C-MYC G-QUADRUPLEX DNA STRUCTURES: CHARACTERIZATION AND PHOTOCLEAVAGE

by

Kimberly J. Long, 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
August 2018

Committee Members:

Sean M. Kerwin, Chair

Wendi M. David

Karen A. Lewis

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

DEDICATION

To Kevin Garcia and my brother, Brandon Long. This would not have been possible without your constant supply of dinners and support.

ACKNOWLEDGEMENTS

I would like to thank the friends I have made in the biochemistry program including Neslor Rodriguez, Brandie Taylor, and Amanda Bohannon. Thank you for being a great emotional support team when things didn't go as planned. I would also like to thank my committee members, Dr. Karen Lewis and Dr. Wendi David, I appreciate your guidance throughout my project and the time you have dedicated to helping me be successful in my research. Thank you to Dr. Steve Whitten, for allowing me to use your equipment and for taking the time to answer my questions. Finally, I would like to thank Dr. Sean Kerwin, I appreciate your open door and the support and guidance you have provided me throughout my journey in this program.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF ABBREVIATIONS

Abbreviation	Description
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
ssDNA	Single-stranded DNA
FUSE	Far upstream element
NMR	Nuclear magnetic resonance
DSBs	Double strand breaks
Ph	Philadelphia chromosome
ABL1	Tyrosine-protein kinase
BCR	Breakpoint cluster region
FRET	Förster resonance energy transfer
CD	Circular dichroism
ROS	Reactive oxygen species
TMPyP4	5,10,15,20-tetra(N-methyl-4-pyridyl) porphyrin
NMM	*N*-methylmesophophyrin IX
cMYC22m	5'TGAGGGTGGGTAGGGTGGGTAG'3
F21T	5'FAM-dGGG(TTAGGG)$_3$-TAMRA'3
cMYCHS	5'TGGGAGGGGCGCTTATGGGGAGGGT'3
FAM	Fluorescein
TAMRA	Carboxytetramethylrhodamine
NHE III	Nuclease hypersensitive element III
T_m	Melting temperature
F-cMYC22m-T	5'FAM-cMYC22m-TAMRA'3
F-cMYCHS-T	5'FAM-cMYCHS-TAMRA'3
MRE	Molar residue ellipticity
DMS	Dimethyl sulfate
8oxoG	Guanine oxidized at C8
Azide-cMYC22m	5'Azide-cMYC22m'3
APS	Ammonium persulfate
SCE	Sodium cacodylate buffer with EDTA

I. INTRODUCTION

1.1 Non-canonical DNA

All known life forms use a four-base genetic code in order to store information. These four bases make up a polymer known as deoxyribonucleic acid (DNA). The sequence of these four bases dictates the resulting ribonucleic acid (RNA) transcripts, translated protein sequences, and subsequent structures. Additionally, the dynamic structure of the genome also allows for more variety and control over transcription and translation.

The canonical form of DNA is the B-form double helix with Watson-Crick base pairing: guanine with cytosine and adenine with thymine. However, DNA has the ability to adopt various non-canonical structures, such as cruciform, Z-DNA, H-DNA, and G-quadruplexes.[1] Triplex DNA is a structure that consists of duplex DNA and a single strand portion that binds the major groove of the duplex through Hoogsteen bonding.[1] These structures are considered H-DNA if the single stranded portion is from one of the duplex strands. In order for triplex structures to occur intramolecularly, they must be formed from sequences of homopurine or homopyrimidines and require a mirror repeat.[1,2] The mirror repeat of H-DNA folds over to form Hoogsteen bonding and is divided by an interlinking region which will form a loop region (Figure 1.1).[1, 2]

G-quadruplexes are formed by stacking of planar guanine tetrads that are held together by extensive Hoogsteen hydrogen bonding (Figure 1.2A).[3, 4] *In vitro* studies have shown that G-quadruplexes have the ability to form from two or four separate DNA strands, or they may form intramolecularly from a single strand, which is assumed to be

the more biologically relevant mode.[5,6] G-quadruplexes have been found to fold into a multitude of topologies, even within a single sequence. This variety arises from a number of factors that affect the formation.[7] The loop regions, or nucleotides not involved in the Hoogsteen bonding, vary in length and sequence composition and are one factor affecting G-quadruplex structure.[3, 8-9] Another factor which dictates topology of the G-qudruplex is the conformation of the glycosydic bond of guanine nucleotides. Glycosydic bonds may adopt either a syn or anti conformation, which determines whether the strands will align parallel or antiparallel to each other (Figure 1.2B).[10, 11] Monovalent cations play an important role in stabilizing G-quadruplex structures by coordinating with the electronegative O6 atom of the guanine residues on the interior of the G-quadruplex.[10] Potassium ions have been shown to provide the most stability to G-quadruplexes, however, other cations also have the ability to stabilize these structures.[12] Other environmental factors, such as pH and molecular crowding, have also demonstrated a contribution to the structural diversity seen in G-quadruplexes.[11, 13]

The formation of a G-quadruplex within chromatin is dependent upon transition from B-DNA to single-stranded DNA (ssDNA), which is assisted by negative supercoiling.[14, 15] Negative supercoiling of DNA occurs both during replication and transcription. Local unwinding occurs in DNA during transcription, which causes positive supercoiling downstream and negative supercoiling upstream of the transcription or replication machinery.[16] DNA supercoiling has been shown to play a role in regulating transcription. This upstream negative supercoiling causes torsional stress that may encourage strand separation in promoter regions.[16] When the c-*MYC* gene is being

transcribed, torsional stress allows opening of the duplex DNA at the far upstream element (FUSE), which is located upstream of promoter 1 (P1).[14, 16]

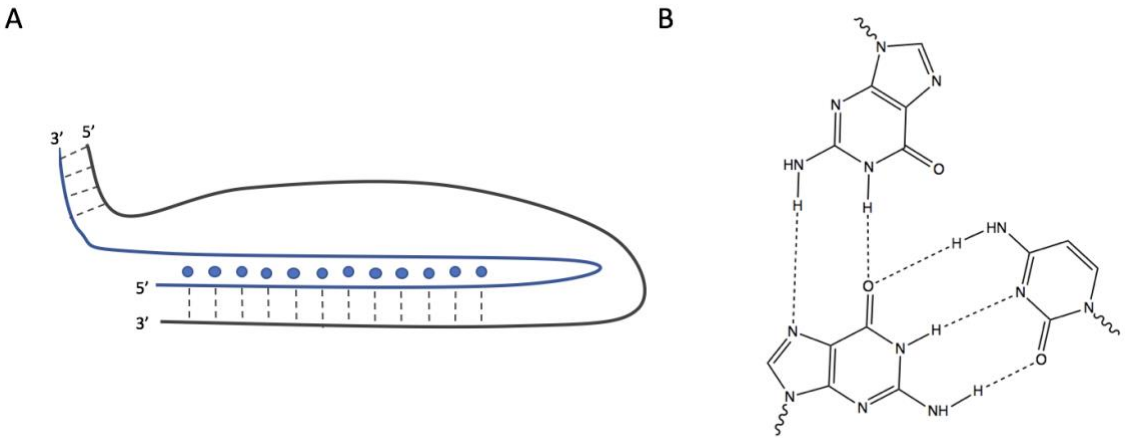

Figure 1.1. H-DNA Structure. (A) Unimolecular, anti-parallel triplex DNA structure. Dashed lines represent Watson-Crick hydrogen bonding, blue dots represent Hoogsteen hydrogen bonding. (B) Example of triplex bonding (G•G-C).

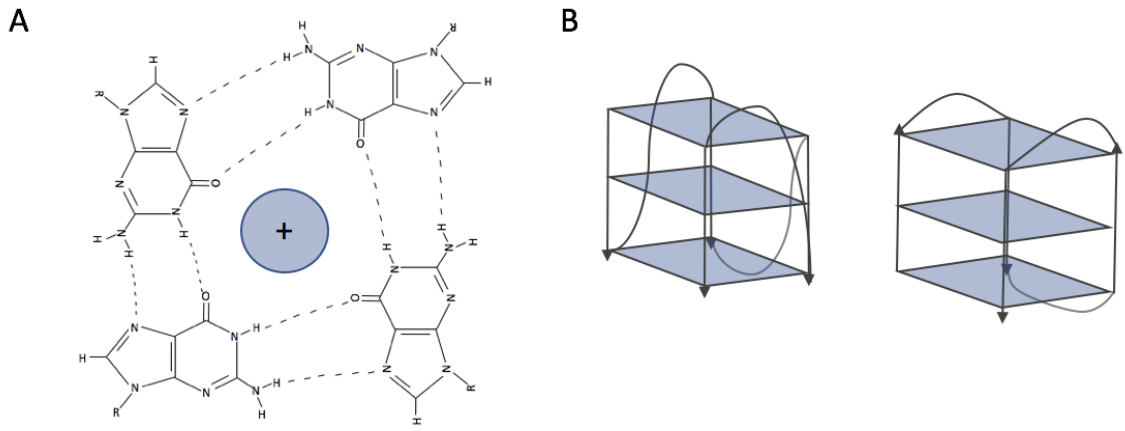

Figure 1.2. G-quadruplex Structure. (A) Hoogsteen bonding of G-tetrad. Bonding occurs between N1 and O6 and between N2 and N7. R represents the deoxyribose sugar of the backbone. (B) Parallel (left) and anti-parallel (right) stranded G-quadruplex diagram.

1.2 Potential G-quadruplexes Within The Genome

G-quadruplexes are proposed to form in specific G-rich regions of DNA, which are over-represented in regulatory regions, such as promoters. This suggests that G-quadruplexes may play a role in biological functions.[3] Previous studies have used bioinformatics to search the human genome for regions that have the potential to form G-quadruplex.[14, 17] The formula used to find potential intramolecular G-quadruplexes within the genome is $G_XN_{L1}G_XN_{L2}G_XN_{L3}G_X$, where X can be any number greater than or equal to 3 and N is 1-7 nucleotides, which will form the loop regions.[18] One such study showed a 230-fold enhancement of putative G-quadruplex sequences in nuclease hypersensitive regions within promoters relative to all genomic DNA.[17] Efforts have also been made to show experimentally that these structures form in cells. Some antibodies, such as BG4 and 1H6, which are specific to G-quadruplex structure, have been used to detect the formation of G-quadruplexes in genomic DNA and telomeres.[19, 20] Experiments involving 1H6 showed increased immunofluorescence when cells were treated with G-quadruplex stabilizing ligands and especially when stained during S phase.[19] Nuclear magnetic resonance (NMR) structures of G-quadruplexes from the promoter regions of MYC, BCL-2 and KIT proto-oncogenes have been resolved.[21-23] Alternatives to NMR to have also been used to experimentally visualize the formation of G-quadruplexes from promoter regions of other proto-oncogenes such as HIF-1α (a transcription factor for VEGF), RB1 and PDGFA.[21,22]

Genetic "hotspots" are regions of DNA that are especially susceptible to mutation, including chromosomal translocations arising from DNA double-strand breaks (DSBs).[24] This type of genetic mutation may be the underlying cause of some cancers.[25] Such is the

case with the Philadelphia Chromosome (Ph), which is a translocation often seen in chronic myeloid leukemia as well as in other leukemias. The translocation seen in Ph positive leukemia results in a fusion of tyrosine-protein kinase (ABL1) and breakpoint cluster region (BCR) genes.[26] Chromosomal translocations are also seen in diffuse large B-cell lymphoma, Burkitt lymphoma, T-acute lymphoblastic leukemia and other cancers.[25] It is suspected that non-canonical DNA structures may be the cause for some of these breaks in hotspot regions.[24] The formation of these structures during replication and transcription can stall polymerases, and, if not unwound by helicases, may cause futile cycling of repair systems.[27] Given the potential roles that G-quadruplexes play in carcinogenesis and cancer cell growth, these structures are of interest for furthering our understanding of cancer development, progression, and treatment.[28]

1.3 *c-MYC* hotspot G-quadruplex

The *MYC* gene is a highly-regulated proto-oncogene, which plays an important role in regulation of the cell cycle, growth, and survival.[28] The transcript of the *c-MYC* gene has been reported to act as a transcriptional activator of genes involved in cell growth, as well as a repressor of genes involved in growth arrest.[29] Translocation events in c-*MYC* have been reported to cause dysregulation of the gene, which is present in many cancers, including lung, colon, and breast.[28] Burkitt Lymphoma is one such cancer with a characteristic translocation occurring within the hotspot of c-*MYC*.[15, 30] The translocation in Burkitt Lymphoma occurs between human chromosome 8, which contains the c-*MYC* gene, and chromosome 14.[15] This hotspot in the promoter region of the c-*MYC* gene has a number of G-rich elements which may form non-canonical DNA structures.[6]

The sequence within this hotspot region contains consensus sequences for both G-quadruplex and H-DNA structures. Due to the sequence overlap for the two structures, it is hypothesized that an intramolecular G-quadruplex and H-DNA structure may be in competition for formation. The formation of these structures is hypothesized to have biological relevance by facilitating genomic instability and the translocation events seen in Burkitt Lymphoma.

The first goal of this project is *in vitro* characterization of the G-quadruplex from the hotspot region of *c-MYC*. Due to the dynamic nature of G-quadruplexes, it is difficult to study these structures using traditional structural analysis methods, such as NMR or crystallography.[31, 32] Therefore, other methods including Förster resonance energy transfer (FRET), circular dichroism (CD) and UV-melting, are to be employed in order to gain insight on the topology and molecular composition of the G-quadruplex structure.

1.4 G-quadruplex Photocleavage

Compounds that effect the cleavage of nucleic acids have been used in order to study primary and secondary DNA structure and as potential cancer therapeutics.[20, 30, 33, 34] Cleavage of nucleic acid structures may arise in a variety of ways. One direct cleavage mechanism is hydrogen abstraction from deoxyribose sugar, which, through rearrangement, leaves breaks in the DNA backbone. This type of direct cleavage is seen in reactions with enediynes and hydroxyl radicals.[33] Cleavage may also occur through indirect means. In this case, the compound may cause a lesion on the DNA base which causes instability and results in a strand breakage upon treatment with heat and a secondary agent, such as piperidine.[33]

Photocleavage agents are a class of nucleic acid cleavage agents which induce digestion of DNA upon irradiation with light. These compounds are appealing due to the control over timing and location in which the DNA cleavage reaction occurs. After absorbing light, these molecules may generate a reactive oxygen species (ROS), which then cleaves DNA. This is considered a type 2 photocleavage reaction.[33] Type 1 photocleavage agents exhibit a direct transfer of electrons from the excited photosensitizer to the nucleotide base. A superoxide radical is usually a secondary product of this reaction.[33, 35] Porphyrins are a common type of compounds studied for their photochemical reactions with DNA.[36]

Detection of G-quadruplexes within cells is desirable because it allows for visualization of the formation of these structures. This can give insight into where G-quadruplexes form, what conditions are favorable for folding, and their general topologies. In addition, probing the genome for G-quadruplex structures may give insight into the effects of their formation. Photocleavage agents may be good candidates as molecular probes for detecting G-quadruplex structures *in vivo.* However, for a G-quadruplex photochemical probe to be useful, it must preferentially cleave the G-quadruplex structure over other DNA structures, requires selectivity for particular G-quadruplex topologies and must prefer the native G-quadruplex topology over a topology induced by binding.

