrid of anti-parallel and parallel in all other ions tested. However, when looking at the G-triplex, Hum15, a parallel conformation can be seen with Ca^{2+} ions, defined by a positive peak at 260 nm and a negative peak at 240 nm. All other ions tested appear to show a hybrid topology (13).

Using UV-vis spectroscopy researchers have identified different forms of non-canonical DNA structure, including G-quadruplex DNA by using techniques in thermal difference spectra (TDS) (19). This technique reveals structural data about the sample by taking advantage of the hypochromic effect. Single strand DNA has a higher absorbance than double strand DNA or when DNA is forming some other structure. Therefore, by subtracting the spectrum of un-melted DNA from the spectrum of melted DNA, a thermal difference spectrum is obtained. This thermal difference spectrum can be unique to specific non-canonical DNA structures. This technique has been used to identify other non-canonical DNA structures such as Hoogsteen DNA duplexes, pyrimidine triplexes, and even G-quadruplex DNA (19). However, it has not been used to obtain structural data on G-triplex intermediate DNA.

TBA13 alongside a Förster resonance energy transfer (FRET) is a physical process in which the excited state energy of one fluorophore, the donor, is able to be transferred to an adjacent fluorophore, the accepter, in the ground state. This is only possible when the two fluorophores are within range of each other (6). FRET has been

used in molecular beacons of duplex DNA, reporter molecules, and in DNA structure studies (2, 6). Although, FRET has been used extensively in G-quadruplex DNA folding studies, it has not been used to obtain structural data on G-triplex intermediate DNA.

Goals

G-quadruplex DNA structure and function is well-studied and has been the target for nanotechnology research, oncology research, as well as medicinal research (14, 17, 23). There is however, very little known about the structure of G-triplex DNA and its effects *in vivo*. G-triplex DNA is an intermediate to the G-quadruplex DNA structure. G-triplex intermediate DNA has a major groove present where a fourth strand of DNA can fold to become G-quadruplex DNA. If a small molecule binds within this major groove it may stabilize the G-triplex while preventing the folding of a fourth strand of DNA and perturb G-quadruplex folding. *Thus, the hypothesis of this work is that selective binding by a small molecular weight ligand to the major groove will stabilize the G-triplex intermediate and prevent G-quadruplex folding.* If this hypothesis is correct, the results of this work may provide an important tool for further studying G-quadruplex structures in biological roles and may also provide an avenue of treatment for genetic diseases associated with G-quadruplexes. *The goals of this project are to characterize different G-triplex structure topologies and to find small molecular weight ligands that can selectively bind to the major groove and stabilize the G-triplex intermediates and prevent G-quadruplex folding.* The characterization of G-triplex structure topologies is important since G-triplex structure, as with the G-quadruplex, may be affected by environmental conditions and strand sequence composition. Studies of G-triplex topology and stabilization was be performed by comparing circular dichroism (CD) spectra and CD-

derived melting temperatures (Tm) with UV-vis spectroscopy thermal difference spectra (TDS) and UV-vis-derived Tm of a collection of model G-triplex forming oligonucleotides. This collection contains various oligonucleotides that present the G-triplex sequence motif, $G_{2-5}N_{1-7}G_{2-5}N_{1-7}G_{2-5}$, as well as truncated thrombin binding aptamer (TBA13) and human telomere sequences. Potential small molecular weight ligands were identified by an *in silico* virtual screen using molecular docking software, Autodock Vina (28). The only known NMR structure of a G-triplex is of T1 (5). Therefore, T1 acted as the model receptor in the virtual screen. Potential G-triplex stabilizing ligands were identified by searching a large sub-set (1.2 million compounds) from a database of commercially available ligands, ZINC12 (11, 12). A subset of the ligands predicted to have the highest affinity for TBA13 was selected to experimentally determine their ability to bind to the G-triplex and prevent G-quadruplex folding. UV-spectroscopy Tm assay with unlabeled TBA13 alongside a FRET bound Tm assay employing a dual-labelled T1 oligonucleotide was used to determine if the potential ligands bound to and stabilized the TBA13 G-triplex.

II. MATERIALS AND METHODS

Materials

The DNA oligonucleotides used in this work are listed in Table 1. All DNA oligonucleotides were synthesized and obtained from Integrated DNA technologies Inc (Coralville, Iowa, USA). Ultrapure water was purified through a Barnstead MicroPure ST by Thermo Scientific (Waltham, Massachusetts, USA). Cacodylic acid and KCl were obtained from Sigma-Aldrich (Saint Lewis, Missouri, USA). NaCl was obtained from Fisher Scientific (Hampton, New Hampshire, USA). $CaCl_2$ was obtained from VWR (Radnor, Pennsylvania, USA). $MgCl_2$ was obtained from MilliporeSigma (Burlington, Massachusetts, USA). All samples tested were made to a solution of either 15 µM or 30 µM of DNA in a 10 mM lithium cacodylate buffer at a pH of 7.3 with no added ions or 100 mM ion solution (NaCl, KCl, $CaCl_2$, or $MgCl_2$). UV-assays (TDS, Tm, and Ligand Assays) were performed on a Jasco V-730 Spectrophotometer (Victoria, British Columbia, Canada). Ligand 42783800 (also called T-75) was synthesized in house. Ligand 5,6-diphenyl-3-(2-pyridyl)-1,2,4-triazine-4,4 disulfonic acid monosodium salt hydrate (S1) was obtained from Alfa Aesar (Haverhill, Massachusetts, USA). Ligand 5,5',6,6'-tetrahydroxy-3,3,3',3'-tetramethyl-1,1'-spirobliindane (S2) was obtained from TCI Chemicals (Portland, Oregon, USA). Ligand 4-methylumbelliferyl-B-D-glucuronide hydrate (S3) was obtained from Chem-Impex (Wood Dale, Illinois, USA). Ligand tetracycline hydrochloride (TeT) was obtained from VWR (Radnor, Pennsylvania, USA).

Table 1. Table of target G-triplex forming oligonucleotides. and their respective sequence used during TDS and Tm studies. T1 was only used in FRET assays.

Oligo	Sequence	Oligo	Sequence
TBA13	TGGTTGGTGTGGT	G5T4	TGGGGGTTTTGGGGGTTTTGGGGT
TBA13-Rev	TGGTGTGGTTGGT	G4T4	TGGGGTTTTGGGGTTTTGGGGT
Htel15	TGGGTTAGGGTTAGGGT	G3T4	TGGGTTTTGGGTTTTGGGT
Htel15-Rev	TGGGATTGGGATTGGGT	G3T3	TGGGTTTGGGTTTGGGT
Htel15-A	AGGGTTAGGGTTAGGGT	G3T2	TGGGTTGGGTTGGGT
Htel15-A-Rev	TGGGATTGGGATTGGGA	G3T1	TGGGTGGGTGGGT
Htel15-delT	TGGGTTAGGGTTAGGG_	G2T4	TGGTTTTGGTTTTGGT
T1	FAM-TGGTTGGTGTGG-TAM		

Thermal Difference Spectra and Temperature Melts

TDS were performed by first heated the samples to 95°C in a water bath then allowing the bath to slowly cool overnight. Once cooled to room temperature the samples were transferred to a 4°C fridge for 2-3 hours. The samples were then placed in the spectrophotometer and allowed to equilibrate for 10 min at 5°C. The UV absorbance spectra was measured from 220 nm to 350 nm in 0.5 nm intervals at a scan speed of 100 nm/min for 5°C and 95°C. To get the TDS spectra, the 5°C spectra was subtracted from the 95°C spectra, baseline corrected, and peak normalized.

Tm was performed between the 5°C spectra and 95°C spectra of the TDS. The temperature was controlled at a ramp rate of 0.5 °C/min. Measurements were taken at 260 nm and 295 nm at every 0.5°C. Once at 95°C the samples were allowed to equilibrate for 10 min before taking the 95°C spectra for the TDS. Tm was determined by non-linear

regression of the 295 nm wavelength by Bolzman Sigmoidal curve fitting function from Graphpad Prism software (San Diego, California, USA).

***In Silico* Screen**

In silico screen of a subset of the ZINC15 database (San Francisco, CA, USA) for potential TBA13 binding ligands was performed using Autodock Vina 1.1.2 (Scripps Research Institute, CA, USA) on the 32 Node Sun Grid (Texas State University, San Marcos, TX, USA). The structure of the G-triplex was obtained from the ^{1}H NMR-based structure of Cerofolini et al. (5), PDB code = 2MKO (1). This structure reveals that this G-triplex intermediate has a major groove present where the fourth strand would be present in the corresponding G-quadruplex. This groove is an ideal target for a compound to bind in order to prevent the fourth strand of the G-quadruplex from folding. Therefore, this major groove was the selected target for the *in silico* screen (Figure 6). In addition to defining this box region for docking, the exhaustiveness parameters for Autodock Vina was set at 16. The ligand database for the *in silico* screen consisted of 1.2×10^6 compounds selected from ZINC15 using the following criteria: Commercially available, MW = 200-350 Da, predicted $Log_{(p)} = 2–3.5$, and no more than one negative charge predicted at pH 7.4. These ligands were grouped into 12 sets based upon the tranches in the ZINC15 database. Within each set, the ligands were docked with the T1 G-triplex and the top 3 scoring binding modes for each ligand, based on predicted binding affinity were saved. The top three ligands of each set were selected and these top hits were ranked based on their top predicted binding affinity to obtain a list of the 23 overall best ligands.
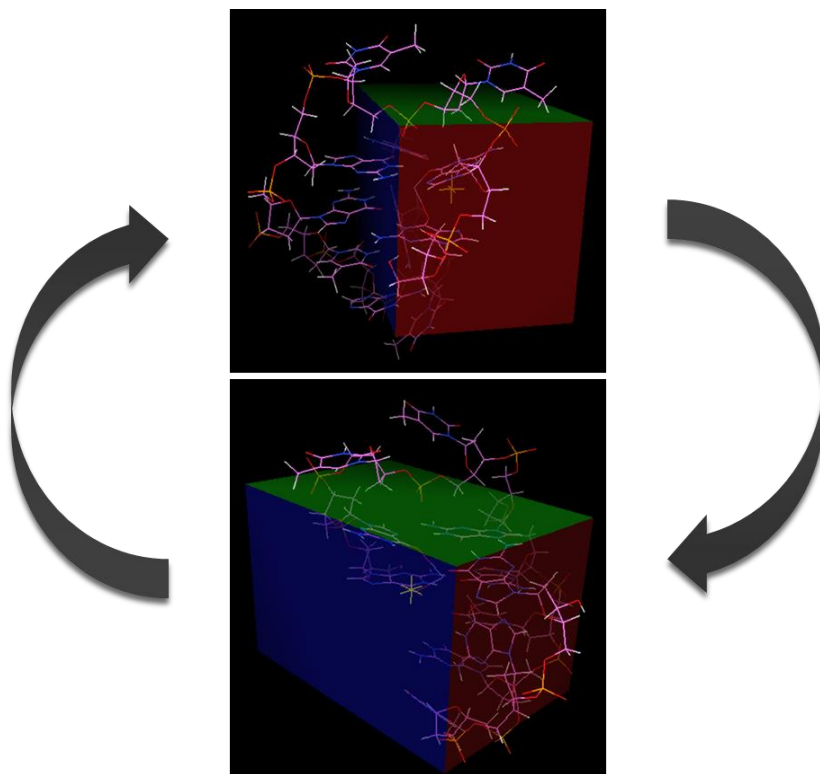
Figure 6. Targeting the G-triplex intermediate major groove *in silico*. Visual representation of three-dimensional targeting of the T1 G-triplex major groove. Arrows represent rotation of G-triplex and three-dimensional targeting of the major groove.

Ligand Stability Assays

Ligand induced stability was tested by Tm of TBA13 or F-T1-T and potential ligands using UV-spectroscopy and FRET. For UV-spectroscopy samples were made of 5 µM of TBA13 in 100 mM KCl with either no ligand or 10 µM added before or after annealing. The Tm was performed by measuring 260 nm and 295 nm wavelengths from 5 °C to 95 °C at a ramp rate of 0.5 °C/min. FRET samples consisted of 100 nM of FAM-TAMRA labeled T1 (F-T1-T), with FAM at the 5' end and TAMRA at the 3' end, in 100 mM KCl with either no ligand or 100 µM of target ligand added after annealing. The FAM "donor" signal was excited at 495 nm and detected at the emission wavelength of

520 nm. TAMRA "accepter" was excited at 544 nm and detected at the emission

wavelength of 576 nm (Figure 7). The Tm was determined by non-linear regression of

the FAM emission wavelength of 520 nm by Bolzman Sigmoidal curve fitting function

from Graphpad Prism software.
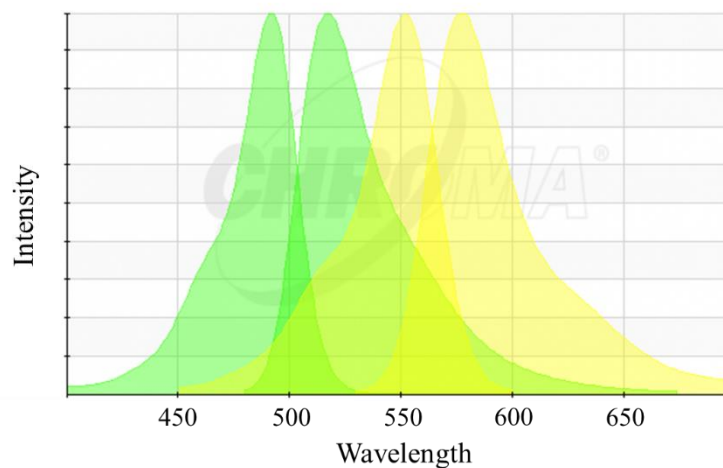
Figure 7. Chroma-graph of FRET system. Green is FAM excitation and emission
spectrum. Yellow is TAMRA excitation and emission spectrum.