The compound 5,10,15,20-tetra(*N*-methly-4-pyridyl) porphyrin (TMPyP4) has the ability to bind G-quadruplex structures and induce a specific photocleavage pattern based on its interaction.[30] However, TMPyP4 photocleavage is not specific to G-quadruplex DNA and also readily cleaves duplex DNA.[13] Therefore, it is important to continue the

search for nucleic acid cleavage agents which are more specific to the G-quadruplex structure over other secondary structures and are selective toward certain G-quadruplex topologies.[13]

A previous study in our lab group showed that the compound *N*-methyl mesoporphyrin IX (NMM) displays preferential photocleavage activity against G-quadruplex DNA compared to duplex DNA.[37] Furthermore, NMM shows some preference for cleavage of the G-quadruplex formed from an oligonucleotide encompassing a portion of the c-*MYC* promoter sequence (cMYC22m) compared to another G-quadruplex forming oligo, F21T.[37] This photocleavage agent also showed increased cleavage for the c-MYC22m G-quadruplex pre-annealed with NMM before irradiation.[37] This implies that the G-quadruplex can be induced into a non-native conformation by the compound, but the compound more efficiently photocleaves the native G-quadruplex topology. This is suggestive of a compound with potential as a topology-selective G-quadruplex cleavage probe. However, the mechanism by which NMM effects photocleavage of G-quadruplex DNA is unknown.

The second goal of this project is to further investigate the photocleavage properties of NMM in order to elucidate the reaction mechanism. To accomplish this, the sequence of the DNA products from NMM-induced photocleavage of the G-quadruplex will be examined. The pattern of the photocleavage will give insight into the mechanism. These reactions will be done using the original 5'fluorescein (FAM) and 3'5-carboxytetramethylrhodamine (TAMRA) fluorophore system used in a previous study as well as an alternative system with 5'BoDIPY, in order to test for the involvement of the fluorophores in the reaction.

II. *C-MYC* HOTSPOT G-QUADRUPLEX CHARACTERIZATION

2.1 Background

The *c-MYC* hotspot region of the human genome contains the sequence 5'-TGGGAGGGGCGCTTATGGGGAGGGT-'3 (cMYCHS), which is predicted to form an intramolecular G-quadruplex. Additionally, the complement to this strand is assumed to allow the formation of H-DNA as well (Figure 2.1). This sequence is found upstream of the promoter region within the nuclease hypersensitive element (NHE III). The NHE III is a guanine-rich region which has been shown to be involved in the majority of control over *c-MYC* transcription.[38] Genomic instability has been observed in this region of the *c-MYC* gene, leading to DSBs and, consequently, translocations as seen in Burkitt lymphoma.[30] Due to the predicted non-canonical structures that may form from this sequence, it is hypothesized that the G-quadruplex and H-DNA may be in competition for formation. It is also hypothesized that these non-canonical guanine-rich structures play a role in genomic instability in the hotspot region. The goal of this project is to characterize the G-quadruplex from the c-*MYC* hotspot region using UV-melting, FRET, and CD methods.

A well-studied G-quadruplex forming oligonucleotide will be used as a control for this project. This positive control G-quadruplex is a modified sequence of a region found downstream from cMYCHS. The sequence of this 22-mer oligonucleotide, referred to as cMYC22m, is 5'-TGAGGGTGGGTAGGGTGGGTAG-'3, where a G to T mutation was introduced at the thymine in position eleven and twenty of the sequence. This sequence was mutated in order to force it into a single topology so NMR structural

studies could be performed, and the mutated sequence has been used in many other studies.[11, 39]

It has been shown that oligonucleotides folded into G-quadruplexes have an increased absorbance at 295nm, which is not seen in single-stranded DNA, and is indicative of Hoogsteen bonding.[12] The melting temperature (T_m) of a G-quadruplex can be found by monitoring the decrease in absorbance at 295nm over a range of temperatures as the G-quadruplex melts.[12] The T_m correlates with the thermodynamic stability of the G-quadruplex.[12, 40] For these reasons, UV-melting curves will be employed for cMYCHS G-quadruplex characterization.

The fluorescently-labeled cMYCHS oligo will also be used for FRET studies in order to characterize the folding of the oligonucleotide into a G-quadruplex structure. In a FRET system of fluorophores, a donor fluorophore may be excited by some wavelength of light and will subsequently emit light within its specific emission range.[41] When the donor and acceptor molecules are in close proximity, the acceptor will absorb the light emitted by the donor.[41] This system may be monitored by observing the quenching of the donor molecule or by monitoring the specific emission of the acceptor molecule. In this project, a single-stranded 5'-FAM-cMYCHS-TAMRA-3' oligonucleotide will be used for detection of folding. FAM has an absorbance maximum of 492nm and an emission maximum of 517nm, while TAMRA has an absorbance maximum of 565nm and an emission maximum of 580nm.[42, 43] Fluorescence will be measured over a range of temperatures in order to produce FAM or FRET-folding curves, from which T_m will be calculated.

Characterization of the cMYCHS G-quadruplex will also be carried out with CD experiments. It has been shown that different G-quadruplex structures give rise to a characteristic CD spectra based on *anti* or *syn* glycosidic bond angles.[44, 45] By monitoring the CD signal in the range of 220-300nm over a range of temperatures it will be possible to follow the folding and unfolding of the G-quadruplex.[45, 46] Characteristic negative or positive peaks are predicted at approximately 260nm or 240nm depending on parallel or antiparallel G-quadruplex alignment.[45] A positive peak at about 260nm and negative at 240nm is associated with a parallel alignment, while the inverse of this is true for anti-parallel alignment.[45] Using CD measurements, thermodynamic properties may be calculated using Van't Hoff analysis.[40] However, these calculations assume a two-state function, therefore it must be determined if the G-quadruplex forms through intramolecular interactions before this method may be implemented.[40] CD measurements taken over a range of temperatures will be used to determine the molecularity and topology of cMYCHS.

Folding curves obtained from spectroscopic measurements can also give information on whether the G-quadruplex is formed through intramolecular or intermolecular interactions, by monitoring changes in T_m when the concentration of oligonucleotide is varied.[12] The T_m of intramolecular G-quadruplexes will not fluctuate with changes in oligonucleotide concentration, while the T_m of intermolecular G-quadruplexes will.[12] Intramolecular G-quadrupexes will conform to a sigmoidal curve because they occur through two-state folding. Whereas intermolecular G-quadruplexes may be formed with significant intermediates, which would be more than two-states.[47]

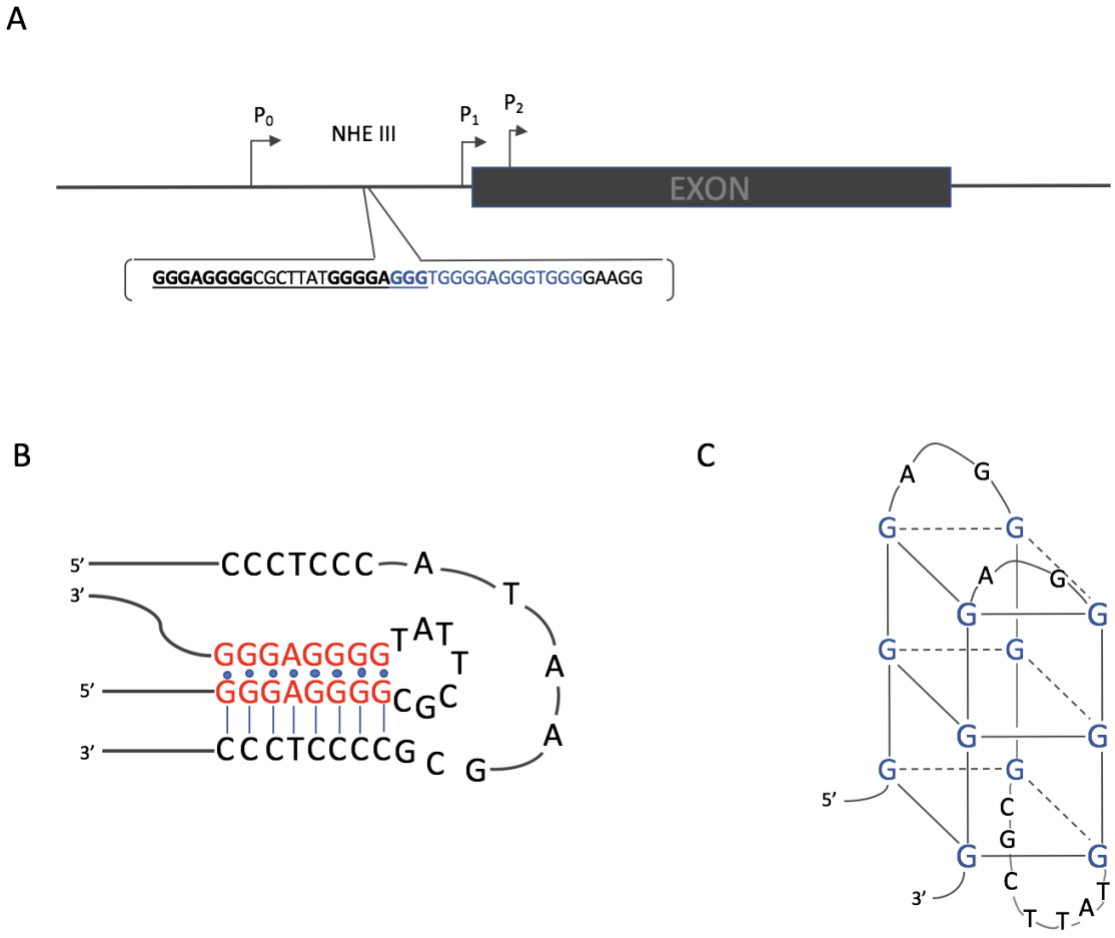

Figure 2.1. Non-canonical DNA within *c-MYC* Hotspot. (A) Overlapping non-canonical DNA sequences. Underlined portion is the sequence of cMYCHS, which may form either H-DNA or G-quadruplex. Bold nucleobases represent those involved in Hoogsteen bonding of H-DNA. The blue region is the wild type sequence of cMYC22m. (B) Possible intramolecular folding of H-DNA from cMYCHS. Blue lines represent Watson-Crick bonding in the duplex portion of the structure and blue dots signify Hoogsteen hydrogen bonding. (C) Potential folding of G-quadruplex from cMYCHS.

2.2 Materials and Methods

Short DNA oligonucleotides cMYC22m (5'-TGAGGGTGGGTAGGGTGGGTAG-3'), F-cMYC22m-T (5'FAM-TGAGGGTGGGTAGGGTGGGTAG-3'TAM), cMYCHS (5'-GGGAGGGGCGCTTATGGGGAGGG-3'), and F-cMYCHS-T (5'FAM-TGGGAGGGGCGCTTATGGGGAGGGT-3'TAM) were obtained from IDT and used without further purification. Chemicals used include cacodylic acid acquired from Sigma-Aldrich, LiOH from Fisher Scientific, EDTA from J.T. Baker, KCl from Fisher Scientific, and HCl from EDM Millipore.

All UV-melting data was acquired using Jasco V-730 UV/vis spectrophotometer. Fluorescence measurements were taken using a Jasco FP-8300 spectrofluorometer. Samples were degassed using the Turbo Sonic 6000 sonicator made by Raytech. All CD data was collected using Jasco J-710 circular dichroism spectropolarimeter. The spectropolarimeter was purged with N_2 gas, purchased from Airgas.

UV-melt curves

A stock solution of Li-cacodylate buffer was prepared with acid at 20mM, EDTA at 0.2mM EDTA, LiOH at 20mM, and the solution was brought to the pH of 7.4 by the addition of LiOH or HCl. A stock solution of 1M KCl was used to make dilutions for each sample. All UV-melting samples were run in 10mM lithium cacodylate buffer, which was dilute from the stock solution and KCl was added to samples in varying concentrations.

Parameters for UV-melt measurements were varied in order to find optimal conditions. Measurements were collected at 295nm over a range of temperatures from

95°C to 20°C, with up to 10°C variations for the lower temperature. A forward and reverse cycle of each sample was taken, with the forward going from 95°C to 20°C and reverse going from 20°C to 95°C. Samples were allowed to equilibrate at the start temperature for 10min before the cycle began. Data measurements were collected every 1°C and the ramp rate was set to 1°C/min for most UV-melts, if varied it will be stated. Blanks for each run were taken with 10mM Li-cacodylate buffer with 15mM KCl. The Jasco Spectra Manager Spectra Analysis software was used to calculate the T_m of the samples by the second derivative or median function. Indicated data sets have been smoothed by averaging each data point with the one preceding it. The cMYC22m oligonucleotide was used as a positive control for UV-melt measurements. Control experiments were performed in order to establish optimal conditions for melts. They were also run in conjunction with cMYCHS in order to monitor working conditions of the experiment.

Due to the melting curve of cMYCHS forming multiple curves and not conforming to a sigmoidal curve, the melting temperatures could not be calculated. Therefore, calculations to compare severity of hysteresis between samples were not done through comparison of melting temperatures. Instead, the hysteresis for UV-melts was found by calculating the temperature difference between five reverse data points and five corresponding forward data points, and these measurements were averaged for each sample. The standard deviation was also calculated for these measurements.

Förster Resonance Energy Transfer

Oligonucleotide solutions were prepared in 10mM Li-cacodylate buffer with 0.1mM EDTA and varying amounts of KCl. Measurements were collected over a range

of temperatures from 95°C to 20°C and the reverse of this cycle, with up to 10°C variations for the lower temperature. Samples were allowed 10min to equilibrate at the start temperature of each cycle before the cycle began. Data was collected for excitation at 494nm and emission at 517nm (FAM), excitation at 494nm and emission at 580nm (FRET) for all samples, some samples also had data collected at excitation at 555nm and emission at 580nm (TAMRA). The sample collection rate was fixed to collect a measurement every 1°C and the ramp rate was set to 1°C/min. Degassing of samples was performed by placing sample tubes in sonicator on degas setting then carefully transferring to cuvettes. Melting temperature analysis was conducted by Van't Hoff analysis, with the online nonlinear lease squares regression (curve fitter) on statpages.info was used to find best fit data. $T_{1/2}$ for F-cMYCHS-T samples were estimated by establishing folded and unfolded baselines, then finding the linear equation for the transitional portion of the graph. The temperature located at the halfway point of the transitional line between the two baselines was ascribed as the $T_{1/2}$ of the sample. In order to compare the severity of hysteresis of the T_m between samples the difference of the forward T_m from the reverse T_m of each sample was found and reported.

Circular Dichroism

Samples of cMYCHS were prepared in 10mM Li-cacodylate buffer with 25mM KCl. The machine was purged with nitrogen gas for 30min before acquiring any data and N_2 gas was also supplied throughout the entire run. Data pitch was set to 0.5nm, scan mode was set to continuous, with a scan speed of 20nm/min, response of 2sec, and band width set to 1.0. Each data point is the accumulation of 8 measurements, which are averaged together. Measurements were taken over the range of 290nm to 220nm with

differing temperatures set for runs. For each run, the machine was set to the desired temperature and allowed to reach this temperature, then it was allowed 10min to equilibrate at this temperature. CD data was converted from milli-degrees to molar residue ellipticity (MRE) based on the number and molecular weight of residues in the oligonucleotide as well as concentration of the sample.

2.3 Results

UV-melts

The G-quadruplex-forming oligonucleotide, cMYC22m, was used as a positive control for UV-melt measurements. Control experiments were performed in order to establish optimal conditions for melts. Melts of the control were also run in conjunction with the cMYCHS oligonucleotide in order to monitor working conditions of the experiment.

cMYC22m was first run as a 5μM sample with 50mM KCl (Figure 2.2A). The average T_m for this sample was calculated to be 81.5°C. There is a slight difference in Tm observed depending upon whether the sample was undergoing cooling (forward) or heating (reverse). This hysteresis will be discussed in a later section. In order to demonstrate the stabilizing effect of potassium on G-quadruplex structures, samples of 5μM cMYC22m were then run with varying KCl concentrations of 50mM, 25mM, 20mM, 15mM, and 10mM (Figure 2.2B). The average T_m for these samples were calculated to be 80.7°C, 75.4°C, 74.6°C, 71.4°C, and 69.7°C respectively. Increased KCl concentration correlated with increased an T_m for the sample. UV-melt data was also collected for cMYC22m at varying oligonucleotide concentrations. This was done in order to test for a stabilized T_m of an intramolecular G-quadruplex when oligonucleotide

concentrations were changed. These samples were run with 15mM KCl, with oligonucleotide concentrations of 2.5µM, 5µM, and 10µM of cMYC22m (Figure 2.3). For ease of comparison, the data for these samples was normalized to a range of 0 to 1. The T_m was calculated for the forward cycle of each of these three samples, and were found to be 72.5°C, 71.6°C, and 71.1°C, respectively. The average T_m for these samples was calculated to be 71.7°C ±0.7°C.

There is a difference in the forward and reverse curves of this data, known as hysteresis. The forward cycle for this data starts at 95°C where all G-quadruplex structures are assumed to be unfolded, and as the temperature decreases the G-quadruplex anneals. Once the samples reach the lowest temperature they were allowed 10min to equilibrate at this temperature. However, the kinetics of these G-quadruplex structures has been shown to be effected by ion concentration, type of ion present, and the time allowed for annealing/melting, and also oligonucleotide concentration for intermolecular G-quadruplexes.[47, 48] The rate constants of association and dissociation cycles may differ depending on these factors, and previous studies have revealed the rates are faster at lower temperatures.[47] Therefore, the ramp rate of 1°C/min is relatively fast and may not facilitate maximal folding in the forward cycle, which starts at the high temperature. This results in a difference in T_m between the forward and reverse cycles. A slower ramp rate of $\leq$ 0.1°C/min would result in a minimal difference of T_m between the two cycles, however, this would take a significant amount of time. Therefore, a ramp rate of 1°C/min or 2°C/min was used for this preliminary data and the average hysteresis is reported (Table 2.2). Due to the cMYCHS curves not conforming to a single sigmoidal curve, the hysteresis is not reported as a difference in T_m for the UV-melts, but as an average

difference between forward and reverse data points along the curve. Points along the curve which displayed a gap were found, and five points which had similar absorbance values on the forward and reverse curves were used to find the difference in temperature between the points. The points used for these measurements and the average difference in the absorbance values of each set of points used is shown in Table 2.1. The average of the five temperature differences was taken for each sample and is reported as the hysteresis in Table 2.2.

Samples of cMYCHS in 25mM KCl were run with the oligonucleotide concentrations at 2.5µM, 5µM, and7µM (Figure 2.4 and 2.5). The data shown in figure 2.4 was also normalized to a range of 0 to 1, to allow for easy visual comparison. These measurements, however, did not conform to a sigmoidal curve, therefore a T_m was not calculated for these data. The average hysteresis between the forward and reverse curves for these samples was calculated and reported in Table 2.2. A 5µM sample of cMYCHS was also run with 100mM KCl to look for a possible stabilizing effect from increased potassium (Figure 2.6). This data also did not follow a sigmoidal curve, and likewise no T_m was calculated. The difficulty in interpretation of this data led to the fluorescent studies. In addition to the data not conforming to a normal curve, a large hysteresis was also displayed between the forward and reverse cycles.

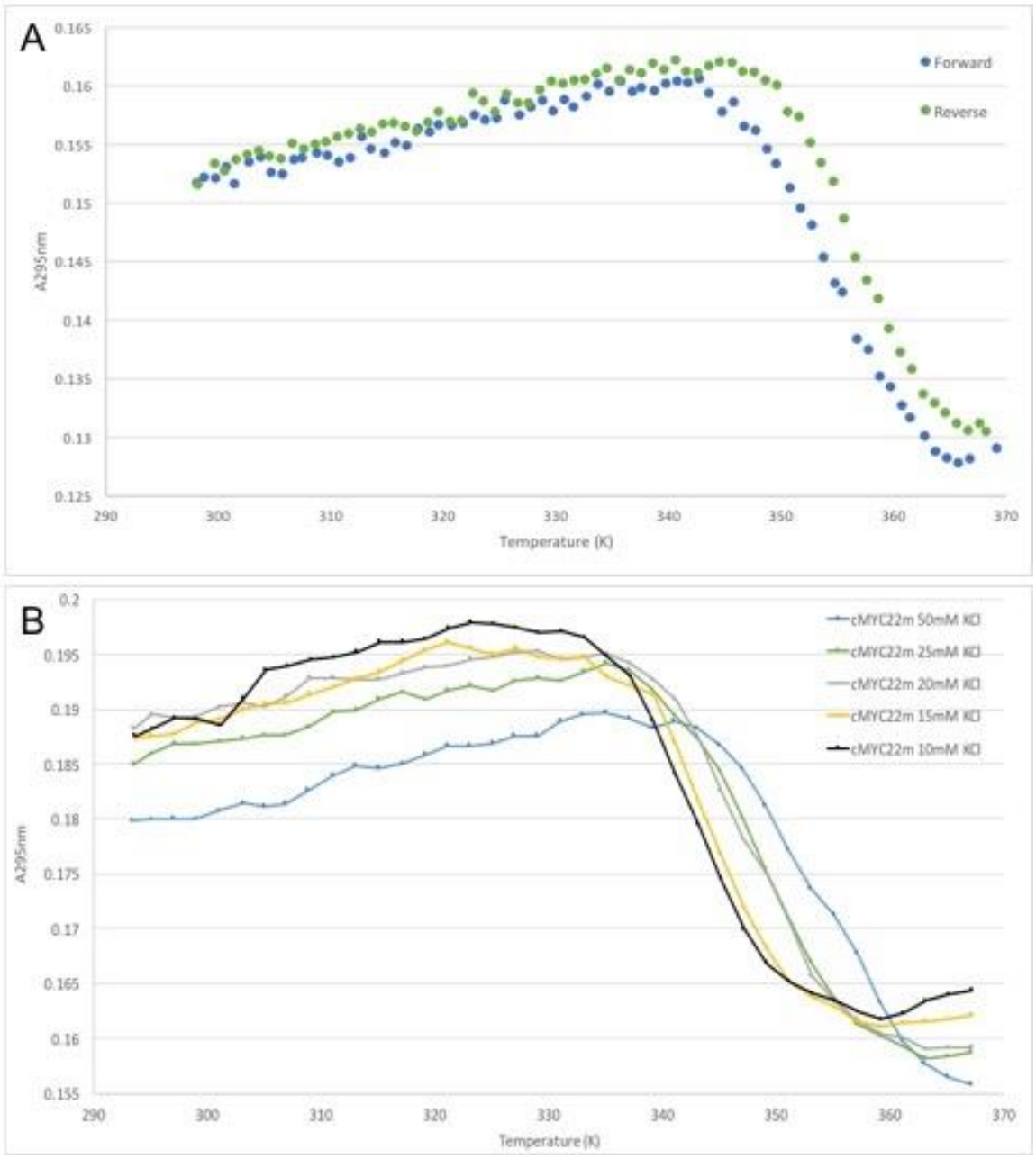

Figure 2.2. cMYC22m UV-folding Curves. Smoothed data curves collected for 295nm absorbance of (A) 5µM cMYC22m in Li-cacodylic buffer with 50mM KCl with measurements taken every 2°C. The forward measurements are from 90°C to 25°C, the reverse of this cycle is labeled reverse. The points represent the average of two data sets. (B) 5µM cMYC22m in Li-cacodylic buffer with KCl concentrations of 50mM, 25mM, 20mM, 15mM, and 10mM. Measurements are the forward cycle from 95°C to 20°C with measurements taken every 2°C

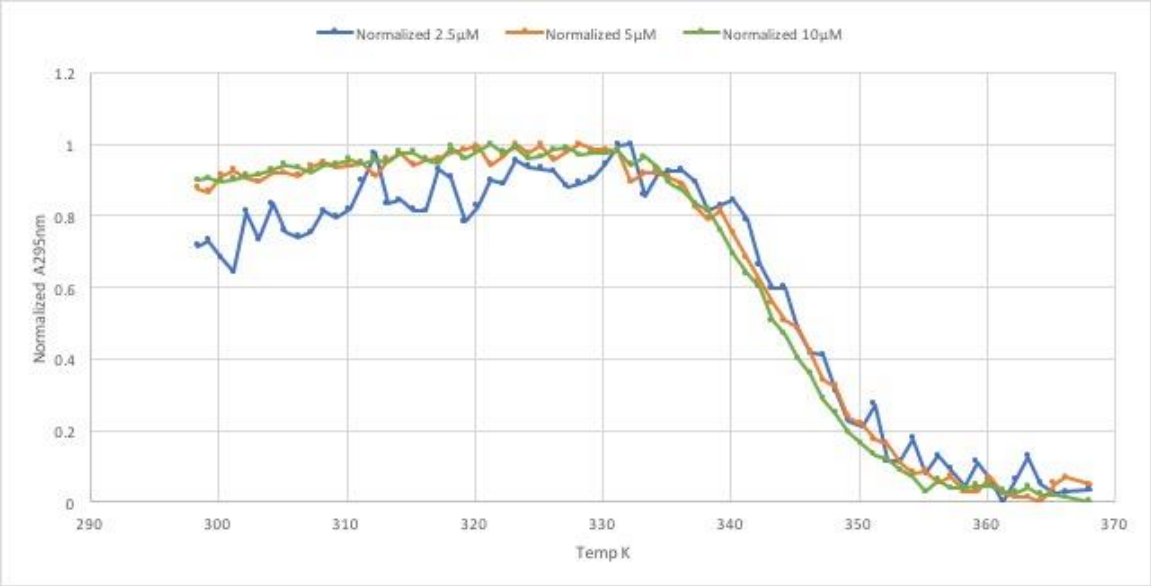

Figure 2.3. Normalized A295nm Folding Curves for Various Concentrations of cMYC22m. Three concentrations of cMYC22m, 2.5µM, 5µM, and 10µM, normalized to a range of 0 to 1, run in 15mM KCl. Run from 95°C to 15°C.

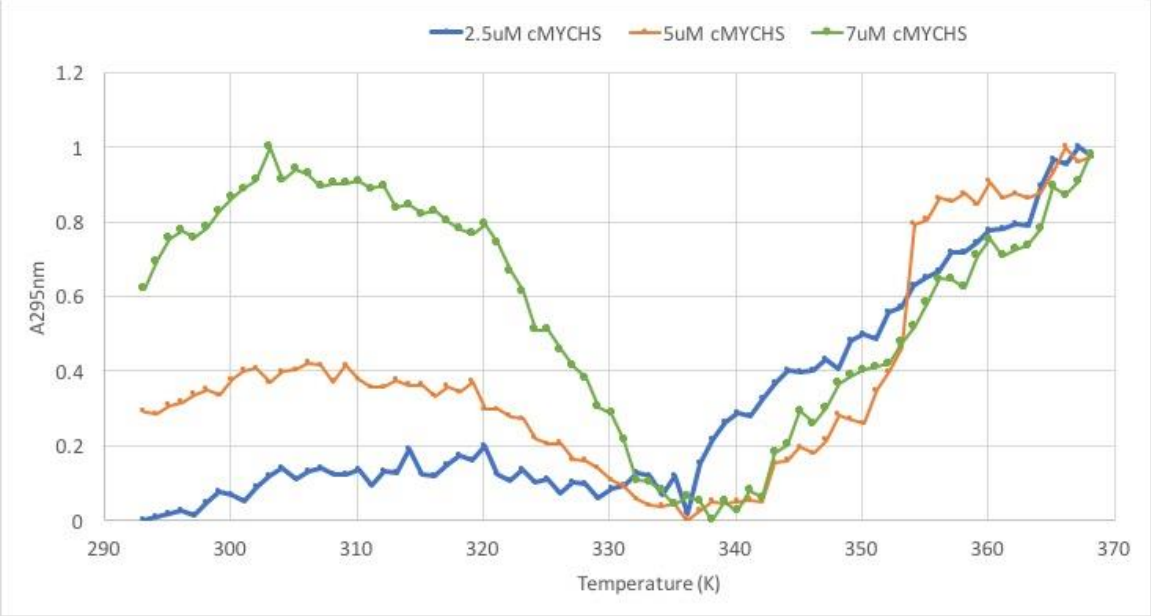

Figure 2.4. Normalized A295nm Folding Curves for Various Concentrations of cMYCHS. Three concentrations of cMYCHS oligonucleotide normalized to start at absorbance of 1, normalized to a range of 0 to 1, run in 25mM KCl. Run from 95°C to 15°C.

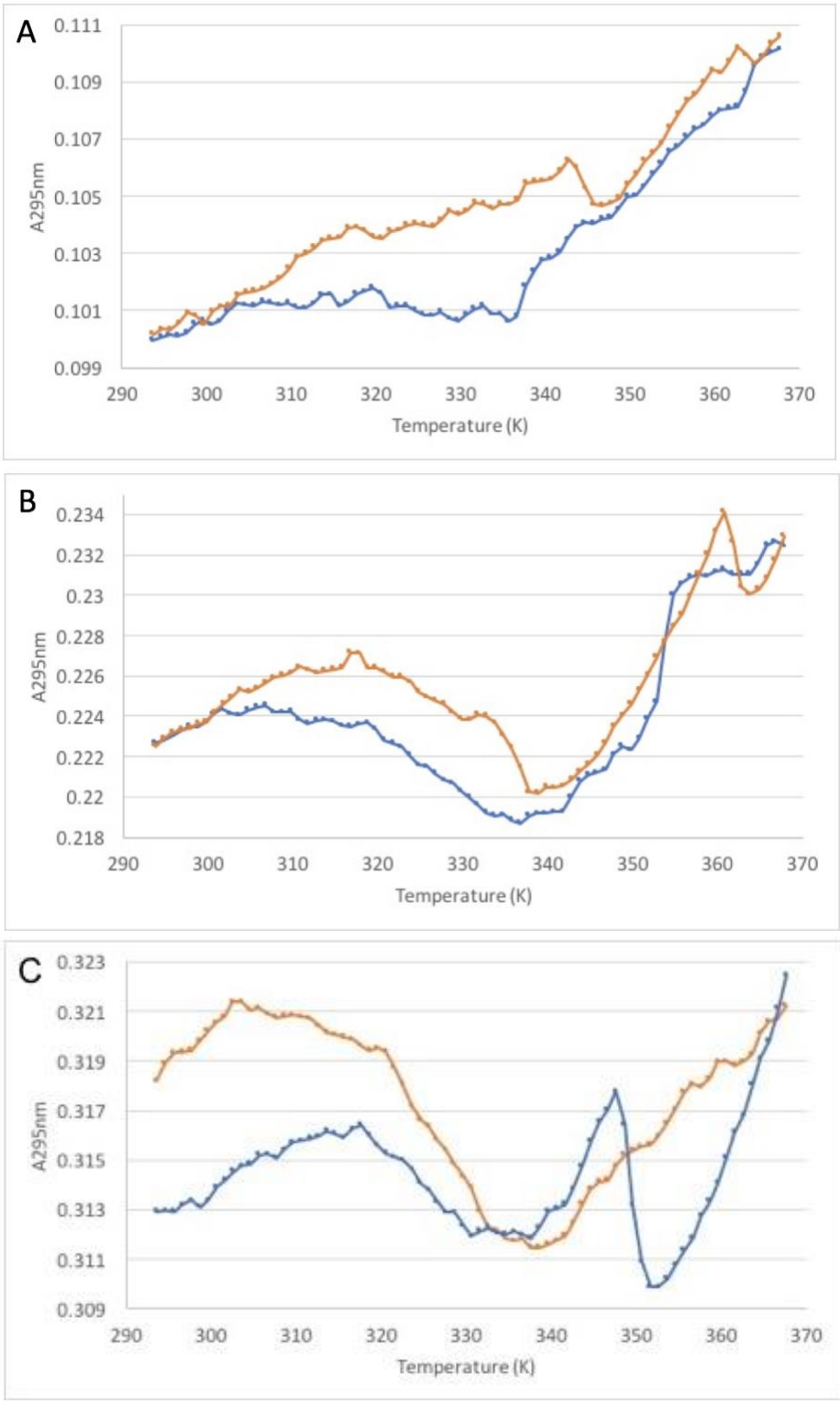

Figure 2.5. Representative Forward and Reverse Folding Curves of Three cMYCHS Concentrations. Blue is the forward cycle (95°C to 15°C), orange is the reverse cycle (15°C-95°C). (A) 2.5µM cMYCHS with 25mM KCl. (B) 5µM cMYCHS with 25mM KCl. (C) 7µM cMYCHS with 25mM KCl. A large hysteresis is seen in all samples.