III. RESULTS

Structure and Stability

Selection of DNA oligonucleotides was based on well-studied G-quadruplex forming sequences found *in vivo* (Htel-series) and experimental sequences which vary in T-loop length and G-tract length (GNTN series). TBA13 was chosen due to the well-studied G-quadruplex, Thrombin Binding Aptamer. In addition to the only NMR solution structure of a G-triplex is of T1, a truncated version of TBA in which one G-tract is missing. Reverse sequences were chosen because of the reversal of T-loop length in which the long and short loops are switched in position.

DNA annealing was crucial in order to get formation of the most thermodynamically stable G-triplex structure. A longer annealing process is thought to give the G-triplex more time to find its most stable structure. This is in contrast to a fast annealing method where the kinetic G-triplex structure may be the product and not as stable. In order to accomplish this once the solution samples were mixed they were placed in a hot water bath at 95°C. The bath was allowed to slowly cool overnight to room temperature. After cooling the samples were moved to a 4°C fridge for 2-3 hours. This allowed the samples to slowly find their most stable structure.

Table 2. Table of average melting temperature of nucleotides used. Standard deviation of temperature melt iterations is in superscript. Asterisks represent there was no standard deviation due to only having one melting obtainable temperature. Questions marks indicate melting temperature was indeterminate. All temperatures are in °C.

Average Melting Temperature of Oligonucleotides					
DNA	Li^+	Na^+	K^+	Ca^{2+}	Mg^{2+}
TBA13 15 // 30 μM	$8.4^{+/-1}$ // 8.6^*	$12.6^{+/-0.2}$ // $9.1^{+/-3}$	$13.1^{+/-1}$ // $12.8^{+/-0.4}$	$14.3^{+/-0.5}$ // $15.2^{+/-1}$	$13.9^{+/- 0.3}$ // $18.4^{+/- 0.3}$

Table 2. Continued.

TBA13-Rev 15 // 30 µM	$12.0^{+/-1.0}$ // $12.8^{+/-4}$	$16.1^{+/-0.6}$ // $15.5^{+/-\ 0.6}$	$15.2^{+/-0.2}$ // $15.0^{+/-0.8}$	$19.2^{+/-0.5}$ // $19.9^{+/-0.9}$	$17.4^{+/-0.5}$ // $17.5^{+/-2}$
Htel15 15 // 30 µM	$16.0^{+/-2}$ // $13.1^{+/-2}$	$28.0^{+/-1.0}$ // $28.0^{+/-0.4}$	$43.2^{+/-1.0}$ // $44.3^{+/-1.0}$	$43.8^{+/-0.5}$ // $45.5^{+/-\ 17}$	$24.5^{+/-3}$ // $23.3^{+/-4}$
Htel15-Rev 15 // 30 µM	$8.3^{+/-0.8}$ // ???	$12.9^{+/-0.1}$ // $12.6^{+/-0.5}$	$38.2^{+/-1}$ // $35.6^{+/-2}$	$57.0^{+/-0.3}$ // $57.4^{+/-0.3}$	$13.3^{+/-0.5}$ // $12.7^{+/-3}$
Htel15-A 15 // 30 µM	$15.0^{+/-6}$ // $21.3^{+/-10}$	$29.0^{+/-0.6}$ // $30.6^{+/-2}$	$40.7^{+/-1}$ // $42.0^{+/-0.7}$	$66.2^{+/-0.6}$ // $66.3^{+/-0.7}$	$34.2^{+/-12}$ // $54.9^{+/-36}$
Htel15-Rev-A 15 // 30 µM	$8.2^{+/-0.3}$ // $17.9^{+/-8}$	$16.3^{+/-0.8}$ // $16.7^{+/-8}$	$40.1^{+/-6}$ // $42.4^{+/-2}$	$59.6^{+/-0.2}$ // $62.0^{+/-4}$	$18.0^{+/-3}$ // $23.2^{+/-19}$
Htel15d-elT 15 // 30 µM	$11.0^{+/-6}$ // $13.6^{+/-0.1}$	$26.6^{+/-0.3}$ // $27.7^{+/-1}$	$44.3^{+/-2}$ // $44.4^{+/-0.8}$	$46.8^{+/-0.1}$ // $48.2^{+/-3}$	$17.3^{+/-5}$ // $28.0^{+/-13}$
G5T4 15 // 30 µM	??? // $16.6^{+/-2}$	55.8^{*} // $64.8^{+/-14}$	76.6^{*} // $78.5^{+/-4}$	80.6^{*} // $80.1^{+/-2.3}$	81.0^{*} // $81.8^{+/-2}$
G4T4 15 // 30 µM	$9.1^{+/-3}$ // $12.4^{+/-2}$	$47.6^{+/-0.5}$ // $48.9^{+/-0.0}$	$82.6^{+/-0.5}$ // $81.3^{+/-0.5}$	$61.6^{+/-0.8}$ // $61.3^{+/-1}$	$47.3^{+/-3}$ // $51.4^{+/-0.5}$
G3T4 15 // 30 µM	$9.6^{+/-0.1}$ // $9.1^{+/-4}$	$18.2^{+/-0.4}$ // $18.9^{+/-2}$	$34.6^{+/-12}$ // $38.6^{+/-4}$	$34.4^{+/-0.4}$ // $40.7^{+/-5}$	$25.0^{+/-3}$ // 48.8^{*}
G3T3 15 // 30 µM	$12.6^{+/-4}$ // 9.67^{*}	$17.4^{+/-1}$ // 17.4^{*}	$41.2^{+/-5}$ // 38.4^{*}	$47.1^{+/-5}$ // 42.3^{*}	$27.7^{+/-6}$ // 39.7^{*}
G3T2 15 // 30 µM	$15.5^{+/-0.2}$ // $13.5^{+/-6}$	$17.3^{+/-0.2}$ // $18.0^{+/-0.8}$	$50.9^{+/-0.1}$ // $53.3^{+/-0.1}$	$49.3^{+/-0.2}$ // $49.1^{+/-0.6}$	$27.4^{+/-2}$ // 24.8^{*}
G3T1 15 // 30 µM	$25.3^{+/-0.5}$ // $26.8^{+/-2}$	$28.6^{+/-0.6}$ // $39.4^{+/-3}$	$64.6^{+/-2}$ // $69.3^{+/-0.1}$	$57.1^{+/-1}$ // $64.4^{+/-1}$	48.5^{*} // $41.4^{+/-16}$
G2T4 15 // 30 µM	$8.0^{+/-1}$ // ???	12.3^{*} // ???	$12.0^{+/-2}$ // 11.8^{*}	12.0^{*} // ???	??? // 8.7^{*}

The primary goal of the temperature melt assay is to determine the effects of sequence and buffer conditions on thermal stability of the G-triplexes. For the Tm assay samples were transferred from the fridge after the annealing process to the UV-spectrometer. It is crucial for the UV-spectrometer to be pre-cooled to 5°C before this transfer, so as to not heat up the samples prematurely. The Tm assay started at 5°C and ramped up to 95°C at a rate of 0.5°C /minute. Data was collected at 0.5°C intervals at the wavelengths of 260 nm and 295 nm. In obtaining the melting temperature the 295 nm

wavelength was used because it was generally more cooperative then the 260 nm wavelength. Melting temperature calculation was performed by non-linear regression from the Bolzman sigmoidal curve fit from Graphpad Prism. Tm assays were performed in triplicates however, as of writing this work most of the DNA melting temperatures are in doublets. Standard deviation of the triplet/doublets was used to estimate experimental error in the melting temperatures.

Reviewing the melting temperatures in Table 2 there are a few trends that can be observed. It is best to look at these trends first by reviewing the naturally derived sequences (TBA13-Htel-series) and comparing buffer conditions, sequence effects from reversal of sequence, and addition of purine or pyrimidine to the 5' or 3' end of the oligonucleotide. Then reviewing a systematic series of varying G-tracts and T-loops.
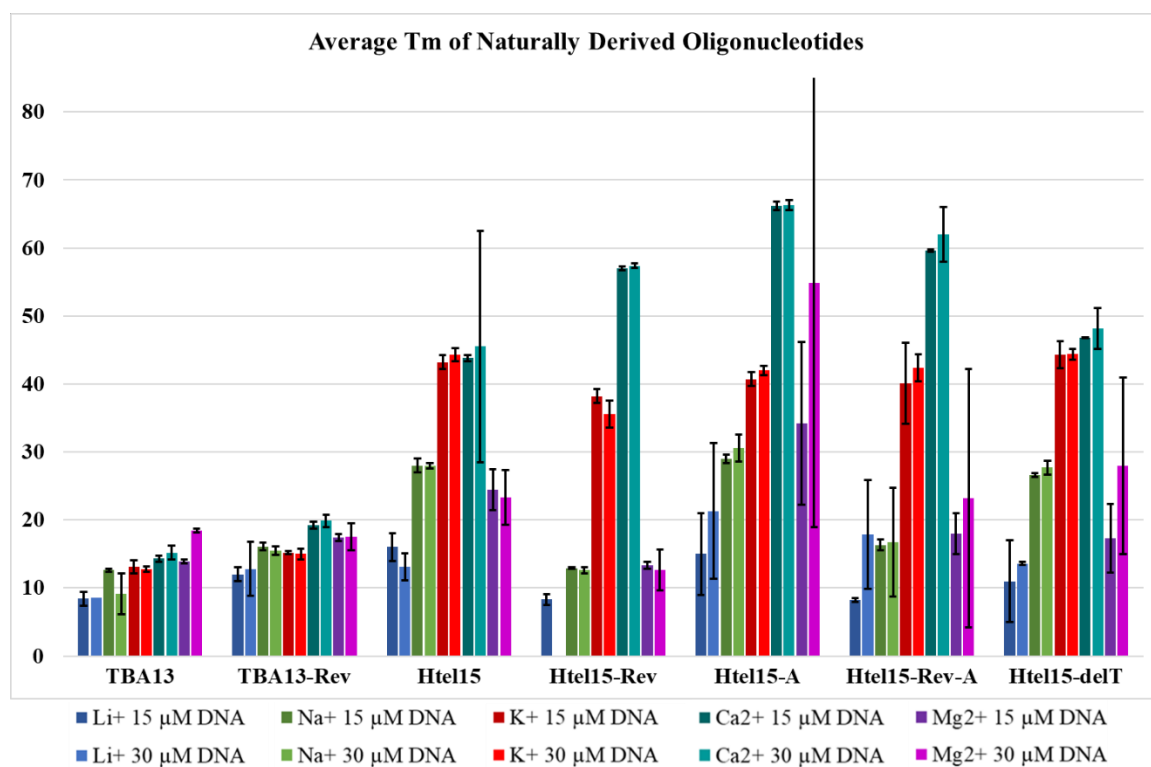

Figure 8. Bar graph representing the average Tm of naturally derived G-triplex forming oligonucleotides. Error bars represent standard deviation.

When looking at the naturally derived sequences TBA13, TBA13-Rev, Htel15, Htel15-Rev, Htel15-A, Htel15-Rev-A and Htel15-delT (Figure 8), it appears that there is an order to melting temperature depending on what cation is present. This is in the order of $Ca^{2+} > K^+ > Mg^{2+} \geq Na^+ > Li^+$. It is also notable that replicate melts conducted in Mg^{2+}, Na^+, and Li^+ tended to yield more disparate Tm values, which is reflected in the higher standard deviations reported for these conditions compared to melts performed in K^+ and Ca^{2+}. This is especially noticeable at the higher DNA concentration in the presence of Mg^{2+}.

Within this series of oligonucleotides, the strand composition also affects thermal stability. One of these trends is that the Htel-derived oligonucleotides, which each have G-tracts of three guanines have higher melting temperatures than the TBA-derived oligonucleotides, with only two guanines per G-tract (GGG > GG). The increase from a G-tract of two guanines to three seems to double the overall melting temperature when Ca^{2+} and K^+ are present (TBA13 vs Htel15). Additionally, the nature of the terminal residues also impacts the melting temperature of these oligonucleotides. Switching the first thymine in Htel15 to adenine increased the melting temperature in the presence of Ca^{2+} by 20°C (Htel15 vs Htel15-A). This effect can also be seen with Mg^{2+} however, the standard deviation for Htel15-A in Mg^{2+} is large. Switching the last thymine in Htel15-Rev to adenine there is an increase in the melting temperature in the presence of Na^+ (Htel15-Rev vs Htel15-A).

Strand direction also appears to have an effect of the melting temperature. The reverse of TBA13 appears to have an increased melting temperature in the presence of

19

every cation (TBA13 vs TBA13-Rev). This is most noticeable with Ca^{2+} and Na^+. The reverse of Htel15 and Htel15-A appears to lower the melting temperature in the presence of almost every cation. In the presence of Na^+ and Mg^{2+} the melting temperature is almost half of the non-reverse sequence (Htel15 and Htel15-A vs Htel15-Rev and Htel15-Rev-A). Htel15-Rev however, has an increase in melting temperature in the presence of Ca^{2+} when compared to Htel15.