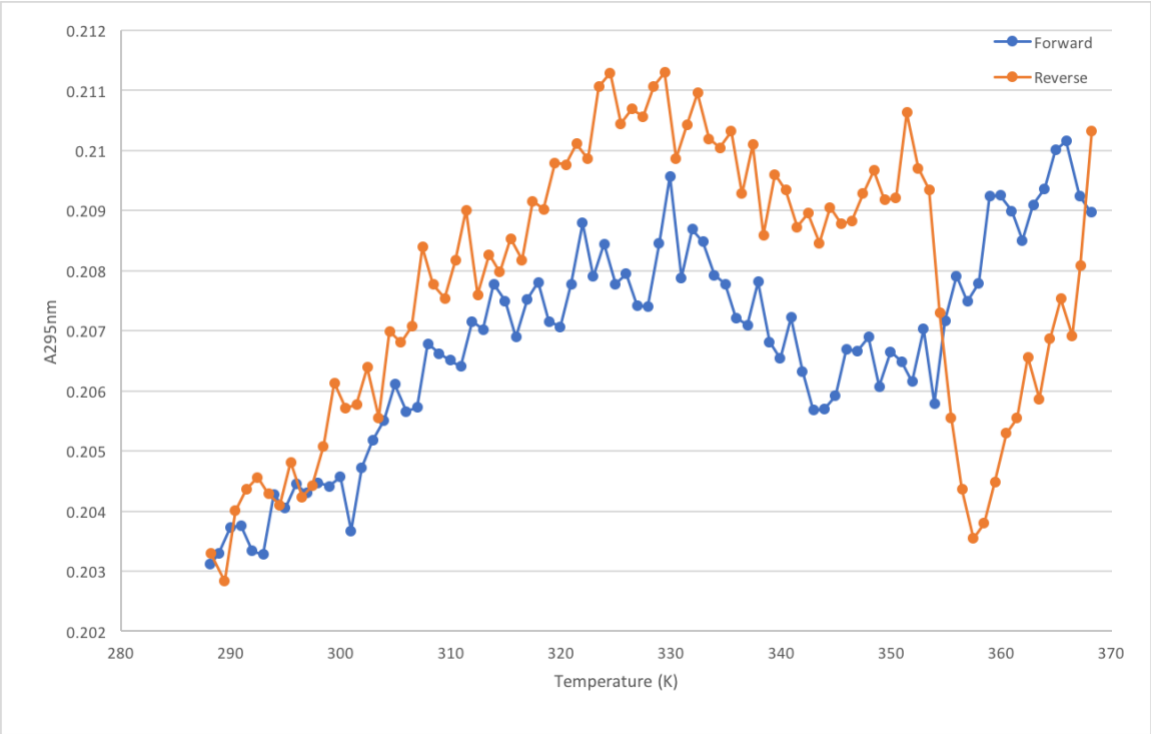

Figure 2.6. Folding Curve of 5µM cMYCHS with 100mM KCl. Blue data points are from the forward cycle 95°C-15°C), orange data points are from the reverse cycle (15°C-95°C). Points represent the average of two data sets.

Table 2.1. Points Used for Hysteresis Calculations.

Sample Type	Oligonucleotide Concentration (μM)	KCl Concentration (mM)	Ramp rate (°C/min)	Forward Points (°C)	Reverse Points (°C)	Average Difference in Absorbance
cMYC22m	2.5	15	1	341.6, 343.6, 345.6, 347.6, 348.6	343.7, 346.7, 349.7, 350.7, 351.7	$2.9*10^{-4}$
cMYC22m	5	15	1	339.7, 342.7, 343.7, 347.7, 348.7	345.6, 348.6, 350.6, 354.6, 356.6	$6.7*10^{-4}$
cMYC22m	10	15	1	336.8, 338.8, 342.8, 343.8, 346.8	341.5, 342.5, 346.5, 348.5, 351.5	$1.6*10^{-3}$
cMYCHS	2.5	25	1	338.12, 339.12, 342.12, 343.13, 344.11	309.19, 310.19, 313.19, 322.18, 333.19	$9.0*10^{-5}$
cMYCHS	5	25	1	319.14, 323.14, 326.13, 330.14, 331.15	329.67, 335.67, 336.67, 337.67, 338.68	$7.1*10^{-4}$
cMYCHS	7	25	1	324.64, 326.64, 327.64, 329.65, 330.65	317.16, 318.16, 320.16, 324.15, 326.16	$6.6*10^{-4}$
cMYCHS	5	100	1	329.9, 331.9, 332.9, 340.9, 343.9	339.4, 341.4, 343.4, 354.4, 355.4	$6.1*10^{-5}$

Table 2.2. Hysteresis of cMYC22m and cMYCHS from UV-melts at Various Oligonucleotide Concentrations.

Sample type	Oligonucleotide Concentration (μM)	KCl concentration (mM)	Ramp rate (°C/min)	Average Hysteresis (°C)
cMYC22m	2.5	15	1	3.1
cMYC22m	5	15	1	6.7
cMYC22m	10	15	1	4.3
cMYCHS	2.5	25	1	23.7
cMYCHS	5	25	1	9.7
cMYCHS	7	25	1	6.7
cMYCHS	5	100	1	10.9

Fluorescence

Fluorescence intensity was measured for FAM and FRET signals of a sample of 100nM F-cMYC22m-T with 10mM, 15mM, and 20mM KCl (Figure 2.7). Using Van't Hoff analysis for the forward measurements of the FAM signal, the T_m for the 10mM, 15mM, and 20mM KCl samples was calculated to be 64.22°C ±0.557°C, 67.52°C ±0.08°C, and 68.9°C ±0.115°C, respectively (Figure 2.7). This same experiment was performed with 100nM F-cMYCHS-T with 10mM and 15mM KCl (Figure 2.8). This data showed a large hysteresis in both the FAM and FRET signals, which was not seen in the cMYC22m sample. The $T_{1/2}$ was estimated to be 44°C for the 15mM KCl sample and 42°C for the 10mM KCl.

The hysteresis of T_m for each sample is reported in Table 3.3. This data quantifies the hysteresis seen in Figure 2.8 A and C. The hysteresis seen in cMYCHS samples is significantly greater than that seen for cMYC22m.

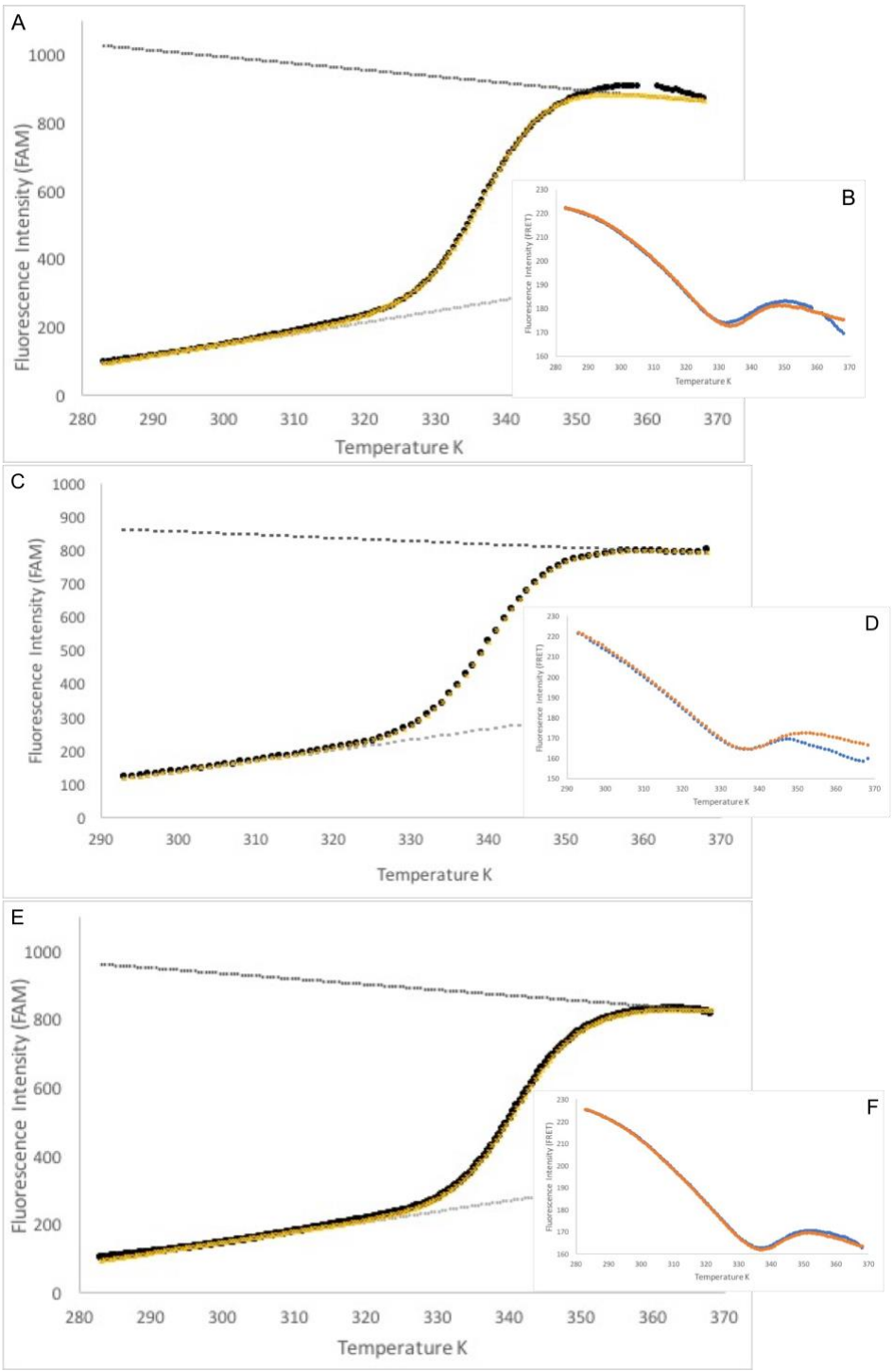

Figure 2.7. FAM and FRET Folding Curves for F-cMYC22m-T with Ranging KCl Concentrations Samples of F-cMYC22m-T (100nM) were analyzed by monitoring FAM fluorescence (A, C, E) and FRET emission (B, D, F) in the presence of 10mM (A and B), 15mM (C and D) or 20mM (E and F) KCl. Ramp rates were 0.5°C/min (A, B, E, and F) or 1°C/min (C and D) Dashed lines represent folded (light gray) and unfolded (dark gray) baselines used for Van't Hoff analysis. Black circles are forward fluorescence data collected, yellow Xs represent Van't Hoff analysis best fit line. Blue data points are from the forward run, orange data points are from the reverse run.

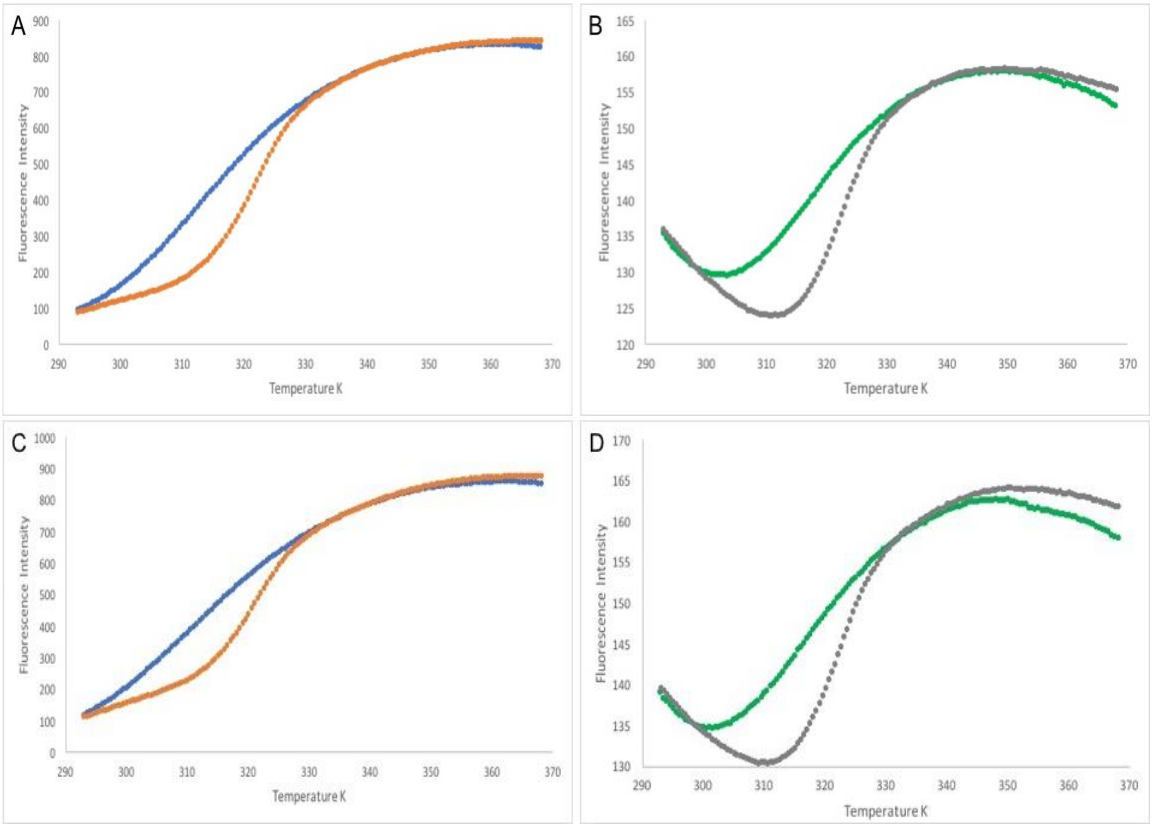

Figure 2.8. FAM and FRET Folding Curves for F-cMYCHS-T 15mM and 10mM KCl. Samples of F-cMYCHS-T (100nM) were analyzed by monitoring FAM fluorescence, forward (blue) and reverse (orange) curves for FAM (A and C) and FRET emission, forward (green) and reverse (gray) curves (B and D) in the presence of 15mM (A and B) or 10mM (C and D). Ramp rates were 0.5°C/min.

Table 3.3. T_m Hysteresis of cMYC22m and cMYCHS with Various KCl Concentrations.

Sample type	Oligonucleotide Concentration (nM)	KCl concentration (mM)	Ramp rate (°C/min)	T_m Hysteresis (°C)
cMYC22m	100	10	0.5	0.5
cMYC22m	100	15	1	0.9
cMYC22m	100	20	0.5	0.5
cMYCHS	100	10	0.5	7.8
cMYCHS	100	15	0.5	6.6

Circular dichroism

CD measurements were taken for a 6μM sample of cMYCHS with 25mM KCl. Measurements were taken at 20°C and 95°C in order to compare melted and folded, and the MRE was plotted against wavelength (Figure 2.9). These same conditions were also used to collect CD data over a range of six temperatures from 20°C to 95°C (Figure 2.10A). This data was plotted against temperature by averaging the points at 263-264nm and 239-241nm, which correspond to the positive and negative peaks, for each of the temperature runs used (Figure 2.10B). From the six temperature range data, few points were present in the transition slope for analysis of T_m. Therefore, CD measurements were also taken for another 8 temperatures for the 6μM cMYCHS sample (Figure 2.10C). The average of the peak points for this data was also plotted against temperature (Figure 2.10D).