Concentration of oligonucleotide mostly does not have an effect on the melting temperature. However, differences can be seen in the presence of Mg^{2+}. TBA13 has a 4°C increase in melting temperature when going from 15 to 30 µM of DNA in the presence of Mg^{2+}. Htel15-A, Htel15-Rev-A, and Htel15-delT have significant increases in melting temperature when going from 15 to 30 µM of DNA in the presence of Mg^{2+}. There is however, a large amount of standard deviation of Htel15-A, Htel15-Rev-A, and Htel15-delT.
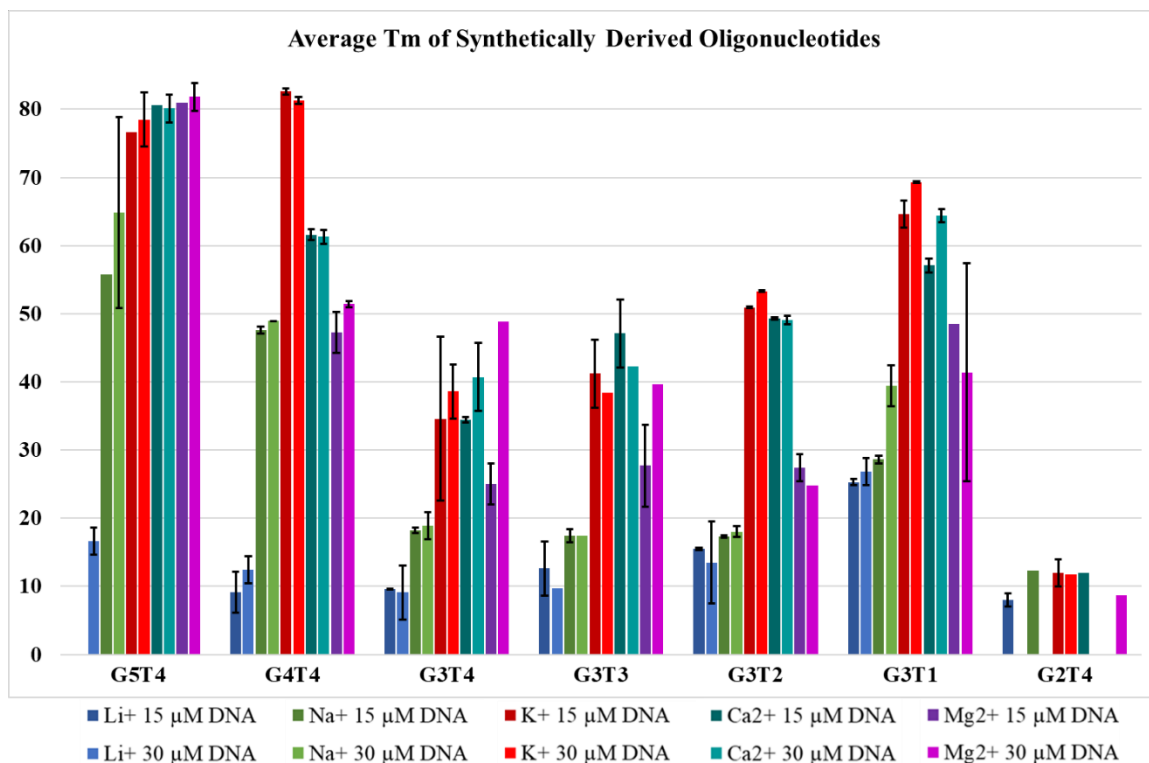
Figure 9. Bar graph representing average Tm of synthetically derived G-triplex forming oligonucleotides. Error bars represent standard deviation.

Focusing on the synthetically derived sequences, G5T4, G4T4, G3T4, G3T3, G3T2, G3T1, and G2T4 (Figure 9), there are few things that can be observed. As seen before with TBA13 and the Htel15 series, more guanines in a G-tract correlate with higher overall melting temperatures when the cations Na^+, K^+, Ca^{2+}, or Mg^{2+} are present. When looking at the T-loops, the shorter the T-loop is the higher the overall melting temperatures are (G3T4 vs G3T3 vs G3T2). The only exception to this is in the presence of Na^+ and Mg^{2+}. With a T-loop of 3 or higher the melting temperatures in the presence of different cations fall in the order of $Ca^{2+} > K^+ > Mg^{2+} > Na^+$. With a T-loop of equal to or less than 2, melting temperatures in the presence of K^+ tend to be more stable than Ca^{2+}, moving the order to $K^+ > Ca^{2+} > Mg^{2+} > Na^+$.
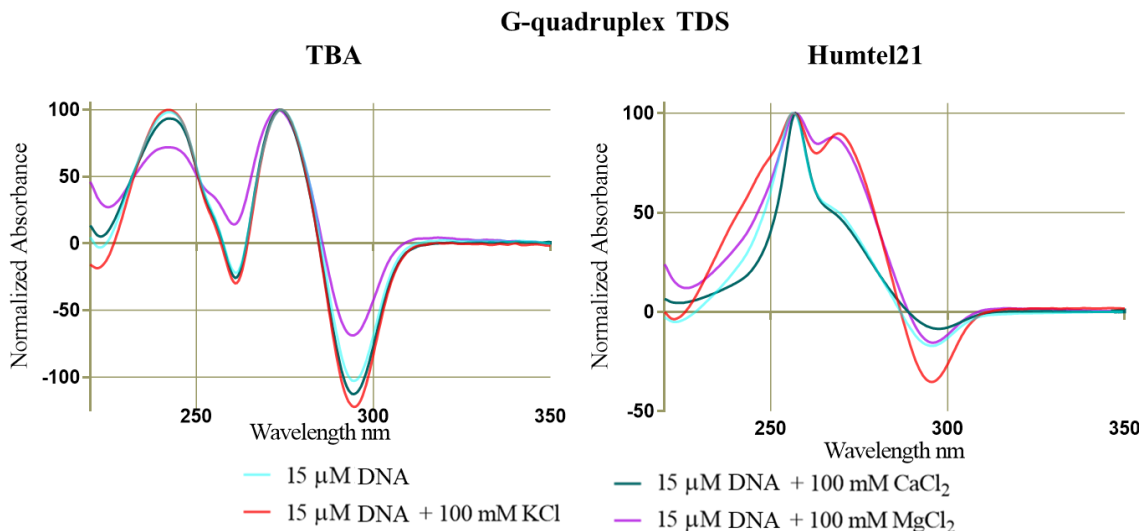
Figure 10. TDS spectra of G-quadruplex DNA. Left is an example of an anti-parallel G-quadruplex. Right is an example of a hybrid G-quadruplex.

Combining temperature melts and thermal difference spectra the formation of DNA secondary structure can be assessed. The goal of this method is to find a pattern that relates each TDS signal to CD topology data in order assign topology of each oligonucleotide. Looking at the spectra from 275 nm to 230 nm, classification of DNA secondary structure can be assessed. Thermal difference spectra were performed and obtained in conjunction with the Tm assay. After annealing and once the samples were loading into the spectrophotometer the samples were scanned at 5°C from 350 nm to 220 nm at a rate of 100 nm/min with data intervals of 0.5 nm. Once the Tm assay reached a temperature of 95°C the samples were equilibrated for 10 min. After 10 min the samples were scanned again at 95°C under the same parameters as the scan at 5°C. The 5°C spectrum was then subtracted from the 95°C to where a crude TDS was made. After subtraction, baseline correction was performed to make sure the area from 350 nm to 320 nm was on the baseline at 0 Abs. Peak normalization was then performed in order to better view the TDS signals when on the same figure.

22

TDS data from G-quadruplex forming sequences can have stark differences

(Figure 10). The formation of G-quadruplex DNA is marked by a negative 295 nm peak.

With TBA there are two large positive peaks at 270 nm and 245 nm and one small

negative peak at 260 nm in the presence of KCl and $CaCl_2$. In the presence of $MgCl_2$

there is a minimum at 260 nm however, it stays positive. Humtel21 has positive peaks at

270 nm and 255 nm in the presence of KCl and $MgCl_2$. In the presence of $CaCl_2$

Humtel21 only has one positive peak at 255 nm and a shoulder at 270 nm. There are six

different families of thermal difference spectra that were observed for those

oligonucleotides and buffer solutions. These families can be differentiated by the

presence or absence of a negative peaks at 295 nm or 230 nm, and the shape of the

spectral curves between 275 nm and 230 nm.
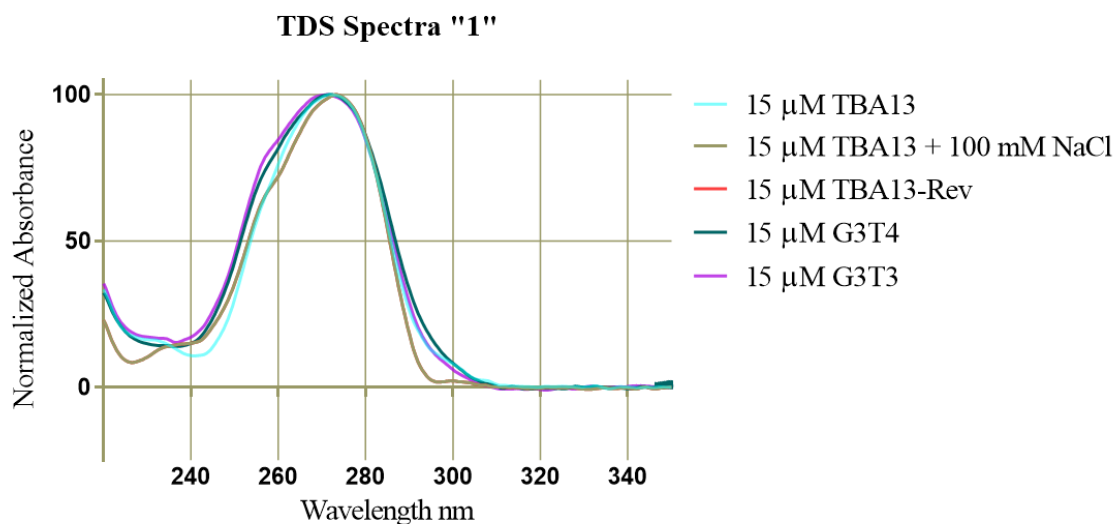
Figure 11. Example of TDS for spectrum family 1.

TDS Spectra "2"

Legend:
- 15 µM TBA13 + 100 mM KCl
- 15 µM TBA13 + 100 mM CaCl$_2$
- 15 µM TBA13-Rev + 100 mM KCl
- 15 µM TBA13-Rev + 100 mM CaCl$_2$
- 15 µM Htel15-Rev + 100mM NaCl

X-axis: Wavelength nm
Y-axis: Normalized Absorbance
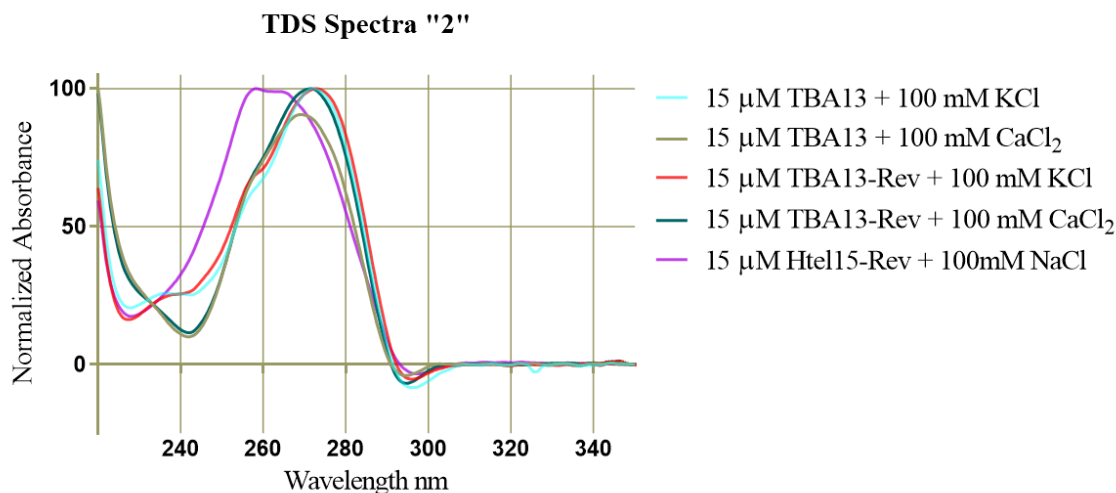
Figure 12. Example of TDS for spectrum family 2.

Reviewing TDS data from all 14 oligonucleotides reveals 6 different spectra present in the 15 µM group. The first family has no negative values and one large peak at 270 nm followed by an area of low differential absorbance from 245 nm (Figure 11). This spectrum is primarily present when there is only Li$^+$ present in the samples. The second family of spectra is similar to the first with one large peak around 270 nm followed by an area of low differential absorbance from 245 nm to 225 nm (Figure 12). However, unlike the first set of spectra, the second displays a slight negative peak at 295 nm.

24

TDS Spectra "3"

Legend:
- 15 μM Htel15 + 100 mM KCl
- 15 μM Htel15-Rev + 100 mM KCl
- 15 μM G3T3 + 100 mM KCl
- 15 μM G3T2 + 100 mM KCl
- 15 μM G3T1 + 100 mM NaCl
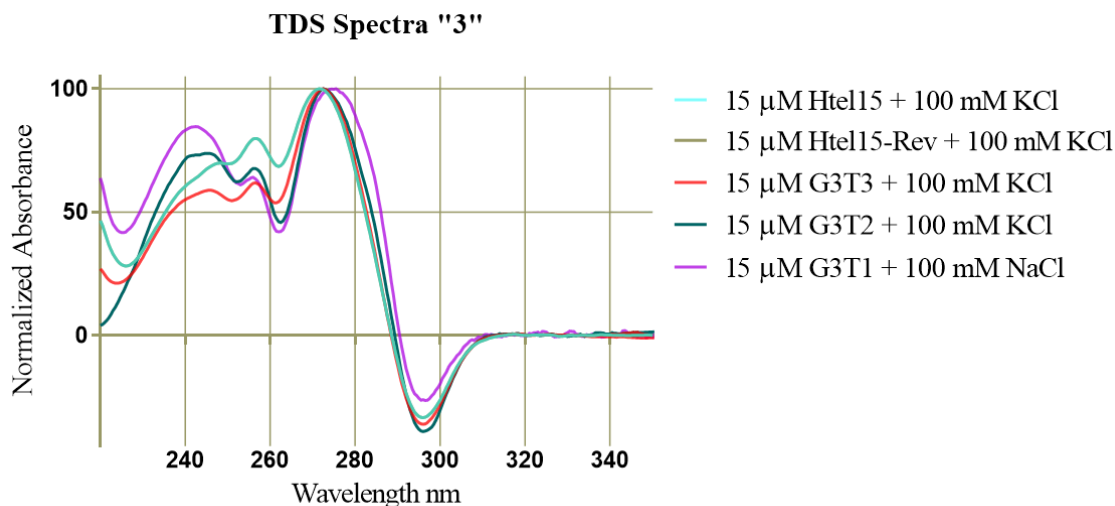
Figure 13. Example of TDS for spectrum family 3.

TDS Spectra "4"

Legend:
- 15 μM Htel15 + 100 mM NaCl
- 15 μM Htel15-A + 100 mM NaCl
- 15 μM Htel15-A + 100 mM KCl
- 15 μM Htel15-delT + 100 mM KCl
- 15 μM G3T2 + 100mM $CaCl_2$
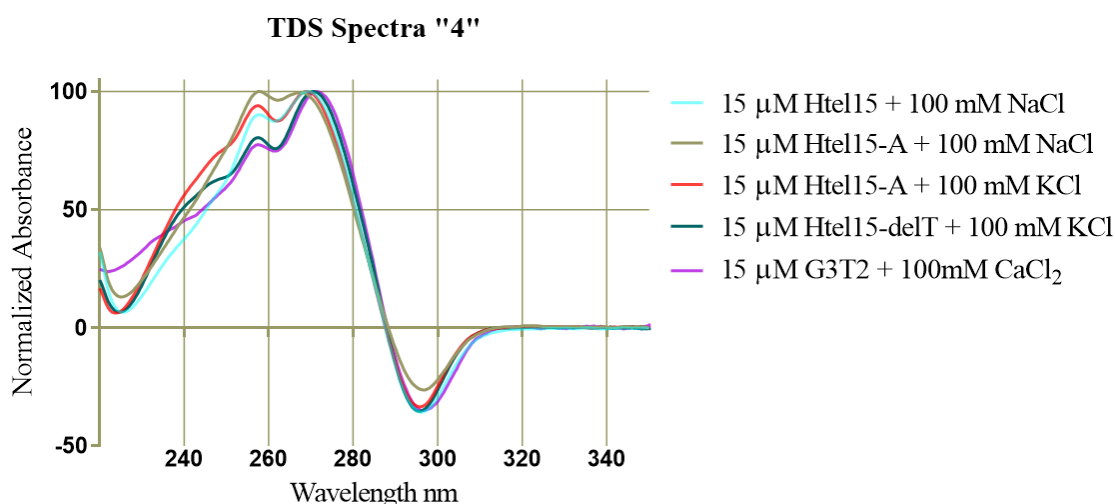
Figure 14. Example of TDS for spectrum family 4.

The third family spectra largely resemble that of a G-quadruplex (Figure 13). There is a well-defined negative peak at 295 nm, a large positive peak at 270 nm, a smaller positive peak at 260 nm, and a short broad positive peak at 245 nm. The fourth family of spectra looks as though it is a morph between the second and third family of spectra (Figure 14). In the third spectra there is a well-defined negative 295 nm peak, a large positive peak at 270 nm, and a smaller second positive peak at around 260 nm.

After the second peak at 260 nm there is a decrease in absorbance diagonally toward a

minimum at 225 nm.

TDS Spectra "5"
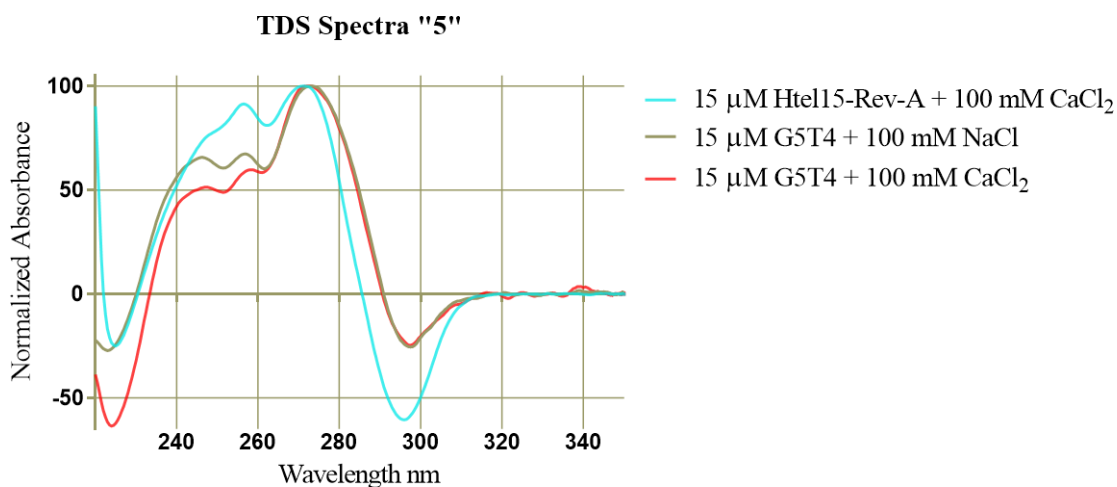
Figure 15. Example of TDS for spectrum family 5.

TDS Spectra "6"

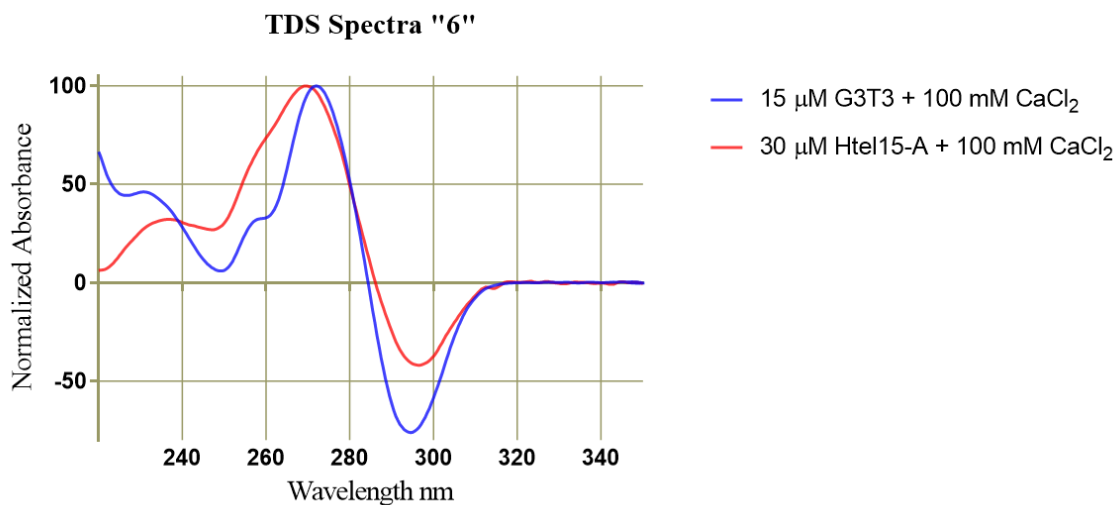**Figure 16.** Example of TDS for spectrum family 6.

The fifth family of spectra is very similar to the fourth family with a well-defined

negative peak at 295 nm, a large positive peak at 270 nm, a smaller positive peak at 260

nm, and a short broad positive peak at 245 nm (Figure15). Unique to the fifth family is

that after the positive peak at 245 nm there is a sharp decrease in absorbance until a negative peak forms at 225 nm. The sixth family of spectra seen is unique to 15 µM G3T3 with 100 mM $CaCl_2$ and 30 µM Htel15-A with 100 mM $CaCl_2$ (Figure 16). In this family there is a well-defined negative peak at 295 nm, a large sharp peak at 270 nm, a positive minimum at 250 nm, and a broad positive peak at 230 nm. With 15 µM G3T3 with 100mM $CaCl_2$ there is also a positive shoulder present at 255 nm.

Looking at TDS and UV Tm data, there is a trend present. Generally, families 3-6 observed from TDS data are more thermally stable then families 1-2. There are a few exceptions when looking at Htel15, 30 µM Htel15-Rev, 15 µM Htel15-A, Htel15-delT, and G3T4 in Ca^{2+}. It is unclear whether this exception is sequence dependent or because of buffer conditions.

Comparing the folding and topology of the various oligonucleotides based on TDS (15 and 30 µM of oligonucleotide) and CD spectra (30 µM of oligonucleotide) that were obtained by the Tuesuwan group in Thailand in the table below (Table 3) there is overall no discernable pattern between the TDS data and the CD data. When comparing thermal stability, determined by UV (30 µM) and CD (30 µM), the melting temperatures are relatively close for the more stable sequences (Htel15, Htel15-Rev, G5T4, G4T4, G3T4, G3T3, and G3T1). However, it is notable that when just Li^+ is present there is a stark difference. The melting temperatures obtained by UV are much lower than those obtained by CD. This can be observed in buffers containing Na^+ (Htel15, Htel15-Rev, G3T3 and G3T2) and Mg^{2+} (Htel15 and Htel15-Rev) as well. In buffer conditions containing K^+ or Ca^{2+} the melting temperatures are very similar. When comparing TDS families and CD topology of sequences with very close, ~3 ºC, melting temperatures a

pattern appears to emerge. TDS families 3 and 4 are mostly associated with parallel

topologies obtained via CD.

Table 3. Comparison of TDS data with CD data. TDS numbers represent assigned TDS spectra. A number followed by a dash then either 15 or 30 represents spectra that is different based on concentration of DNA. Example: Htel15A with KCl, 3-15 / 4-30, would represent that 15 µM Htel15-A presents the signal 3 TDS and 30 µM Htel15-A presents the signal 4 TDS. Letters next to the CD rows represent CD topology data gathered by colleagues in Thailand. A represents anti-parallel, P represents parallel, H represents hybrid, SS represents single strand, and U represents unknown. TDS Tm data is from 30 µM oligonucleotide concentration. All temperatures are in ºC.

Table Comparing TDS Data and CD Data										
TBA13	**Li**	**Na**	**K**	**Ca**	**Mg**	**G5T4**	**Li**	**Na**	**K**	**Ca**	**Mg**
TDS	1	1	2	2	2	**TDS**	1	5-15 / 4-30	3-15 / 4-30	5-15 / 3-30	4
TDS Tm	8.6	9.1	12.8	15.2	18.4	**TDS Tm**	16.6	64.8	78.5	80.1	81.8
CD	A	A	A	H	H	**CD**	H	P	P	P	P
CD Tm	???	???	17.5	47.2	???	**CD Tm**	25.8	64.1	67.5		82.7
TBA13-Rev	**Li**	**Na**	**K**	**Ca**	**Mg**	**G4T4**	**Li**	**Na**	**K**	**Ca**	**Mg**
TDS	1	2	2	2	2	**TDS**	1	3-15 / 4-30	3-15 / 4-31	3	3
TDS Tm	12.8	15.5	15.0	19.9	17.5	**TDS Tm**	12.4	48.9	81.3	61.3	51.4
CD	A	H	H	HP	HP	**CD**	HP	U	HP	P	P
CD Tm	???	???	20.6	22.1	16.7	**CD Tm**	28.3	46.8	87.7	62.3	48.0
Htel15	**Li**	**Na**	**K**	**Ca**	**Mg**	**G3T4**	**Li**	**Na**	**K**	**Ca**	**Mg**
TDS	1	3	4	2	1	**TDS**	1	3	3-15/ 4-30	1	1
TDS Tm	13.1	28.0	44.3	45.5	23.3	**TDS Tm**	9.1	18.9	38.6	40.7	48.8
CD	P	H	HA	P	P	**CD**	H	H	H	HP	H
CD Tm	34.6	59.1	43.0	44.4	52.4	**CD Tm**	???	18.7	37.9	35.1	41.0
Htel15-Rev	**Li**	**Na**	**K**	**Ca**	**Mg**	**G3T3**	**Li**	**Na**	**K**	**Ca**	**Mg**
TDS	1	1-15 / 2-30	4	4-15 / 1-30	1	**TDS**	1	3	4	6-15 / 3-30	1
TDS Tm	???	12.6	35.6	57.4	12.7	**TDS Tm**	9.67	17.4	38.4	42.3	39.7
CD	P	H	HP	P	P	**CD**	H	H	H	P	HP
CD Tm	26.8	49.1	45.7	58.9	51.9	**CD Tm**	43.1		43.8	41.8	31.2

Table 3. Continued.