CD measurements were also taken for 20μM samples of cMYCHS in 10mM Li-cacodylic buffer with EDTA and 25mM KCl. Data was collected over the range of wavelengths from 220-290nm for temperature points of 30°C, 45°C, 50°C, 55°C, 60°C, 70°C, 75°C, 85°C. This data was collected on two occasions and the average of the data points was taken, converted to MRE, and plotted against wavelength (Figure 2.11A). Just as was done for the 6μM sample, the average of the peak points was plotted versus temperature (Figure 2.11B).

The $T_{1/2}$ for both 6μM and 20μM was calculated by establishing folded and unfolded baselines. Due to the stabilized MRE between 20°C and 40°C seen in Figure 2.10B, it is assumed that the G-quadruplex would be folded at temperatures of 30°C or

lower. Therefore, for both samples the folded baseline was determined to be the point at the lowest temperature, 30°C. The unfolded baseline for the 6µM and 20µM samples was established as the point at the higher temperature. A linear line for the transitional portion of the graph was then plotted, and the temperature which was position at the halfway point of the transitional line between the two baselines was assigned as the $T_{1/2}$ of the sample (Figure 2.12). Approximate $T_{1/2}$ values for 6µM and 20µM cMYCHS were found to be 52°C and >60.5°C respectively. The $T_{1/2}$ value is estimated to be greater than 62°C due to the lack of a clear unfolded baseline.

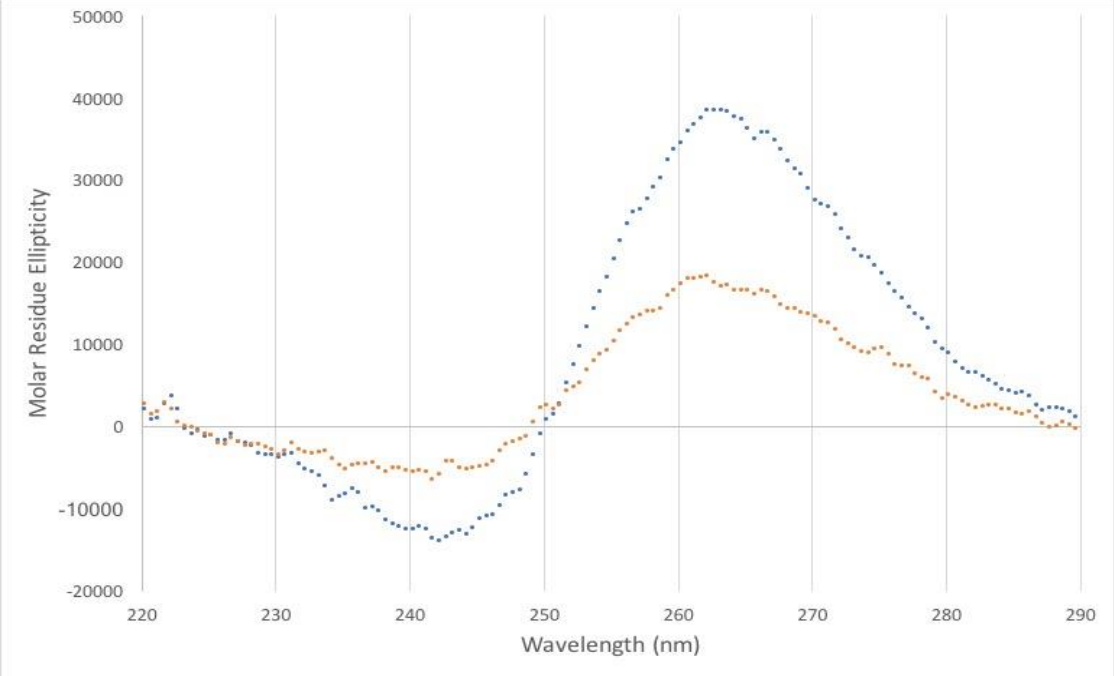

Figure 2.9. CD of 6µM cMYCHS 20°C and 95°C. Blue data points are from 20°C, orange data points are from 95°C. Positive 260nm peak and negative 240nm peak, with a large decrease in absorbance going from 20°C to 95°C.

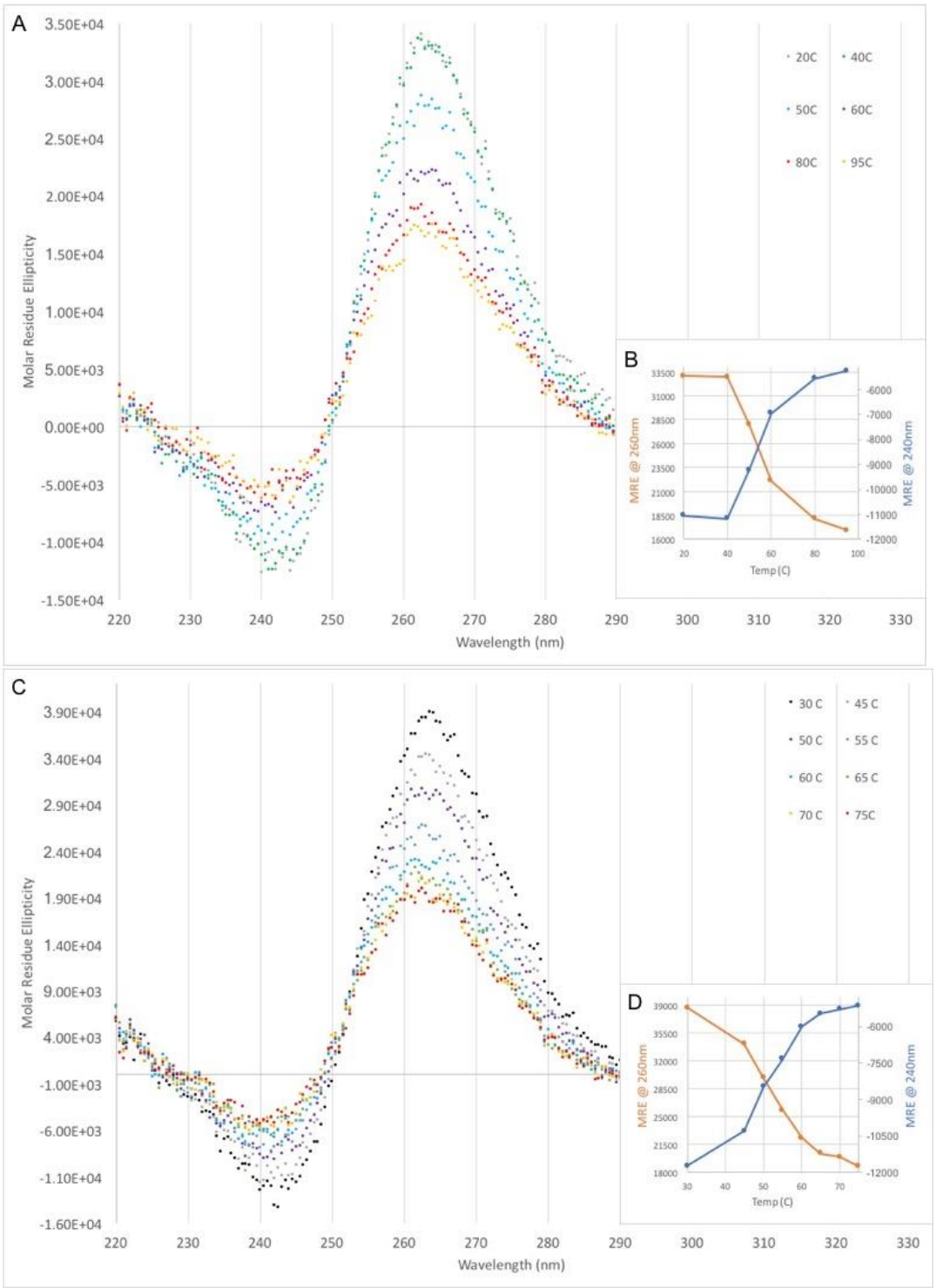

Figure 2.10. CD of 6µM cMYCHS Range of Temperatures. CD data collected from 220nm to 290nm and converted to MRE. (A) Data over 6 temperatures (20°C, 40°C, 50°C, 60°C, 80°C, and 95°C) smoothed by averaging each data point with the one preceding it. (B and D) MRE measurements at from 263-264nm averaged and plotted against temperature shown in orange, 239-241nm measurements averaged and plotted against temperature shown in blue. (C) Data over 8 temperatures (30°C, 45°C, 50°C, 55°C, 60°C, 65°C, 70°C and 75°C) smoothed by averaging each data point with the one preceding it.

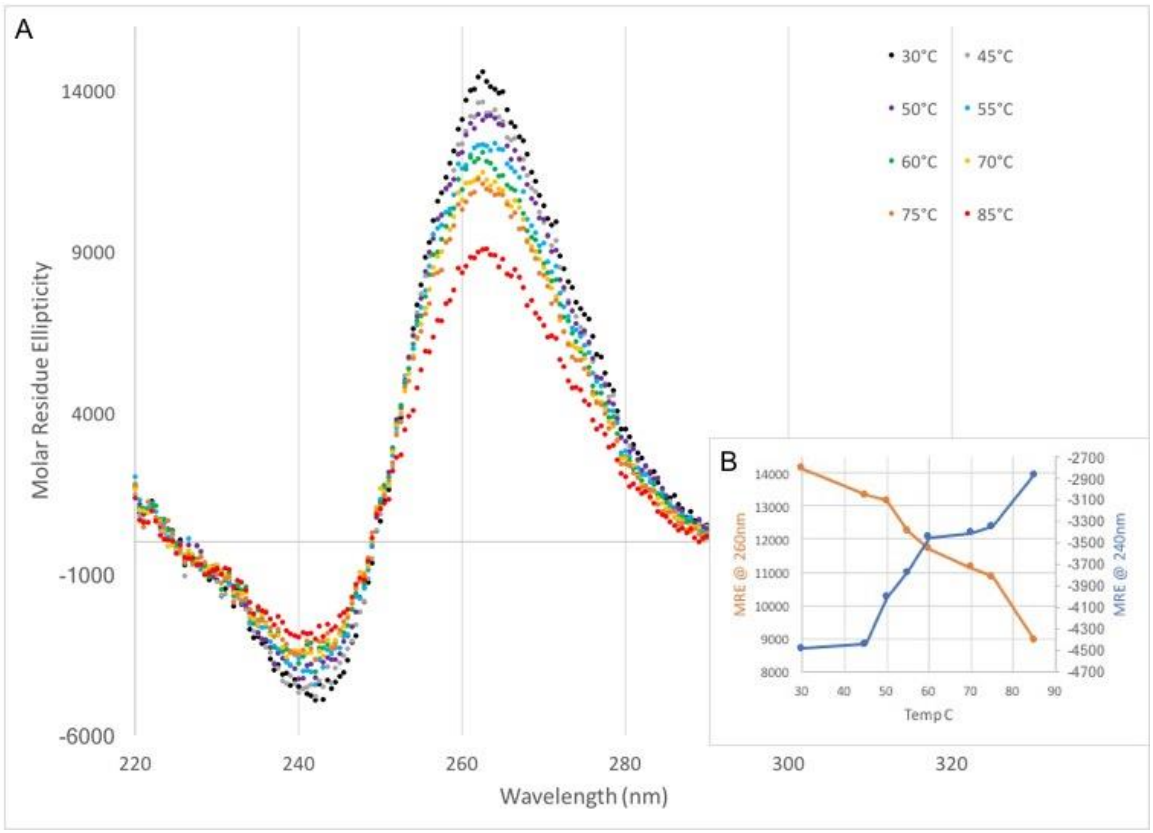

Figure 2.11. CD of 20µM cMYCHS Range of Temperatures. Average of two CD data sets collected from 220nm to 290nm and converted to MRE. (A) Data over 8 temperatures (30°C, 45°C, 50°C, 55°C, 60°C, 70°C, 75°C and 85°C). (B) MRE measurements at from 263-264nm averaged and plotted against temperature shown in orange, 239-241nm measurements averaged and plotted against temperature shown in blue.

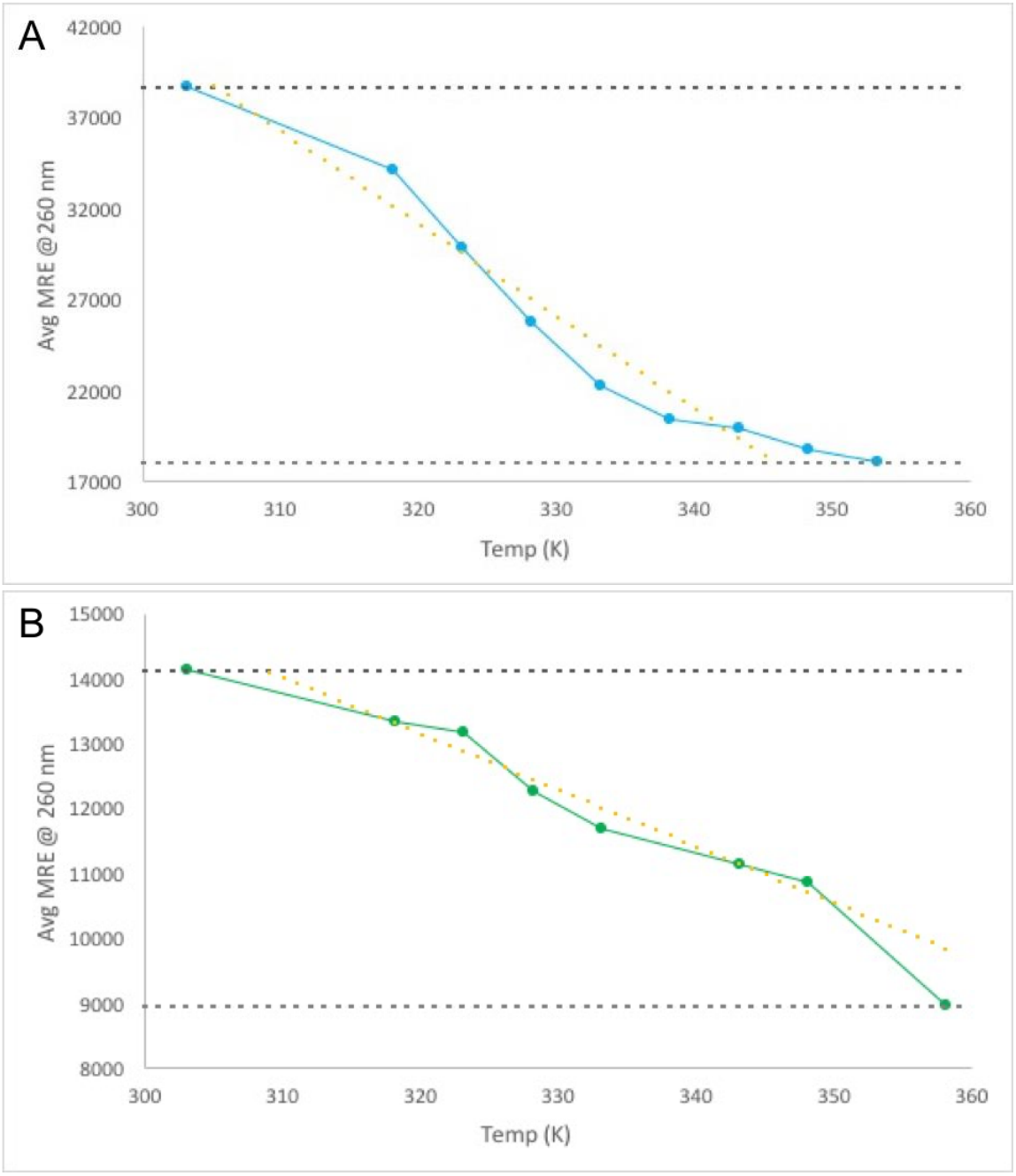

Figure 2.12. $T_{1/2}$ of cMYCHS from CD. Gray dashed lines represent folded (dark gray) and unfolded (light gray) baselines, yellow dotted line is the trend line used to calculate $T_{1/2}$. (A) 6µM cMYCHS $T_{1/2}$ for 260nm. (B) 20µM cMYCHS $T_{1/2}$ for 260nm

2.4 Discussion

UV-melting data was collected for a control, cMYC22m, and the experimental oligonucleotide, cMYCHS. The melting temperature was calculated for cMYC22m for different oligonucleotide concentrations as well as multiple salt concentrations. Melting temperatures for cMYC22m increased as the KCl concentration was increased, which was expected due to the stabilization effect potassium has on G-quadruplex structure. Under fixed KCl concentration, the T_m of the control cMY22m oligonucleotide remained constant. Thus, the data obtained for the control oligonucleotide are consistent with the formation of an intramolecular G-quadruplex. The oligonucleotide cMYCHS, however, did not produce expected results for an intramolecular G-quadruplex. As seen in Figure 2.4-2.6, cMYCHS UV-melts showed multiple peaks, and did not conform to the normal sigmoidal curve seen in intramolecular G-quadruplexes. These data also showed a large hysteresis between forward and reverse temperature cycles, which is consistent with folding intermediates.[12, 47] This data was inconclusive and due to the atypical curves was difficult to analyze, therefore, fluorescence was then utilized.