Htel15-A	Li	Na	K	Ca	Mg	G3T2	Li	Na	K	Ca	Mg
TDS	1	3	3-15 / 4-30	2-15 / 6-30	1	TDS	1	3	4	3	3
TDS Tm	15.3	29.6	41.8	66.1	54.9	TDS Tm	13.5	18.0	53.3	49.1	24.8
CD	P	H	HA	P	P	CD	H	HP	P	P	P
CD Tm						CD Tm	45.6	41.6	52.9	52.8	39.5
Htel15-Rev-A	Li	Na	K	Ca	Mg	**G3T1**	Li	Na	K	Ca	Mg
TDS	1	2	3-15 / 4-30	5	1	TDS	3	4	4	4	4
TDS Tm	17.9	16.7	42.4	62.0	23.2	TDS Tm	26.8	39.4	69.3	64.4	41.4
CD	H	HP	H	P	P	CD	P	P	P	P	P
CD Tm						CD Tm	25.0	38.2	65.9	61.5	46.1
Htel15-delT	Li	Na	K	Ca	Mg	**G2T4**	Li	Na	K	Ca	Mg
TDS	1	3	3	2	1	TDS	1	1	1	1	1
TDS Tm	13.6	27.7	44.4	48.2	28.0	TDS Tm	???	???	11.8	12.0	???
CD	P	H	HA	P	P	CD	SS	SS	SS	HA	HA
CD Tm						CD Tm				???	10.1

In Silico Screen

The major groove in the G-triplex intermediate is a promising target to prevent G-quadruplex folding. By ligand binding in the major groove of G-triplex DNA, the fourth strand of DNA cannot fold into the major groove, thus preventing G-quadruplex folding. The goal of the *in silico* screen is to find a small molecular weight ligand that is able to bind the G-triplex major groove in order to prevent the fourth strand of a DNA from folding into the major groove in order to fold into a G-quadruplex. T1 is the only NMR solution structure of a G-triplex. Therefore, T1 was used as the receptor molecule in the *in silico* docking for the ligand screen. In order to target the major groove it first was selected by highlighting the major groove as seen in Figure 6 using Autodock tools.

29

Compounds from the ZINC15 database were selected based on the criteria: commercially available, MW = 200-350 Da, predicted LogP = 2–3.5, and no more than one negative charge predicted at pH 7.4. The ZINC15 database subset consisted of 1.2×10^6 compounds that were separated into 12 tranches. Each tranche consists of one to ten groups. The compounds were individually docked with T1 using Autodock Vina at an exhaustiveness value of 16. Exhaustiveness is a measure of how thorough the docking program searches for the most stable mode of docking, higher is more thorough. The top three hits from each group from each tranche was then ranked against each other by lowest predicted binding affinity and the top 27 were examined.

Table 4. Table of top 23 ligands with the highest potential binding affinity. The ligands are sorted by their highest predicted binding affinity. Efficiency was calculated by the following equation: $\left(\frac{Mean\ Predicted\ Binding\ Affinity}{Number\ of\ Heavy\ Atoms} \right) \times 100$.

ZINC ID	kcal/mole 1	kcal/mole 2	kcal/mole 3	Number of Heavy Atoms	Daltons	Average Efficiency Score
19325434	-7.5	-7.3	-6.4	16	224.2	44.2
9828011	-7.1	-6.2	-6.2	18	243.2	36.1
164936326	-7	-6.5	-6.2	17	236.2	38.6
6885996	-7	-6.4	-6.3	17	231.2	38.6
97463340	-7	-6	-6	18	245.2	35.2
17836980	-7	-6	-5.9	18	245.3	35.0
42783800	-7	-5.8	-5.8	18	248.3	34.4
4004793	-7	-5.8	-5.2	16	219.2	37.5
371296	-6.9	-6.9	-6.8	18	243.2	38.1
252487033	-6.9	-6.7	-6.6	17	234.2	39.6
253471299	-6.9	-6.6	-6.6	17	234.1	39.4
408516973	-6.9	-6.5	-6	17	239.2	38.0
57612466	-6.9	-6.5	-6	18	245.2	35.9
70631631	-6.9	-6.3	-6.1	17	236.2	37.8
42783801	-6.9	-6.3	-5.6	18	248.3	34.8
42783799	-6.9	-6.2	-5.7	18	248.3	34.8
28711517	-6.9	-6.1	-6	17	225.2	37.3

Table 4. Continued.

15436161	-6.9	-6.1	-5.7	19	249.3	32.8
408516972	-6.9	-6	-5.9	17	239.2	36.9
42783793	-6.9	-6	-5.5	17	236.3	36.1
22143920	-6.9	-5.7	-5.6	18	236.3	33.7
34789262	-6.9	-5.6	-5.5	17	234.3	35.3
15081695	-6.8	-6.8	-6.3	18	246.2	36.9

Table 5. Structures of top potential G-triplex binding compounds.

Zinc ID	Structure	Zinc ID	Structure
19325434		9828011	
97463340		17836980	
371296		252487033	
57612466		70631631	
28711517		15436161	

Table 5. Continued.
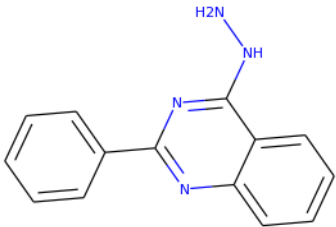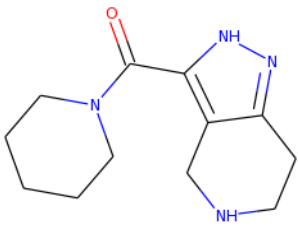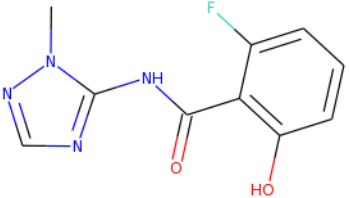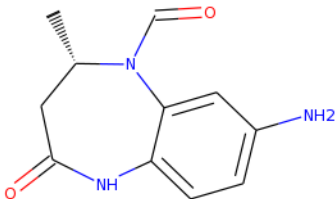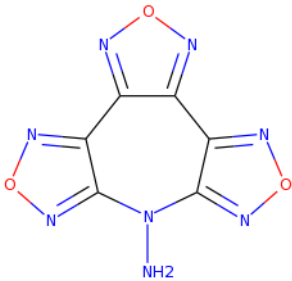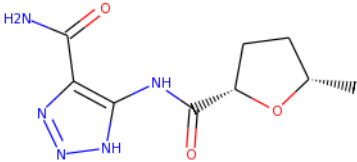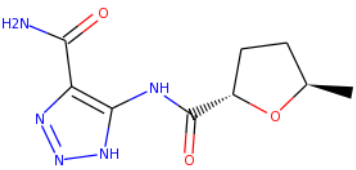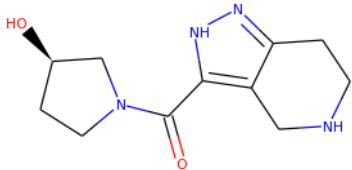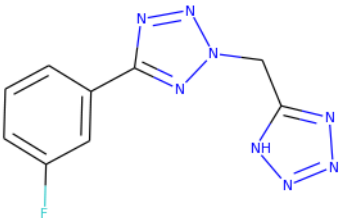
22143920

34789262

164936326

6885996

42783800

4004793

253471299

408516973

42783801

42783799

408516972

42783793

Table 5. Continued.

15081695

The top 27 compounds from *in silico* screening of T1 G-triplex with 1.2×10^6 commercially available compounds from the ZIC15 database were ranked by highest predicted binding affinity (Tables 4 and 5). During analysis, it was discovered that there were four pair of compounds that were identical but in separate tranches. Therefore, the top hits only the compound with the highest predicted binding affinity of each pair was ranked. Structural analysis of the now top 23 compounds revealed that there was a very similar structure shared between four different compounds: ZINCID 42783799, 42783800, 42783801, and 34789262 (Figure 17).
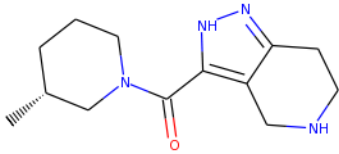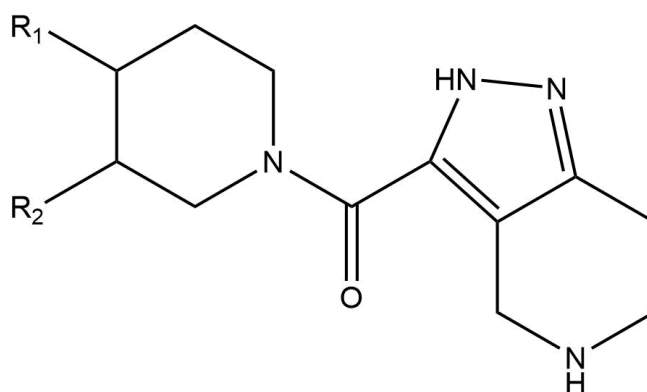

Figure 17. Common structure from *in silico* ligand screen. R_1 is either CH3 or H. R_2 is CH3 with R or S stereochemistry or H.

In silico analysis of structural binding showed that the top compounds are in fact predicted to bind to the T1 G-triplex via the major groove (Figure 18). However, it is interesting to note that one of the top compounds predicted to bind to the T1 G-triplex is predicted to bind in the major groove and intercalate between two G-triads (Figure 19). With ligand 42783800 binding deep in the major groove and with 371296 intercalating between two G-triads a chimeric ligand was designed. The chimeric design combined the 2-ring structure and the carbonyl group of 42783800 with the pyridine and one arm of the 371296 (Figure 20). The chimeric ligand had a lower predicted binding affinity than any of the top ranked compounds from the ZINC15 database screen. It had a predicted binding affinity of -8.5 kcal/mol, -1 kcal/mol lower than the top compound from the *in silico* screen of the ZINC15 database.

Figure 18. Predicted binding of 42783800 to G-triplex DNA. Red is T1 G-triplex DNA and blue is ligand 42783800. Hydrogen bonding is represented by yellow dotted lines.

Figure 19. Predicted binding of 371296 to G-triplex DNA. Red is T1 G-triplex DNA and blue is ligand 371296. Hydrogen bonding is represented by yellow dotted lines.
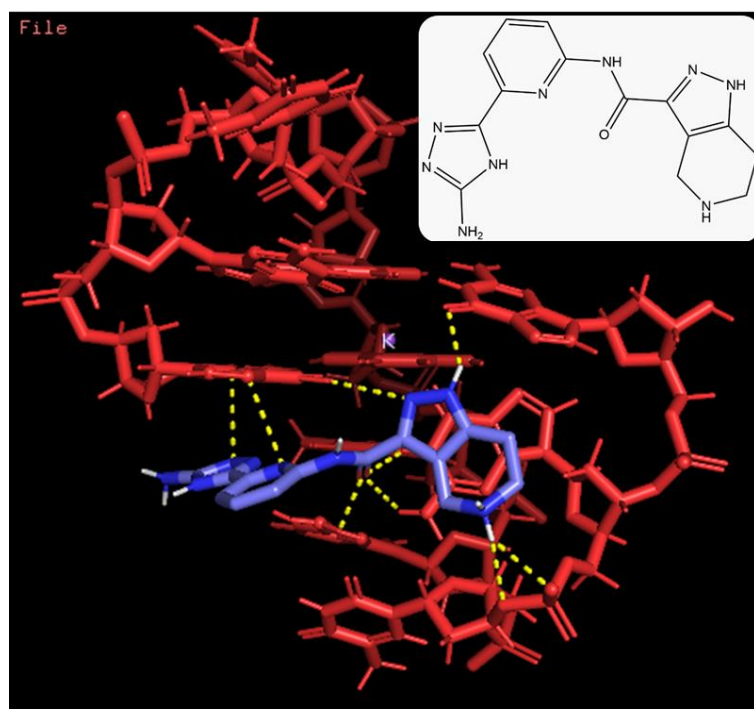
Figure 20. Chimeric ligand binding with G-triplex DNA.

35

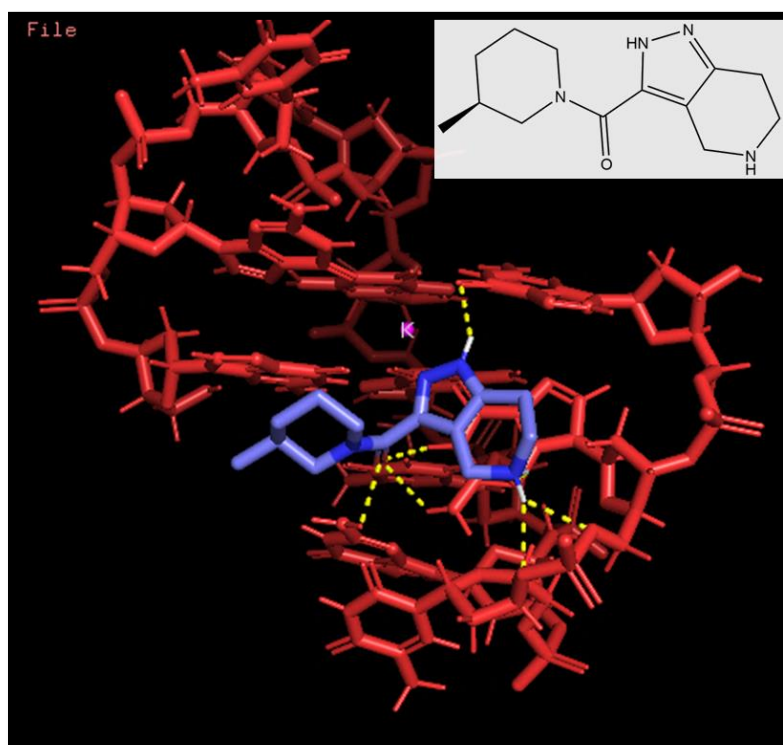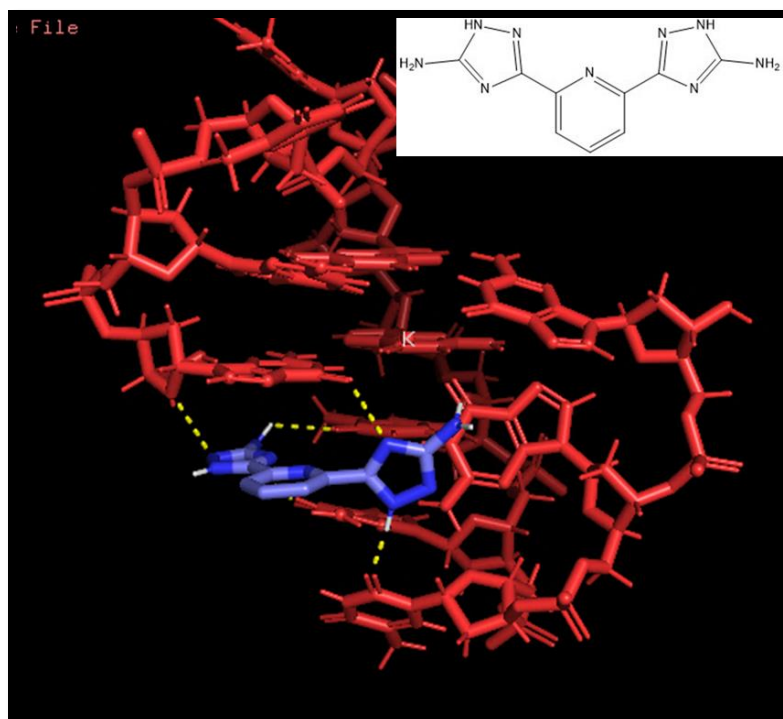**Ligand Binding and Stability**

Ligands from the *in silico* screen were assayed by UV-spectroscopy and with fluorescence spectroscopy, The goal of this experiment is to determine if potential ligands from the *in silico* screen increase the thermal stability of the G-triplex forming oligonucleotides TBA13 and F-T1-T. A similar Tm assay used for determining the thermal stability of the DNA oligonucleotides was utilized by UV-spectroscopy and by FRET to determine the stability of G-triplex forming oligonucleotides in the presence of ligands. Each sample consisted of 5 µM TBA13 or 100 nM F-T1-T in a 10 mM lithium cacodylate buffer at pH 7.3 with 100 mM KCl. Ligand was either co-annealed or added after the annealing process. DNA annealing was performed heating the samples in water bath to 95 ℃, which was allowed to cool over-night, then transferred to a 4 ℃ fridge for 2-3 hours.

The FRET system involved T1 with FAM attached to the 5' end and TAMRA attached to the 3' end (F-T1-T). As F-T1-T folds into a G-triplex, fluorescence excitation energy transfer between the FAM "donor" fluorophore and TAMRA "accepter" fluorophore occurs resulting in decrease in intensity at the emission wavelength, 520 nm, of FAM and an increase in intensity at the emission wavelength, 576 nm, of TAMRA. As the G-triplex unfolds the distance between the FAM and TAMRA fluorophores is increased, this results in a reduction of FRET and an increase in FAM emission signal intensity. The FAM signal has been reported to give more reproducible results, therefore it was used to find the melting temperature as F-T1-T unfolds (**7**). This is evidenced in Figure 22 where the FAM signal can be seen to better show a two-state function based on

the more sigmoidal nature of the curve. The sigmoidal shape of the melting curve allows it to be fitted using non-linear regression in order to find a more accurate melting temperature.

The first ligand to be tested was 34789262. For the ligand assays 34789262 is denoted as T-75. T-75 was initially tested by utilizing an UV-spectroscopy melting assay (Figure 21). Measurement at 295 nm did not show there was a difference in melting temperature of TBA13 when T-75 was present. Measurement at 260 nm however, did show an increase in melting from 25.2 °C to 33.5 °C when T-75 was added before the DNA was allowed to anneal and 35.4 °C when T-75 was added after the DNA annealing process (Table 6). Next T-75 was assayed with F-T1-T by FRET. There was no difference between the Tm of F-T1-T without ligand and the Tm of F-T1-T with T-75.
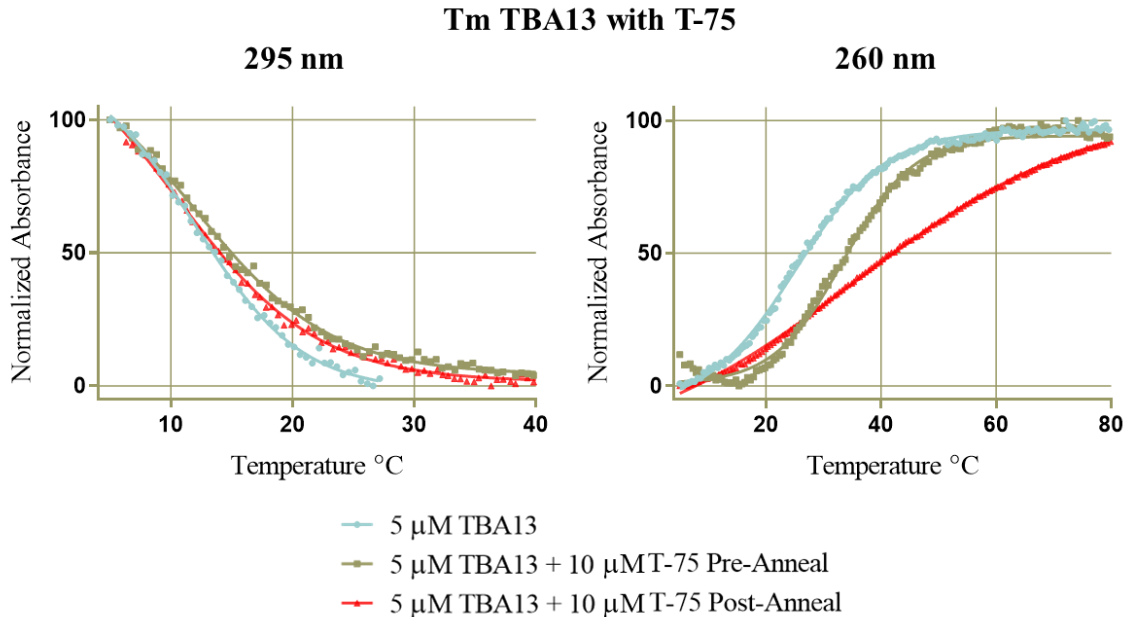
Figure 21. UV Temperature melt of 5 µM TBA13 with 10 µM T-75.
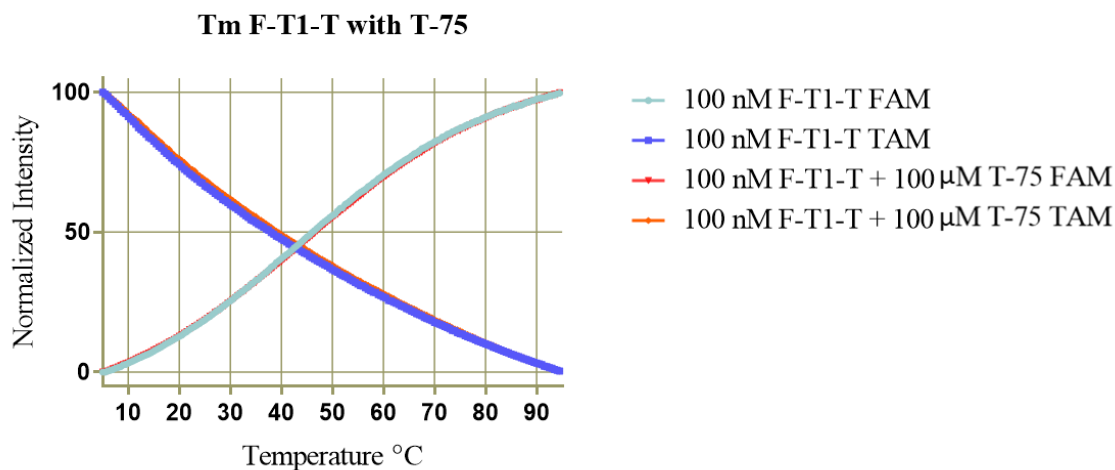
Tm F-T1-T with T-75

Figure 22. Tm of F-T1-T with T-75 using FRET. Ligand was added after DNA was annealed.

Ligands S1, S2, S3, and TeT were hits found through *in silico* screening of commercial vendors by other students in the lab. Ligands S1-S3 did not show a significant change in TBA13 or F-T1-T melting temperature (Figures 23 and 24). The Tm of F-T1-T in the presence of Tetracycline Hydrochloride (TeT) interestingly had a lower melting temperature when added after annealing. F-T1-T by itself had a melting temperature of 45.0 °C however, in the presence of TeT it had a melting temperature of 38.3 °C. Even more interesting is that when TeT was co-annealed with F-T1-T there was a 4.2 ℃ increase in thermal stability.

Tm Ligands Post-Annealed
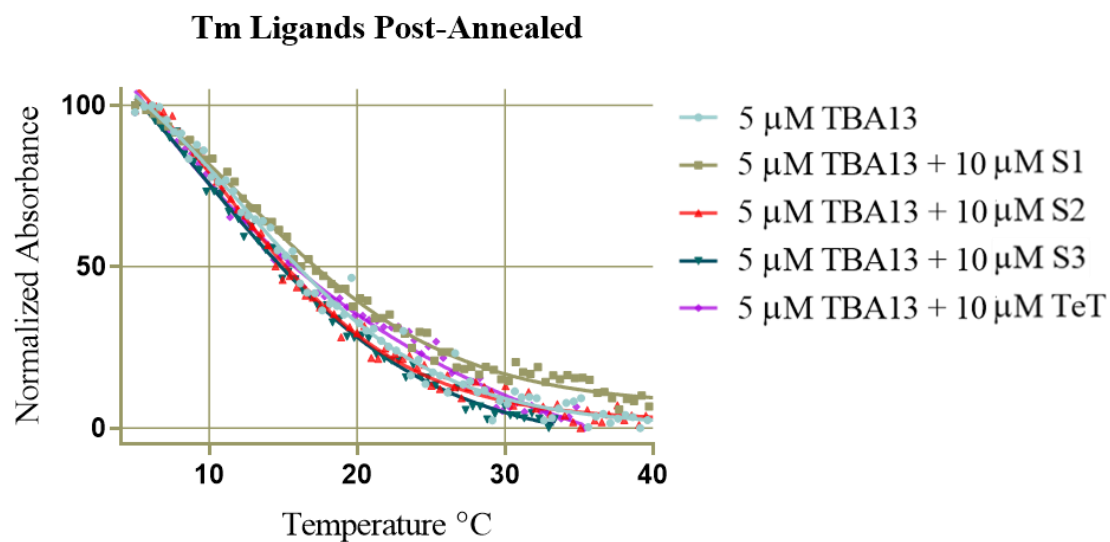
Figure 23. UV Temperature melts of TBA13 with ligands S1, S2, and S3 added after DNA annealing.

FRET Tm of F-T1-T with Ligands Post-Annealed
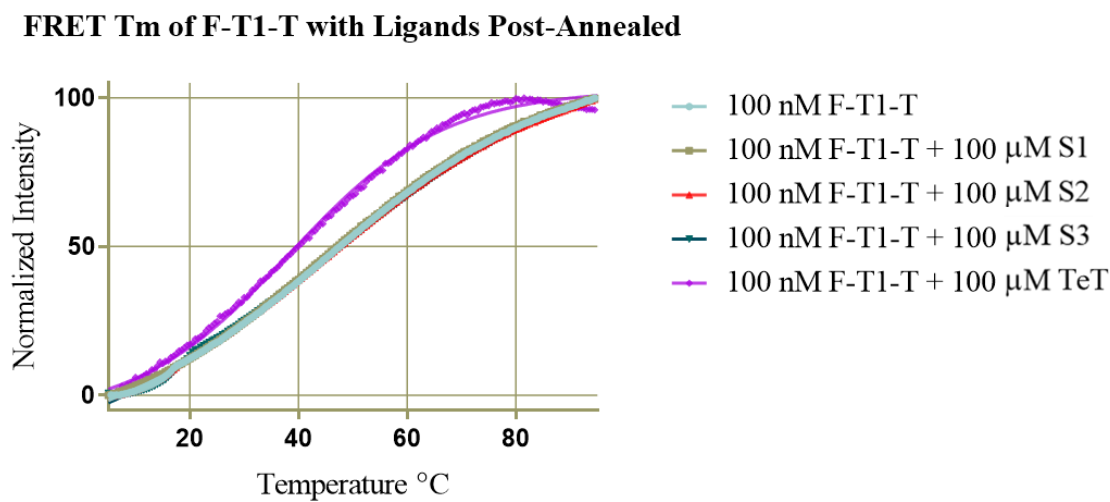
Figure 24. Fluorescence temperature melt of F-T1-T with ligand added post annealing.