As a G-quadruplex labeled with FAM and TAMRA melts, and the distance between the fluorophores increases, the FAM becomes less quenched. This results in increased FAM emission and decreased FRET signal, which was seen in the F-cMYC22m-T melts. However, this same system was applied to F-cMYCHS-T and an increase in both FAM and FRET emissions were seen. When the F-cMYCHS-T FRET signal is compared to that of F-cMYC22m-T it is clear that at the low temperatures, where a G-quadruplex should be folded, the FRET signal of F-cMYC22m-T exceeds that of F-cMYCHS-T by about 60%. This indicates that the fluorophores are at a greater

distance in the F-cMYCHS-T system than in the intramolecular G-quadruplex, F-cMYC22m-T. The larger distance could be due to the formation of a different topology of intramolecular G-quadurplex for cMYCHS versus cMYC22m, or the formation of intermolecular G-quadruplex from cMYCHS. Additionally, a large hysteresis was seen for both FAM and FRET data. In order to further characterize the cMYCHS G-quadruplex folding CD was used.

The CD data for both 6μM and 20μM showed a positive peak at approximately 260nm and a negative peak at 240nm, which is indicative of a parallel stranded G-quadruplex. These data also exhibited a clear decrease in MRE as the sample was heated. The melting temperatures for these two concentrations of oligonucleotide were initially calculated using Van't Hoff analysis. These T_m values, however, were found to be significantly different from each other, demonstrating the G-quadruplex may not form unimolecularly. Due to the assumed two state function in the Van't Hoff, the data was reanalyzed as seen in figure 2.12. The $T_{1/2}$ values estimated for the two concentrations of oligonucleotide were found to have an approximate 10°C difference. These findings along with the large hysteresis seen in both UV and FRET melts led us to the conclusion that the G-quadruplex forms from intermolecular interactions.

This outcome has also led us to infer that the cMYCHS G-quadruplex is likely not in competition with the H-DNA in the c-*MYC* hotspot region. A collaborating lab has used a cMYCHS oligonucleotide extended with the proper sequence for formation of an intramolecular H-DNA structure in order to compare formation of H-DNA and G-quadruplex under two different conditions, H-DNA forming (with Mg^{++}) or G-quadruplex forming (with K^{+}).[6] Using multiple experimental methods, the study found

that the oligonucleotide formed a stable, intramolecular, anti-parallel H-DNA structure under H-DNA conditions.[6] Alternatively, under G-quadruplex conditions, the oligonucleotide formed multiple, less well characterized species, which arose from intra- or intermolecular interactions.[6] This study suggested that the H-DNA and G-quadruplex structures are likely not in completion. Our study supports this finding by demonstrating that the G-rich portion of the cMYCHS does not form a single, stable *intra*molecular G-quadruplex structure. Thus, there is no competition between formation of intramolecular G-quadruplex and H-DNA for the extended cMYCHS construct. Investigation of the H-DNA folding is ongoing.

III.G-QUADRUPLEX PHOTOCLEAVAGE BY NMM

3.1 Background

As discussed in chapter II, the NHE of *c-MYC* is a G-rich region that contains many putative non-canonical secondary structures. One of these non-canonical DNA structures is a G-quadruplex formed by the 5'TGAGGGTGGGGAGGGTGGGGA'3 sequence (Figure 2.1A). Previous studies revealed that this sequence is difficult to characterize, due to its propensity for folding into multiple topologies.[23] However, it was demonstrated that a guanine to thymine mutation at positions 11 and 20 constrained folding to a single parallel-stranded conformation.[23] Molecular dynamic calculations revealed that this mutated parallel-stranded G-quadruplex displayed similar stabilizing interactions and thermodynamic properties to the wild type.[23] Therefore, it is assumed that cMYC22m is acceptable for studying the *c-MYC* G-quadruplex as a target for ligand interactions.

With G-quadruplex structures being implicated in genome instability, it is important to detect these structures within the genome. A previous study of NMM has the ability to induce photocleavage of G-quadruplex structures was performed within our lab (Figure 3.1). This study revealed that NMM exhibits preferential cleavage of cMYC22m in the presence of potassium versus sodium.[37] Considering the formation of the single parallel topology in potassium[23], this preferential cleavage is interpreted as a higher affinity for the parallel stranded G-quadruplex structure over single stranded DNA or alternative G-quadruplex topologies of cMYC22m. NMM also preferentially cleaves cMYC22m over the 21mer human telomeric G-quadruplex.[37] The telomeric G-quadruplex, from the repeat sequence TTAGGG, has been shown to be polymorphic[49, 50],

consistent with the hypothesis that NMM displays specificity for some G-quadruplex topologies above others. This photocleavage agent has also shows an increase in photocleavage activity when the G-quadruplex is pre-annealed as opposed to co-annealed with NMM.[37] This suggests that the compound may induce a different topology when co-annealed and that the cleavage is more specific for the native topology of the cMYC22m G-quadruplex formed in potassium. Taken together, these characteristics suggest that NMM is a qualified candidate as a G-quadruplex probe. However, the mechanism of this cleavage has not yet been explored. Therefore, the goal of this project is to investigate the mechanism by which NMM cleaves cMYC22m upon irradiation with 420nm light. Previous investigations of cMYC22m photocleavage by NMM have been performed with 5'FAM and 3'TAMRA attachments. Porphyrin ligands have been shown to bind G-quadruplexes through planar stacking with the external G-quartets or by loop region interactions.[49] Therefore, the potential involvement of fluorophores attached to the ends of the G-quadruplex sequence must also be investigated.

To examine the potential involvement of fluorophores in the photocleavage reaction, a different type of fluorophore must be introduced. By performing the same experiments but in the presence of a different fluorophore, consistency in the photocleavage pattern between the two fluorophores can be assessed. The fluorophore employed for this investigation is BoDIPY, which is attached to the 5' end of cMYC22m by a copper-free click reaction. This type of click reaction utilizes azides and strained alkynes for a catalyst-free cycloaddition reaction.[51]

NMM shows preferential photocleavage at specific guanine sites of F-cMYC22m-T. Therefore, a guanine ladder is a useful tool in the comparison of 5'FAM 3'TAMRA to

5’BoDIPY. Dimethyl sulfate (DMS) is a compound which induces cleavage by methylating N7 of guanine, leaving an unstable glycosidic bond that can be broken with heat treatment.[34] A secondary agent, such as piperidine, is then able to cleave the sugar-phosphate bond, leading to a break in the backbone. Due to involvement of guanine N7 in Hoogsteen bonding of the G-quadruplex, this site may not be methylated in a folded G-quadruplex.[34] Therefore, the above Maxam-Gilbert method of creating a guanine DNA ladder will be utilized with denatured cMYC22m G-quadruplex to ensure more efficient cleavage at guanine sites. These G-ladders are to be run alongside photocleavage samples on polyacrylamide gels.

When attempting to crack the code of the photocleavage mechanism occurring between NMM and cMYC22m, there are a few tools to help decipher the cryptic pattern. After samples of cMYC22 with NMM are irradiated, they may be heat treated with piperidine in order to complete cleavage of the backbone. If a sample has been oxidized at the 8 position of guanine (8oxoG), which is the most nucleophilic base site, this may not cause complete cleavage of the DNA backbone. Therefore, comparison between samples which have been heat treated with piperidine and untreated samples will reveal the site of the reaction. Although both type I and type II photocleavage reactions may oxidize nucleobases, this piece of information is important in understanding the photocleavage reaction. Another tool which may assist in understanding the photocleavage mechanism is the use of D_2O as a solvent for the reactions. The use of D_2O as a solvent significantly extends the relatively short lifetime of the singlet oxygen, which is produced in type II photocleavage reactions.[52] Subsequently, if an increase in photocleavage is demonstrated in D_2O samples, the reaction likely occurs through type II.

Therefore, both piperidine heat treatment and a D_2O solvent will be employed in this investigation of the photocleavage mechanism of NMM with cMYC22m. A well-studied photocleavage agent, TMPyP4 (Figure 3.1.A), will also be used for comparison between its photocleavage pattern on cMYC22m and that of NMM (Figure 3.1B).

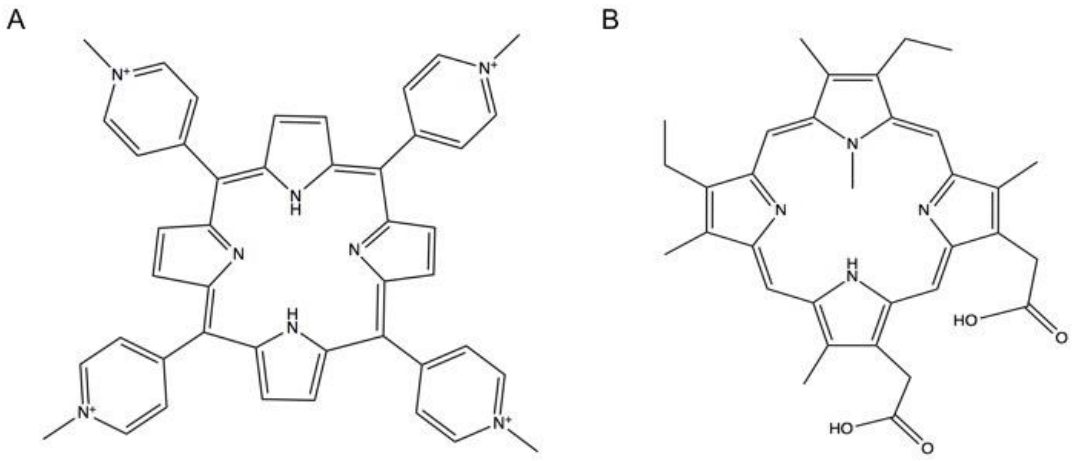

Figure 3.1. Structures of TMPyP4 and NMM. (A) Chemical structure of TMPyP4 (B) Chemical structure of NMM.

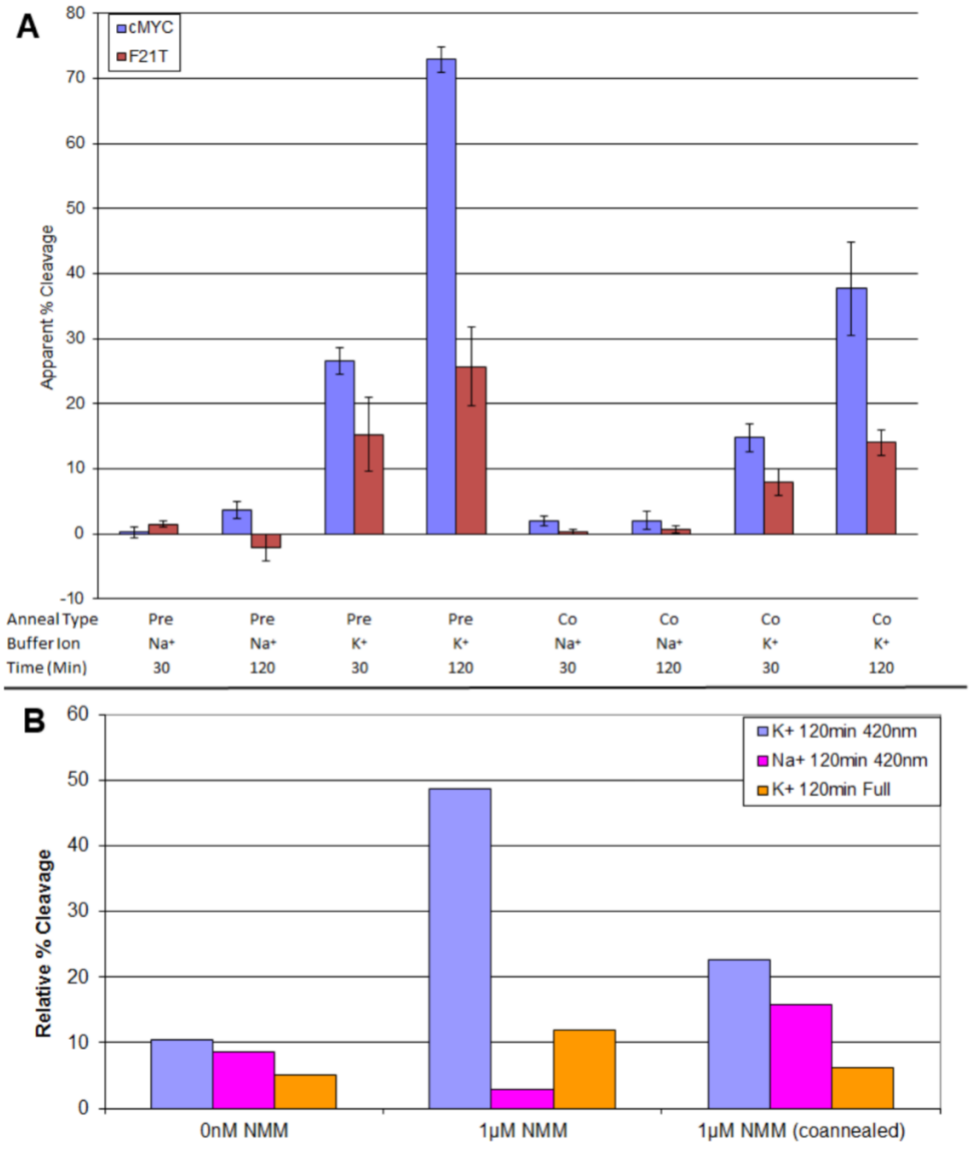

Figure 3.2. Photocleavage of Dual-labeled Oligonucleotides by NMM. Figure used with permission from Dominic McBrayer[37] (A) Comparison of 5'FAM-dGGG(TTAGGG)$_3$-TAMRA-3' (F21T) and F-cMYC22m-T photocleavage by NMM in buffers containing sodium or potassium with G-quadruplex annealed before treatment with NMM (pre) or annealed in the presence of NMM (co) at different time points. (B) Photocleavage of F-cMYC22m-T in sodium or potassium containing buffers with white light or 420nm light.