Table 6. Tm with ligand present. All temperatures are in ºC.

Tm with Ligand				
DNA	**UV Tm**		**FRET Tm (FAM)**	
	Pre-A	**Post-A**	**Pre-A**	**Post-A**
TBA13 260 // 295	25.2 // 13.1	25.2 // 13.1		
F-T1-T			45.0	45.0
With T-75 260 // 295	33.5 // 11.8	35.4 // 10.6	44.2	44.7
With S1	11.4	12.2	45.9	45.4
With S2	12.4	10.9		44.6
With S3	12.4	9.6		44.1
With TeT	12.9	10.3	49.2	38.3

IV. DISCUSSION

Structure and Stability

G-quadruplex DNA is a non-canonical DNA structure that has been shown to play a role in gene transcription and regulation, replication, and in genomic instability leading to double strand breaks. Due to the unique nature of G-quadruplex DNA it has become a target for genetic regulation and oncology research. There is some evidence that an intermediate to the G-quadruplex may exist. This intermediate is thought to be the G-triplex intermediate and is much less studied.

The G-triplex intermediate is of similar structure as the G-quadruplex however, it has one very important difference. Unlike G-quadruplex DNA, the structure oh which is composed of four strands of DNA, the G-triplex is only composed of three strands. This major difference in structure allows the formation of a major groove to be present in G-triplex DNA. Targeting the major groove of G-triplex DNA serves as an applicable target to prevent the formation of G-quadruplex DNA. A ligand that can bind this major groove would thus prevent the fourth strand of DNA from folding into this groove and thus prevent G-quadruplex folding. Different G-triplex structure topology is thought to affect the width and shape of the major groove. Therefore, determining the topology of the G-triplex is a crucial step in finding a ligand that can bind the major groove.

Thermal stability data obtained from melting temperature assay provided a few observable trends. First, buffer conditions have a major impact on overall stability. G-triplex forming DNA sequences tend to be the most stable in Ca^{2+}, followed by K^+. G-triplex forming DNA sequences in the presence of Mg^{2+} and Na^+ are about equal

however, there are a few cases when Mg^{2+} promotes a more stable G-triplex than Na^+. G-triplex forming DNA sequences in the presence of cations are in the following order Ca^{2+} > K^+ > Mg^{2+} ≥ Na^+ > Li^+. The second observable trend is with strand composition and sequence. The longer the G-tracts are in a sequence, the more thermally stable the G-triplex is. This is most likely attributed to more hydrogen bonds holding the structure together, thus making it more stable. Therefore, GGGGG > GGGG > GGG > GG. The chance of intermolecular G-quadruplexes may also increase with increase G-tract length. In the case of G5T4, a white precipitate can be seen at the bottom the cuvette after the annealing process. This white precipitate is thought to be the formation of G-wires in solution. The third trend that can be observed is the shorter the T-loop is the more thermally stable G-triplex structure. Therefore, TTTT > TTT > TT > T. Interestingly, when T-loop length is ≤ 2 the G-triplex forming DNA is more stable in K^+ than in Ca^{2+}. This may be because that with the more sterically hindered T-loop it is more favorable to coordinate with a cation with a larger radius. Thus, also hindering coordination when a cation with a smaller radius. The fourth trend is that the sequence direction also effects the overall thermal stability. This may be because the size and location of the T-loops is changing. Changing the size and location of the T-loops may affect the topology of the G-triplex making either more stable or less. The fifth and final trend is when a terminal thymine is replaced with an adenine (Htel15 vs Htel15-A) cations with a smaller radius increase the stability of the G-triplex. Changing a pyrimidine (C and T) to a purine (A and G) may allow the G-triplex to better coordinate with cations with a smaller radius, thus increasing the thermal stability of the G-triplex.

A total of six different families of spectra were found from the TDS of all 14 oligonucleotides. The first family does not show any negative 295 nm signal and is the spectra for every oligonucleotide in the presence of only Li^+ except for G3T1. Due to there is no negative 295 nm signal and without Na^+, K^+, $Ca2^+$, or $Mg2^+$ present, family 1 is most likely mostly single strand DNA. The lack of a suitable cation to help coordinate the DNA would mean that the DNA strand is not forming a stable secondary structure. Family 2 looks much like spectra of family 1, however, there is a negative 295 nm signal. With the addition of the 295 nm signal there is some secondary structure present. TBA13 and TBA13-Rev TDS are primarily family 2. TBA13 and TBA13-Rev both have Tm less than 20 °C. Therefore, family 2 spectra may consist of a population of folded DNA with secondary structure and unfolding DNA. This is further evidenced with the small negative 295 nm signal family 2 exhibits. Family of spectra 3 is unique because it has a positive peak around 245 nm. This spectrum appears primarily in the presence of K^+. Additionally, the spectrum looks similar to the G-quadruplex DNA TDS spectra published by Mergny et al (19). Family of spectra 4 have a unique spectrum that appears as though it is intermediate of families 2 and 3. It primarily appears in the presence of Na^+ and K^+ however, there are a few times where it appeared in the presence of $Ca2^+$ and $Mg2^+$. Family 5 only appears with Htel15-Rev-A in the presence of Ca^{2+} and with 15 µM G5T4 in the presence of Na^+ and Ca^{2+}. It is unique because it is the only signal to have a negative peak at 225 nm. Family 6 has a sharp positive peak at 270 nm and a positive peak at 230 nm. This spectrum is only present with 15 µM G3T3 in the presence of $Ca2^+$ and 30 µM Htel15-A in the presence of $Ca2^+$. It is also noticeable that families 3-6 are primarily observed when the G-triplex is more thermally stable and families 1-2 are

observed when the G-triplex is less thermally stable. The exception is when looking at Htel15, 30 µM Htel15-Rev, 15 µM Htel15-A, Htel15-delT, and G3T4 in Ca^{2+}, all of which have melting temperatures above 40°C.

The Tuesuwan group in Thailand have CD topological data on all 14 G-triplex forming oligonucleotides that were tested. Their CD data along with UV TDS was performed in order to obtain even more structural data on G-triplex DNA. When comparing thermal stability, determined by UV and CD, the melting temperatures are relatively close for the more stable sequences (Htel15, Htel15-Rev, G5T4, G4T4, G3T4, G3T3, and G3T1). However, it is notable that when just Li^+ is present there is a stark difference. The melting temperatures obtained by UV are much lower than those obtained by CD in the presence of only Li^+. This can also be observed in buffers containing Na^+ (Htel15, Htel15-Rev, G3T3 and G3T2) and Mg^{2+} (Htel15 and Htel15-Rev) as well. In buffer conditions containing K^+ or Ca^{2+} there is little difference between techniques. When comparing TDS families and CD topology of sequences with similar thermal stability, ~3 ºC, it appears that TDS families 3 and 4 match up the best with parallel topologies obtained by CD.

In Silico Ligand Screen

In silico docking of 1.2×10^6 compounds of a subset of the ZINC15 database yielded a handful of potential ligands for the T1 G-triplex. The highest predicted binding affinity from the screen was from ZINCID 19325434 with a predicted binding affinity at -7.5 kcal/mol. Upon further review of the top structures, there was a similar structure between five of the top 23 compounds. The similar structure shared by five of the top compounds is a good indication of potential ligand binding. The top compound, ZINCID

42783800, had a predicted binding affinity of -7.0 kcal/mol. *In silico* docking revealed that 42783800 bound in the major groove of the T1 G-triplex making a good candidate for *in vitro* studies. Another interesting hit was that of ZINCID 371296 with a predicted binding affinity of -6.9 kcal/mol. *In silico* docking revealed that the top mode of binding for 371296 did bind within the major groove of T1 and intercalated between two G-triads.

Due to the major groove binding of 42783800 and the intercalation by 371296 a chimeric ligand was designed. This design took the best part of T1 binding of 42783800 and the intercalated part of 42783800 and combined them into a single compound. This chimeric compound was then docked with T1. The chimeric compound exceeded the top hit from the ZINC15 database screen with a predicted binding affinity of -8.5kcal/mol. This compound is promising however, it has yet to be tested *in vitro*.

Upon initial *in vitro* testing of 34789262 (T-75) with TBA13 by UV melting assay there was no change in melting temperature at 295 nm. Upon further review there was an increase of 8-10 °C when looking at 260 nm. Carrying on to a FRET Tm assay there was no change in F-T1-T melting temperature in the presence of T-75. However, about a month had passed between testing T-75 by UV before FRET. During this time T-75 was stored in an aqueous solution which may have had an effect on the integrity of the T-75 ligand.

Ligands S1, S2, S3, and TeT were hits found from other *in silico* screens of T1 with commercial manufacturers. These, however, were not shown to increase the thermal stability of TBA13 or of F-T1-T. In the case of TeT the thermal stability of F-T1-T actually decreased by 7 °C when added after annealing. However, when TeT was co-

annealed with F-T1-T there was a 4 ºC increase in thermal stability. Only a handful of ligands were able to be tested. Testing of the remaining top hits may prove beneficial in finding other ligands or binding motifs that can selectively bind the G-triplex intermediate in the G-quadruplex folding pathway.

Future Directions

TDS and Tm data obtained from experiments herein have shed some light on G-triplex topology and thermal stability. Mostly what conditions, buffer or sequential, affect G-triplex stability. Researchers can use this in experimental design to order to better study G-triplex function. Topological data obtained by TD and CD may also be built upon in order to fully characterize G-triplex topology. However, more experiments will need to be done to build upon the current data. This data can in-turn be useful in further ligand binding assays or drug discovery.

Although *in silico* screen did provide ligands which may stabilize the T1 G-triplex more experimentation need to be done. With further testing a ligand which can bind to and stabilize the G-triplex intermediate may be found. If so, it would serve as a great tool for studying G-quadruplex and G-triplex DNA *in vivo*.

APPENDIX SECTION

Tm Data

Tm 15 μM TBA13 N1

Tm 15 μM TBA13 N2

Tm 15 μM TBA13 N3

Tm 30 μM TBA13 N1

Tm 30 μM TBA13 N2

Tm 30 μM TBA13 N3

Legend:
- TBA13
- TBA13 + 100 mM NaCl
- TBA13 + 100 mM KCl
- TBA13 + 100 mM CaCl₂
- TBA13 + 100 mM MgCl₂

Tm 15 μM RevTBA13 N1

Tm 15 μM RevTBA13 N2

Tm 15 μM RevTBA13 N3

Tm 30 μM RevTBA13 N1

Tm 30 μM RevTBA13 N2

Tm 30 μM RevTBA13 N3

Legend:
- RevTBA13
- RevTBA13 + 100 mM NaCl
- RevTBA13 + 100 mM KCl
- RevTBA13 + 100 mM CaCl₂
- RevTBA13 + 100 mM MgCl₂

Tm 15 μM Htel15 N1

Tm 15 μM Htel15 N2

Tm 15 μM Htel15 N3

Tm 30 μM Htel15 N1

Tm 30 μM Htel15 N2

- Htel15
- Htel15 + 100 mM NaCl
- Htel15 + 100 mM KCl
- Htel15 + 100 mM CaCl$_2$
- Htel15 + 100 mM MgCl$_2$

Tm 15 μM RevHtel15 N1

Tm 15 μM RevHtel15 N2

Tm 30 μM RevHtel15 N1

Tm 30 μM RevHtel15 N2

- RevHtel15
- RevHtel15 + 100 mM NaCl
- RevHtel15 + 100 mM KCl
- RevHtel15 + 100 mM CaCl$_2$
- RevHtel15 + 100 mM MgCl$_2$

Tm 15 μM Htel15A N1

Tm 15 μM Htel15A N2

Tm 30uM Htel15A N1

Tm 30 μM Htel15A N2

Htel15A
Htel15A + 100 mM NaCl
Htel15A + 100 mM KCl
Htel15A + 100 mM CaCl$_2$
Htel15A + 100mM MgCl$_2$

Tm 15 μM RevHtel15A N1

Tm 15 μM RevHtel15A N2

Tm 30 μM RevHtel15A N1

Tm 30 μM RevHtel15 N2

Tm 30 μM RevHtel15A N3

RevHtel15A
RevHtel15A + 100 mM NaCl
RevHtel15A + 100 mM KCl
RevHtel15A + 100 mM CaCl$_2$
RevHtel15A + 100 mM MgCl$_2$

Tm 15 μM Htel15delT N1

Tm 15 μM Htel15delT N2

Tm 30 μM HteldelT N1

Tm 30 μM HteldelT N2

- HteldelT
- HteldelT + 100 mM NaCl
- HteldelT + 100 mM KCl
- HteldelT + 100 mM CaCl$_2$
- HteldelT + 100 mM MgCl$_2$

Tm 15 μM G5T4 N1

Tm 30 μM G5T4 N1

Tm 30 μM G5T4 N2

- G5T4
- G5T4 + 100 mM NaCl
- G5T4 + 100 mM KCl
- G5T4 + 100 mM CaCl$_2$
- G5T4 + 100 mM MgCl$_2$

50

Tm 15 µM G4T4 N1

Tm 15 µM G4T4 N2

Tm 30 µM G4T4 N1

Tm 30 µM G4T4 N2

G4T4

G4T4 + 100 mM NaCl

G4T4 + 100 mM KCl

G4T4 + 100 mM CaCl₂

G4T4 + 100 mM MgCl₂

Tm 15 µM G3T4 N1

Tm 15 µM G3T4 N2

Tm 30uM G3T4 N1

Tm 30 µM G3T4 N2

G3T4

G3T4 + 100 mM NaCl

G3T4 + 100 mM KCl

G3T4 + 100 mM CaCl₂

G3T4 + 100 mM MgCl₂

Tm 15 µM G3T3 N1

Tm 15 µM G3T3 N2

Tm 30 µM G3T3 N1

Tm 30 µM G3T3 N2

- G3T3
- G3T3 + 100 mM NaCl
- G3T3 + 100 mM KCl
- G3T3 + 100 mM CaCl$_2$
- G3T3 + 100 mM MgCl$_2$