3.2 Materials and Methods

The short oligonucleotide 5'Azide-TGAGGGTGGGTAGGGTGGGTAG-3' (Azide-cMYC22m) was ordered from IDT and used without further purification. 7-(3-((3-Azidopropyl)amino)-3-oxopropyl)-5,5-difluoro-1,3-dimethyl-5H-dipyrrolo[1,2-c:2',1'-f][1,3,2]di was ordered from ABCL, and BDP-FL-DBCO from BraodPharm. BoDIPY-azide was modified in our lab to have a strained alkyne attached, SMK-T-66. Reagents obtained from Sigma-Aldrich include methyl red, methanol, urea, β-mercaptoethanol, transfer RNA, sodium hydroxide, xylene cyanol, dimethyl sulfate (DMS), and piperidine. Reagents supplied by Fisher Scientific include isopropanol and formamide. Acrylamide and sodium acetate were procured from VWR. Bisacrylamide and TEMED were purchased from MP Biomedicals. Agarose and bromophenol blue were obtained from Bio Rad, ammonium persulfate (APS) from Promega, sodium cacodylate from Fullam, EDTA from Baker, O'RangeRuler 5bp DNA ladder from ThermoFisher Scientific, GelGreen from Biotium, NMM from Frontier Scientific, and TMPyP4 from MidCentury.

Equipment used in these experiments is as follows. The Fisher Scientific microcenterfuge Accuspin Micro 17 was used with Microcon YM-3 regenerated cellulose 3000MW filters. The NanoDrop 3000 manufactured by Fisher Scientific was used to test DNA concentrations. The Vacufuge Plus from Eppendorf was used for drying samples. The Luzchem Photoreactor was used for photocleavage reactions, with 420nm lamps from Rayonet. The Owl P10DS gel rig from Thermo Scientific powered by the Power Supply EC5000P from EC Apparatus Corporation was used for polyacrylamide gels. The Molecular Imager PharosFX, manufactured by Bio Rad, was used for gel imaging.

Click Reactions

For MR-cMYC22m click reactions, the Azide-cMYC22m was prepared to 20µM in milli-Q water and MR-alkyne was prepared in methanol to 22µM. For BoDIPY-cMYC22m click reactions, the Azide-cMYC22m was prepared to 10µM and BoDIPY-alkyne was prepared in 20% methanol 80% water to 100µM concentration. For either reaction the two solutions were mixed together in equal amounts and incubated at room temperature for 24hours. This mixture was then filtered through a 3000MW cellulose filter at 13000Xg for 50min. The filter was then inverted and spun again at 1000Xg for 3min. The filter was washed with 100µM milli-Q water, this wash was added to sample volume.

The concentration of the filtered samples calculated using the nanodrop at 260nm, and the molar extinction coefficient of cMYC22m provided by IDT was listed as 226300L/mol•cm. The samples were dried and resuspended in water to a final concentration of 10µM. These samples were then tested on a polyacrylamide gel to ensure the click reaction occurred. Prior to running on a gel, samples were diluted to 1µM in formamide buffer with 3µL dye added before running on gel.

Stock Solution Preparations

A stock solution of 10X TBE was prepared with 108g Tris base, 55g Boric acid, 7.5g EDTA with disodium salt, and filled to 900mL milli-Q H_2O. The solution was then brought to the pH of 8.3 using HCl then brought to final volume of 1000mL with Milli-Q H_2O.

Sodium cacodylate buffer with EDTA (SCE) was made as a stock solution of 250mM sodium cacodylate with 12.5mM EDTA brought to the pH of 8 using HCl or NaOH. Li-cacodylate buffer was prepared as described in chapter II, and KCl was added from a 1M stock in appropriate concentrations. These solutions were filtered through a 0.2µm cellulose filter.

Quench solution was prepared as a 1mL solution containing 1.5M sodium acetate at pH 7, 1M β-mercaptoethanol, 100ug/ml tRNA, and filled to 1mL with milli-Q water. Formamide loading buffer was made as 80% formamide, 10mM NaOH, and 1mM EDTA. Loading dyes xylene cyanol, bromophenol blue, and a mixture of the two were prepared as 10% solutions.

Maxam-Gilbert G-DNA Ladders

F-cMYC22m-T and BoDIPY-cMYC22m, from click reactions, were prepared to 10µM concentrations in milli-Q H_2O. 10µL of the oligonucleotide was added to 150µL 50mM SCE buffer. This sample was then heated to 95°C for 3min before DMS was added. For pure DMS samples, 1µL of DMS was added and pipette up and down then the sample was allowed 3.5min to incubate at room temperature before the addition of 100µL quench solution. For 10% DMS samples, 1.25µL of 10% DMS was added and pipette up and down then allowed 8min to incubate at room temperature before the addition of 100µL quench solution. After quench solution was added, 750µL of isopropanol was added and the sample was stored in -80°C for 1hour. The sample was then pelleted at 13000Xg for 10min, the supernatant was pulled off and the pellet was then washed with 500µL of ice-cold 70% isopropanol, centrifuged at 13000Xg for another 3min, and the supernatant was extracted (pelleting process). The pellet was then resuspended in 120µL

of 0.6M sodium acetate at pH 7 and 500µL of isopropanol was added before the sample was stored at -80°C overnight. After this incubation period the sample was subjected to the pelleting process again before being dried down. Following the speedvac drying of the sample, the sample was resuspended in 70µL of 10% piperidine, which had been freshly distilled, and incubated at 90°C for 30min. After piperidine treatment the samples were washed with 30µL milli-Q water, dried down, washed with another 20µL of milli-Q water before being dried down again. For storage of these samples, they were resuspended in 50µL milli-Q H_2O and divided into 10µL aliquots in 5 tubes. These tubes were subsequently dried down then stored in -20°C. Before running on a polyacrylamide gel, Maxam-Gilbert samples were re-suspended in 10µL formamide buffer with 3µL xylene cyanol dye added.

Photocleavage

Photocleavage samples were prepared in black Eppendorf tubes as 1µM F-cMYC22m-T 50mM Li-cacodylate buffer with 150mM potassium. Either 1µM NMM or 0.25µM TMPyP4 was added to samples. Negative control samples were also prepared with no ligand added. Samples that were to be run in D_2O instead of H_2O were prepared as described above, then dried down, and re-suspended in D_2O. Each sample type was loaded into a row on a black 364 well plate with 20µL/well. Each plate contained samples for only one time-point. The plates were placed into the photoreactor with four 420nm bulbs illuminated, they were allowed either 60, 90, or 120min in the photoreactor. After the appropriate amount of time, the plates were removed and half of each sample was aliquoted into a black tube for treatment with piperidine, the other half was aliquoted

into a black tube storage until run on a gel. All samples were dried down, and those not being treated with piperidine were stored at 4°C.

A 10% solution of freshly distilled piperidine was prepared. A hot water bath was heated to 90°C. Sample lids were taped down and the samples were placed in the hot water bath to incubate for 30min. Following piperidine treatment, samples were dried down and placed in 4°C for storage. All samples were re-suspended in 20µL formamide buffer with 6µL 10% xylene cyanol dye added. Each sample was loaded onto the gel in 12µL amounts.

Polyacrylamide Gel Electrophoresis (PAGE)

A 40% stock solution of 38:2 acrylamide:bisacrylamide solution was prepare by adding 190g acrylamide, 10g bisacrylamide, 50mL 10X TBE and milli-Q water to reach 500mL. A 2% agarose plug was used for gel casting. Denaturing gels were prepared as 60mL of 20% 38:2 acrylamide:bisacrylamide, with 1X TBE, 7M urea, and topped off to 60mL with milli-Q water. This solution was then filtered through a 0.2µm cellulose filter. 200µL of TEMED was added along with 200µL of a 10% APS solution, which was made fresh for each gel.

The following conditions apply to PAGE experiments unless otherwise stated. PAGE was performed with 1X TBE running buffer. The gel was pre-run at 6W for 30min before loading samples. Wells of the gel were then blown out using running buffer, and samples were loaded. After samples were loaded, the gel was run at 6W for 30min, then 4W for 6hours.

Gels were imaged with excitation of 488nm and emission of 530nm (FITC) and/ or excitation of 532nm and emission of 605nm (TAMRA). For gels that contained the

5bp ladder, they were stained with GelGreen for 30min and had triplicate rinses with H_2O for 10min. Stained gels were also imaged on FITC setting.

3.3 Results

Click Reactions

In order to determine their suitability as fluorescent labels for DNA, fluorescent scans of methyl red and two BoDIPY-derived labels were conducted. It was found that methyl red had no fluorescent activity; however, the two BoDIPY-derived compounds displayed good fluorescence signals. The BoDIPY-azide modified in our lab to have a strained alkyne attached, SMK-T-66, showed a similar fluorescence spectra to to BDP-FL-DBCO, a commercially available BoDIPY-based strained alkyne (Figure 3.2). With an excitation wavelength of 488nm, both samples showed peak emission at approximately 510nm. Therefore, both BoDIPY constructs were used as a fluorescent tag for cMYC22m.

Click reactions were preformed between Azide-cMYC22m and either BDP-FL-DBCO or SMK-T-66. After click reactions were performed, they were purified through a spin column and tested on 20% polyacrylamide gels. These gels were performed in order to visualize the reaction and should have shown a single band in the denaturing polyacrylamide gel. However, multiple spots are visible on these gels (Figure 3.3A and B). The imager detection limit for this sample was also tested by running decreasing amounts of Bo-DIPY-cMYC22m (Figure 3.3C). The detection limit is approximately 1pmol for this sample.

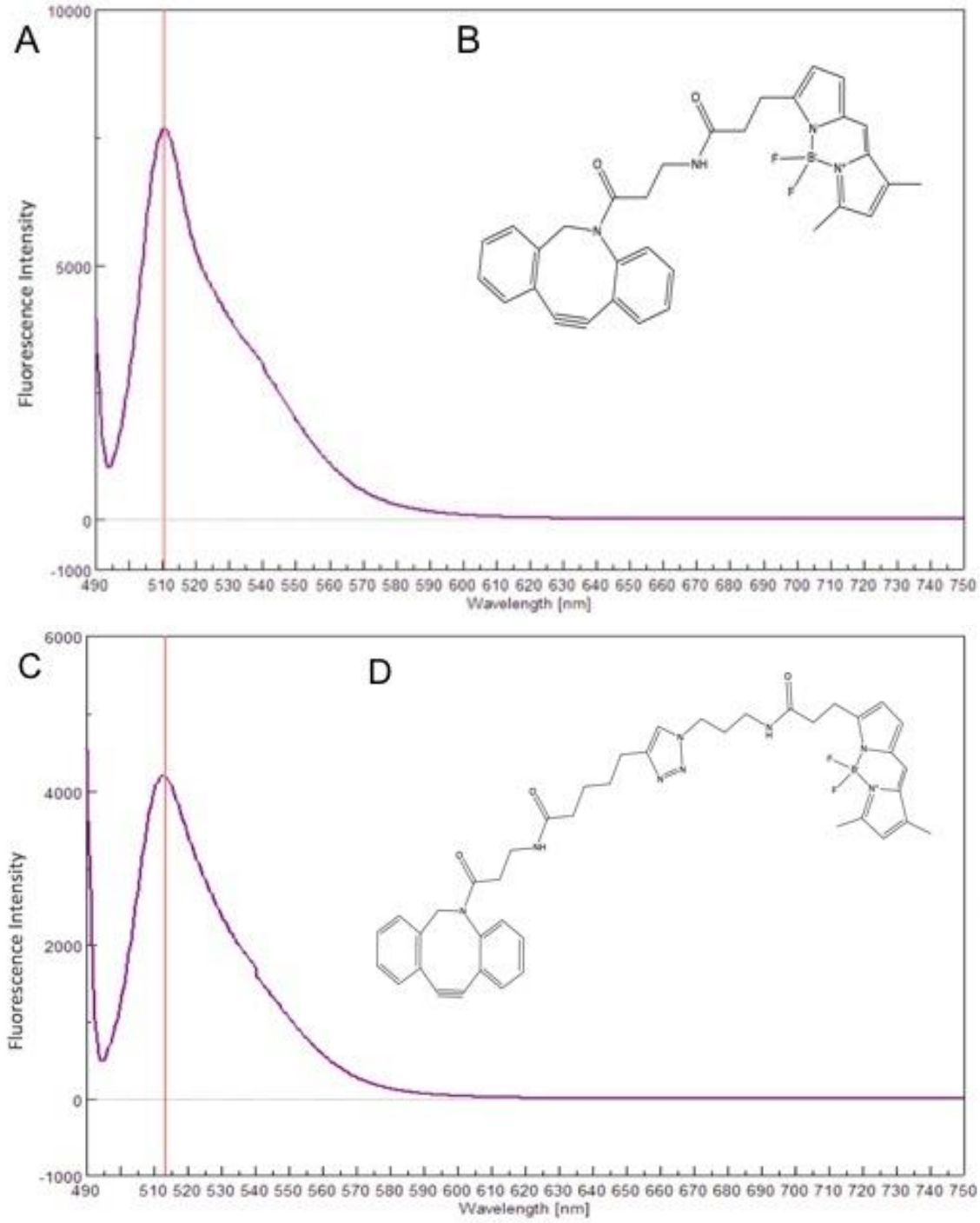

Figure 3.3. Fluorescence Spectra of BoDIPY. (A and C) Fluorescence emission spectra of BoDIPY at excitation wavelength of 488nm. (A and B) BDP-FL-DBCO (C and D) SMK-T-66

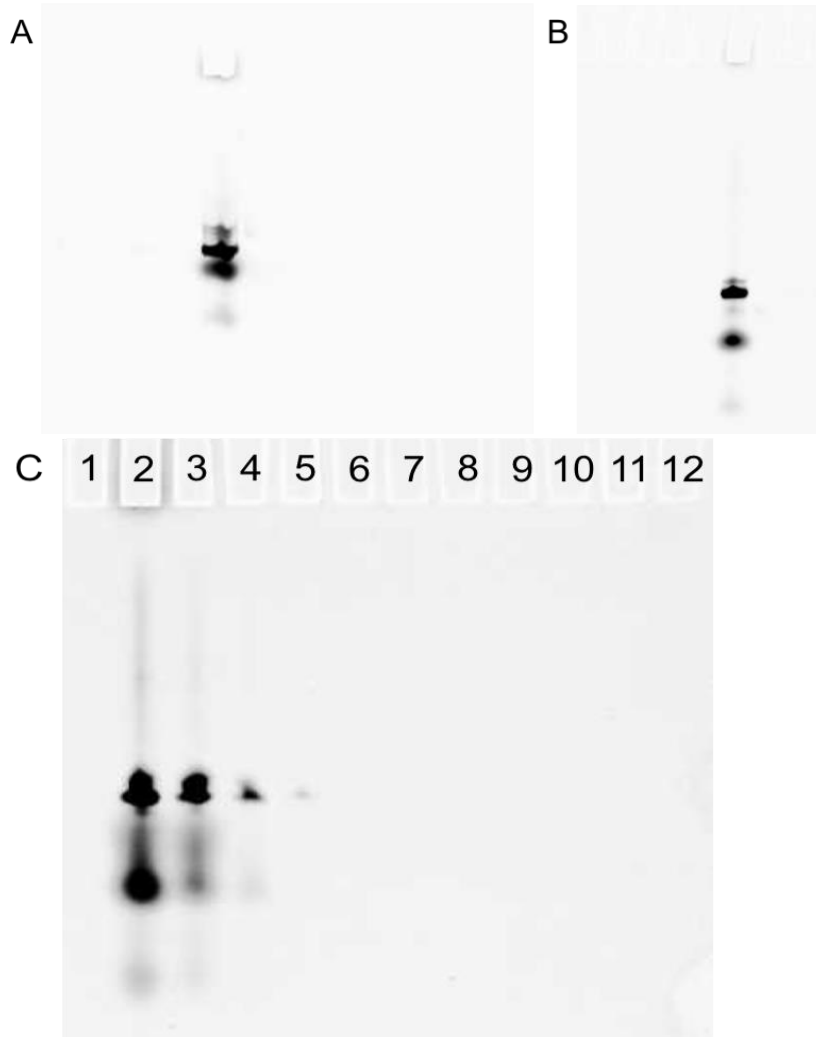

Figure 3.4. Click Reaction of cMYC22m with BoDIPY. (A) 5µM BoDIPY-cMYC22m, using SMK-T-66, run at 10W for 30min and 8W for 2hours (B) 5µM BoDIPY-cMYC22m, using SMK-T-66, run at 10W for 30min and 8W for 2.5hours (C) Click reaction was performed using BDP-FL-DBCO run at 10W for 30min and 8W for 2.5hours; Lane 1: empty, Lane 2: 2µM BoDIPY-cMYC22m, Lane 3: 1µM BoDIPY-cMYC22m, Lane 4: 100nM BoDIPY-cMYC22m, Lane 5: 10nM BoDIPY-cMYC22m, Lane 6: 1nM BoDIPY-cMYC22m, Lane 7: 100pM BoDIPY-cMYC22m Lanes 8-12: empty

Maxam-Gilbert Reactions

Maxam-Gilbert reactions were performed for both BoDIPY-cMYC22m and F-cMYC22m-T. These reactions were originally performed with only 50µL quench solution, which contains sodium acetate, and with 95% ethanol instead of 100% isopropanol. Samples were not detectable in these conditions (Figure 3.5). Therefore, conditions were adjusted to those listed in section 3.2. Adjusted conditions resulted in a

successful G-ladder for F-cMYC22m-T (Figure 3.6 Lanes 6 and 7). Due to poor resolution near the 5' end of the sequence, it is difficult to distinguish some of the guanine cleavage sites. The 100% and 10% DMS reactions of F-cMYC22m-T were equally successful, therefore, both the reactions are used for photocleavage gels. The BoDIPY-cMYC22m Maxam-Gilbert reactions, however, were not visible (Figure 3.5 Lanes 9 and 10).