Tm 15 µM G3T2 N1

Tm 15 µM G3T2 N2

Tm 30 µM G3T2 N1

Tm 30 µM G3T2 N2

- G3T2
- G3T2 + 100 mM NaCl
- G3T2 + 100 mM KCl
- G3T2 + 100 mM CaCl$_2$
- G3T2 + 100 mM MgCl$_2$

Tm 15 µM G3T1 N1

Tm 15 µM G3T1 N2

Tm 30 µM G3T1 N1

Tm 30 µM G3T1 N1

- G3T1
- G3T1 + 100 mM NaCl
- G3T1 + 100 mM KCl
- G3T1 + 100 mM CaCl$_2$
- G3T1 + 100 mM MgCl$_2$

Tm 15 µM G2T4 N1

Tm 30 µM G2T4 N1

Tm 30 µM G2T4 N2

- G2T4
- G2T4 + 100 mM NaCl
- G2T4 + 100 mM KCl
- G2T4 + 100 mM CaCl$_2$
- G2T4 + 100 mM MgCl$_2$

TDS Data

Tm 15μM TBA13

Tm 30 μM TBA13

Legend:
- TBA13
- TBA13 + 100 mM NaCl
- TBA13 + 100 mM KCl
- TBA13 + 100 mM CaCl$_2$
- TBA13 + 100 mM MgCl$_2$

TDS 15 μM RevTBA13

TDS 30 μM RevTBA13

Legend:
- RevTBA13
- RevTBA13 + 100 mM NaCl
- RevTBA13 + 100 mM KCl
- RevTBA13 + 100 mM CaCl$_2$
- RevTBA13 + 100 mM MgCl$_2$

TDS 15 µM Htel15

TDS 30 µM Htel15

- Htel15
- Htel15 + 100 mM NaCl
- Htel15 + 100 mM KCl
- Htel15 + 100 mM CaCl$_2$
- Htel15 + 100 mM MgCl$_2$

TDS 15 µM RevHtel15

TDS 30 µM RevHtel15

- RevHtel15
- RevHtel15 + 100 mM NaCl
- RevHtel15 + 100 mM KCl
- RevHtel15 + 100 mM CaCl$_2$
- RevHtel15 + 100 mM MgCl$_2$

TDS 15 μM Htel15A

TDS 30 μM Htel15A

- — 30 μM Htel15A
- — 30 μM Htel15A + 100 mM NaCl
- — 30 μM Htel15A + 100 mM KCl
- — 30 μM Htel15A + 100 mM CaCl$_2$
- — 30 μM Htel15A + 100 mM MgCl$_2$

TDS 15 μM RevHtel15A

TDS 30 μM RevHtel15A

- — 30 μM RevHtel15A
- — 30 μM RevHtel15A + 100 mM NaCl
- — 30 μM RevHtel15A + 100 mM KCl
- — 30 μM RevHtel15A + 100 mM CaCl$_2$
- — 30 μM RevHtel15A + 100 mM MgCl$_2$

56

TDS 15 μM Htel15delT

TDS 30 μM Htel15delT

- 30 μM Htel15delT
- 30 μM Htel15delT + 100 mM NaCl
- 30 μM Htel15delT + 100 mM KCl
- 30 μM Htel15delT + 100 mM CaCl$_2$
- 30 μM Htel15delT + 100 mM MgCl$_2$

TDS 15 μM G5T4

TDS 30 μM G5T4

- G5T4
- G5T4 + 100 mM NaCl
- G5T4 + 100 mM KCl
- G5T4 + 100 mM CaCl$_2$
- G5T4 + 100 mM MgCl$_2$

57

TDS 15 µM G4T4

TDS 30 µM G4T4

TDS 15 µM G3T4

TDS 30 µM G3T4

TDS 15 µM G3T3

TDS 30 µM G3T3

- G3T3
- G3T3 + 100 mM NaCl
- G3T3 + 100 mM KCl
- G3T3 + 100 mM $CaCl_2$
- G3T3 + 100 mM $MgCl_2$

TDS 15 µM G3T2

TDS 30 µM G3T2

- G3T2
- G3T2 + 100 mM NaCl
- G3T2 + 100 mM KCl
- G3T2 + 100 mM $CaCl_2$
- G3T2 + 100 mM $MgCl_2$

TDS 15 μM G3T1

TDS 30 μM G3T1

G3T1
G3T1 + 100 mM NaCl
G3T1 + 100 mM KCl
G3T4 + 100 mM CaCl$_2$
G3T1 + 100 mM MgCl$_2$

TDS 15 μM G2T4

TDS 30 μM G2T4

G2T4
G2T4 + 100 mM NaCl
G2T4 + 100 mM KCl
G2T4 + 100 mM CaCl$_2$
G2T4 + 100 mM MgCl$_2$

Ligand Stability Data

UV-spectra of Ligands

Legend:
- 10 μM S1
- 10 μM S2
- 10 μM S3
- 10 μM T-75

X-axis: Wavelength nm
Y-axis: Absorbance

Tm Ligands Co-Annealed

Legend:
- 5 μM TBA13 + 10 μM S1
- 5 μM TBA13 + 10 μM S2
- 5 μM TBA13 + 10 μM S3
- 5 μM TBA13 + 10 μM TeT

X-axis: Temperature °C
Y-axis: Normalized Absorbance

Fluoresence of Ligands

Legend:
- 100 μM S1 FAM
- 100 μM S2 FAM
- 100 μM S3 FAM
- 100 μM T-75 FAM
- 100 μM TeT FAM

REFERENCES

1. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, T. N., Bhat, H., Weissig, I. N., & Shindyalov, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research, 28*, 235-242.

2. Bifi, G., Tannahill, D., & Mcafferty, J. (2013). Quantitive Visualization of DNA G-quadruplex Structures in Human Cells. *Nature Chemistry, 5*. doi:10.1038/NCHEM.1548

3. Bochman, M., Paeschke, K., & Zakian, V. (2012). DNA Secondary Structures: Stability and Function of G-quadruplex Structures. *Nat Rev Genet, 13*(11), 770-780.

4. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K., & Neidle, S. (2006). Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Research, 34*(19), 5402-5415. doi:10.1093/nar/qkl655

5. Cerofolini, L., Amato, J., Giachetti, A., Limongelli, V., Novellino, E., Parrinello, M., . . . Luchinat, C. (2014). G-triplex Structure and Formation Propensity. *Nucleic Acids Research, 42*(21), 13393-13404. doi:10.1093/nar/gku1084

6. Clegg, R. M. (2006). *The History of FRET: From Conception Through the Labors of Birth* (Vol. 3). New York: Springer.

7. De Cian, A., Guittat, L., Kaiser, M., Sacca, B., Amrane, S., Bourdoncle, A., . . . Mergney, J.-L. (2007). Flourescence-based melting assays for studying quadruplex ligands. *Methods, 42*(2), 183-195. doi:10.1016/j.ymeth.2006.10.004

8. Eddy, J., & Maizels, N. (2006). Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res, 34*(14), 3887-3896. doi:10.1093/nar/gkl529

9. Huber, M. D., Lee, D. C., & Maizels, N. (2002). G4 DNA unwinding by BLM and Sgs1p: substrate specificity and substrate-specific inhibition. *Nucleic Acids Res, 30*(18), 3954-3961.

10. Huppert, J. L., & Balasubramanian, S. (2007). G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res, 35*(2), 406-413. doi:10.1093/nar/gkl1057

11. Irwin, J. J., & Shoichet, B. K. (2005). ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model, 45*(1), 177-182. doi:10.1021/ci049714+

12. Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., & Coleman, R. G. (2012). ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model, 52*(7), 1757-1768. doi:10.1021/ci3001277

13. Jiang, H., Cui, Y., & Zhoa, T. (2015). Divalent Cations and Molecular Crowding Buffers Stabilize G-triplex at Physiologyically Relevant Temperatures. *Scientific Reports, 9255*(5).

14. Kim, M.-Y., Vankayalapati, H., Shin-ya, K., Wierzba, K., & Hurley, L. H. (2002). Telomestatin, a Potent Telomerase Inhibitor That Interacts Quite Specifically with the Human Telomeric Intramolecular G-Quadruplex. *J. Am. Chem. Soc, 124*(10), 2098-2099. doi:10.1021/ja017308q

15. Lago, S., Nadai, M., Rosseto, M., & Richter, S. N. (2018). Surface Plasmon Resonance kinetic analysis of the interaction between G-quadruplex nucleic acids and an anti-G-quadruplex monoclonal antibody. *BBA, 1862*, 1276-1282. doi:10.1016/j.bbagen.2018.03.002

16. Limongelli, V., Tito, S., & Cerofolini, L. (2013). The G-Triplex DNA. *Angew. Chem. Int. Ed, 52*, 2269-2273.

17. Livshits, G. I., Stern, A., Rotem, D., Borovok, N., Eidelshtein, G., Miglore, A., . . . Porath, D. (2014). Long-range charge transport in single G-quadruplex DNA molecules. *Nature Nanotechnology, 9*, 1040-1046.

18. London, T. B., Barber, L. J., Mosedale, G., Kelly, G. P., Balasubramanian, S., Hickson, I. D., . . . Hiom, K. (2008). FANCJ Is a Structure-specific DNA Helicase Associated with the Maintenance of Genomic G/C Tracts. *J Biol Chem, 283*(52), 36132-36139. doi:10.1074/jbc.M808152200

19. Mergny, J., Lacroix, L., & Amrane, S. (2005). Thermal Difference Spectra: A Specific Signature for Nucleic Acid Structures. *Nucleic Acids Research, 33*(16), e138.

20. Mohaghegh, P., Karow, J. K., Brosh Jr, R. M., Bohr, V. A., & Hickson, I. D. (2001). The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res, 29*(13), 2843-2849.

21. Oganesian, L., Moon, I. K., Bryan, T. M., & Jarstfer, M. B. (2006). Extension of G-quadruplex DNA by ciliate telomerase. *EMBO J, 25*(5), 1148-1159. doi:10.1038/sj.emboj.7601006

22. Paeschke, K., Capra, J. A., & Zakian, V. A. (2011). DNA replication through G-quadruplex motifs is promoted by the S. cerevisiae Pif1 DNA helicase. *Cell, 145*(5), 678-691. doi:10.1016/j.cell.2011.04.015

23. Rezler, E. M., Seenisamy, J., Bashyam, S., Kim, M.-Y., White, E., Wilson, w. D., & Hurley, L. H. (2005). Telomestatin and Diseleno Sapphyrin Bind Selectively to Two Different Forms of the Human Telomeric G-Quadruplex Structure. *J. Am. Chem. Soc, 127*(26), 9439-9447. doi:10.1021/ja0505088

24. Sanders, C. M. (2010). Human Pif1 helicase is a G-quadruplex DNA-binding protein with G-quadruplex DNA-unwinding activity. *Biochemical Journal, 430*(1), 119-128. doi:10.1042/BJ20100612

25. Shay, J. W., & Wright, W. E. (2011). Role of teomeres and telomerase in cancer. *Semin Cancer Biol, 21*(6), 349-353. doi:10.1016/j.semcancer.2011.10.001

26. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., & Hurley, L. H. (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcirption. *PNAS, 99*(18), 11593-11598.

27. Stadbauer, P., Trantírek, L., & Cheatham III, T. (2014). Triplex Intermediates in Folding of Human Telomeric Quadruplexes Probed by Microsecond-scale Molecular Dynamics Simulations. *Biochimie, 105*, 22-35.

28. Trott, A., & Olson, J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of Computational Chemistry, 31*, 455-461. doi:10.1002/jcc.21334

29. Vasquez, K., & Wang, G. (2013). The Yin and Yang of Repair Mechanisms in DNA Structure-induced Genetic Instability. *Mutation Research, 743-744*, 118-131.

30. Wang, G., & Vasquez, K. (2014). Impact of Alternative DNA Structures on DNA Damage, DNA Repair, and Genetic Instability. *DNA Repair (Amst), 19*, 143-151. doi:10.1016/j.dnarep.2014.03.017

31. Watson, J., & Crick, F. (1953). A Structure for Deoxyribose Nucleic Acid. *Nature, 171*, 737-738.

32. Whitmore, L., & Wallace, B. A. (2008). Protein secondary structure analyses from circular dichroixm spectroscopy: Methods and reference databases. *Biopolymers, 89*(5). doi:10.1002/bip.20853

33. Zahler, A. M., Williamson, J. R., Cech, T. R., & Prescott, D. M. (1991). Inhibition of telomerase by G-quartet DNA structures. *Nature*, 718-720. doi:10.1038/350718a0

34. Zhao, J., Bacolla, A., Wang, G., & Vasquez, K. (2010). Non-B DNA Structure-induced Genetic Instability and Evolution. *Cell Mol Life Sci, 67*(1), 43-62. doi:10.1007/s00018-009-0131-2

35. Zybailov, B., Sherpa, M., Glazko, G., Raney, K., & Glazko, V. (2013). G4-quadruplexes and Genome Instability. *Molecular Biology, 47*(2), 197-204.