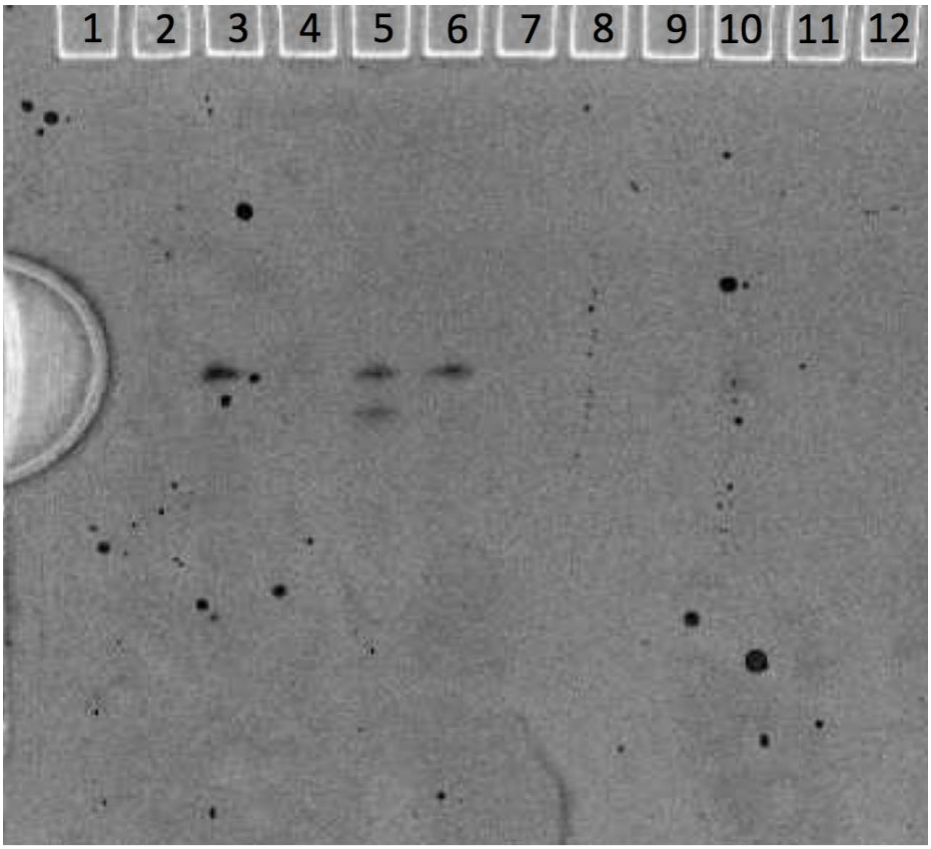

Figure 3.5. Maxam-Gilbert of F-cMYC22m-T and BoDIPY-cMYC22m with 100% and 10% DMS; Original Conditions. Original Maxam-Gilbert conditions of 50µL quench solution and 95% ethanol. Lane 1: empty, Lane 2: 3µL of 10% xylene cyanol and bromophenol blue in 10µL formamide buffer, Lane 3: 100% DMS F-cMYC22m-T, Lane 4: 10% DMS F-cMYC22m-T 5min, Lane 5: 10% DMS F-cMYC22m-T 8min, Lane 6: 10% DMS F-cMYC22m-T 12min, Lane 7: empty, Lane 8: 10% DMS BoDIPY-cMYC22m 5min, Lane 9: 10% DMS BoDIPY-cMYC22m 8min, Lane 10: 10% DMS BoDIPY-cMYC22m 12min, Lane 11-12: empty

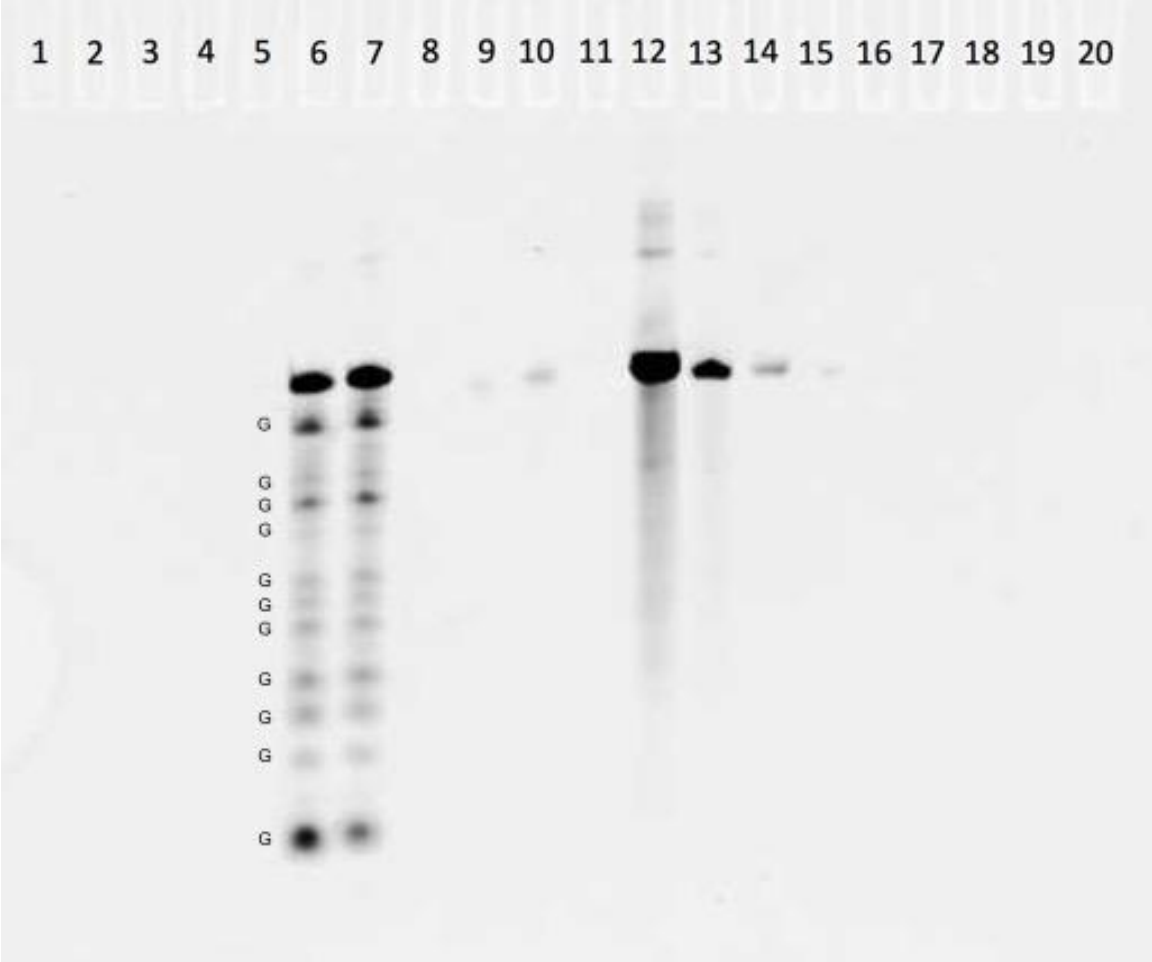

Figure 3.6. Maxam-Gilbert G-ladder of F-cMYC22m-T and Detection Limit. Lane 1-3: empty, Lane 4: 3µL of 10% xylene cyanol and bromophenol blue in 10µL formamide buffer, Lane 5: empty, Lane 6: 100% DMS F-cMYC22m-T, Lane 7: 10% DMS F-cMYC22m-T 8min, Lane 8: empty, Lane 9: 100% DMS BoDIPY-cMYC22m using DBP-FL-DBCO, Lane 10: 10% DMS BoDIPY-cMYC22m using DBP-FL-DBCO, incubated 8min, Lane 11: empty, Lane 12: 1µM F-cMYC22m-T, Lane 13: 100nM F-cMYC22m-T, Lane 14: 10nM F-cMYC22m-T, Lane 15: 1nM F-cMYC22m-T, Lane 16: 100pM F-cMYC22m-T, Lane 17-20: empty.

Photocleavage Reactions

Photocleavage reactions were performed on F-cMYC22m-T. These reactions were conducted with both NMM and TMPyP4, as well as a negative control containing no ligand, with 420nm light irradiation (Figure 3.7). A BoDIPY-cMYC22m 1µM sample and Maxam-Gilbert reaction with the new conditions were also run on this gel (Bo and Bo G lanes, respectively). However, the BoDIPY-cMYC22m click reaction sample displayed the same problems as previously seen, and the Maxam-Gilbert reaction was not

visible for this sample. The photocleavage reactions of F-cMYC22m-T displayed successful cleavage by both TMPyP4 (Ligand type "T") and NMM (Ligand type "N"). The F/T G lane, which contains the G-ladder of F-cMYC22m-T was not fully visible, likely due to the contrast with the photocleavage samples which were at higher concentration.

The original resuspension of the photocleavage samples was into 20µL formamide with 6µL xylene cyanol dye, 12µL of each sample was used to load into the gel. In order to address the poor resolution of photocleavage samples and low visibility of the G-ladder, an additional 10µL of formamide and 3µL of xylene cyanol was added to the remaining photocleavage sample. PAGE was performed with the new sample dilution (Figure 3.8). The guanine ladder is visible in this figure for comparison with the photocleavage samples. This image reveals that although NMM and TMPyP4 have a preference for cleavage at guanines, cleavage occurs at the sites of other nucleobases as well. It is also apparent that NMM and TMPyP4 have similar cleavage patterns. Figures 3.7 and 3.8 also show piperidine heat treatment is required for better resolution of bands.

Photocleavage reactions with NMM and TMPyP4 were also done in D_2O in addition to H_2O. These reactions were irradiated for either 60min, 90min, or 120min. Due to the high intensity of photocleavage band, G-ladders were not visible with the automatic exposure. Therefore, for each photocleavage gel, images of the G-ladders were adjusted to a proper setting to allow for visible bands. These detectible G-ladder lanes were overlaid on the original gel images to allow for easier comparison (Figures 3.9, 3.10 and 3.11 A). In a visual comparison between D_2O and H_2O samples there is no obvious difference in band intensity for either ligand (Figures 3.9, 3.10, and 3.11 B). Reactions of

F-cMYC22m-T with NMM and TMPyP4 reveal many similarities in cleavage pattern. However, there is a notable difference in band intensity at guanines 17 and 22 in the sequence. Samples with NMM show more intense bands at these guanines than samples using TMPyP4 as the photocleavage agent. Photocleavage reactions irradiated for 60min or 90min instead of the maximal 120min also showed less intense cleavage bands, which is expected based on previous fluorescent studies[37].

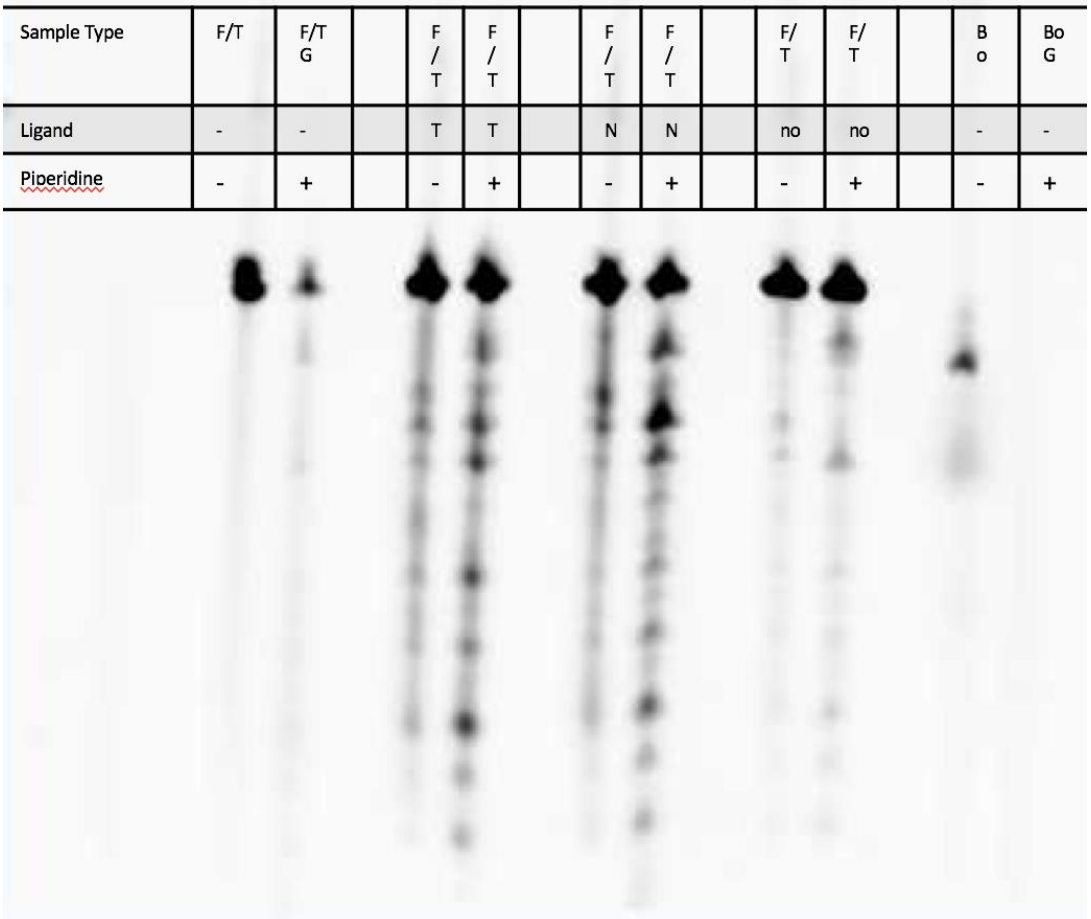

Figure 3.7. Photocleavage of F-cMYC22m-T by NMM and TMPyP4, and BoDIPY-cMYC22m Sample. Sample types include F-cMYC22m-T (F/T), F-cMYC22-T G-ladder (F/T G), BoDIPY-cMYC22m using DBP-FL-DBCO (Bo), and BoDIPY-cMYC22m G-ladder using DBP-FL-DBCO (Bo G). Ligands are TMPyP4 (T), NMM (N), or no ligand (no).

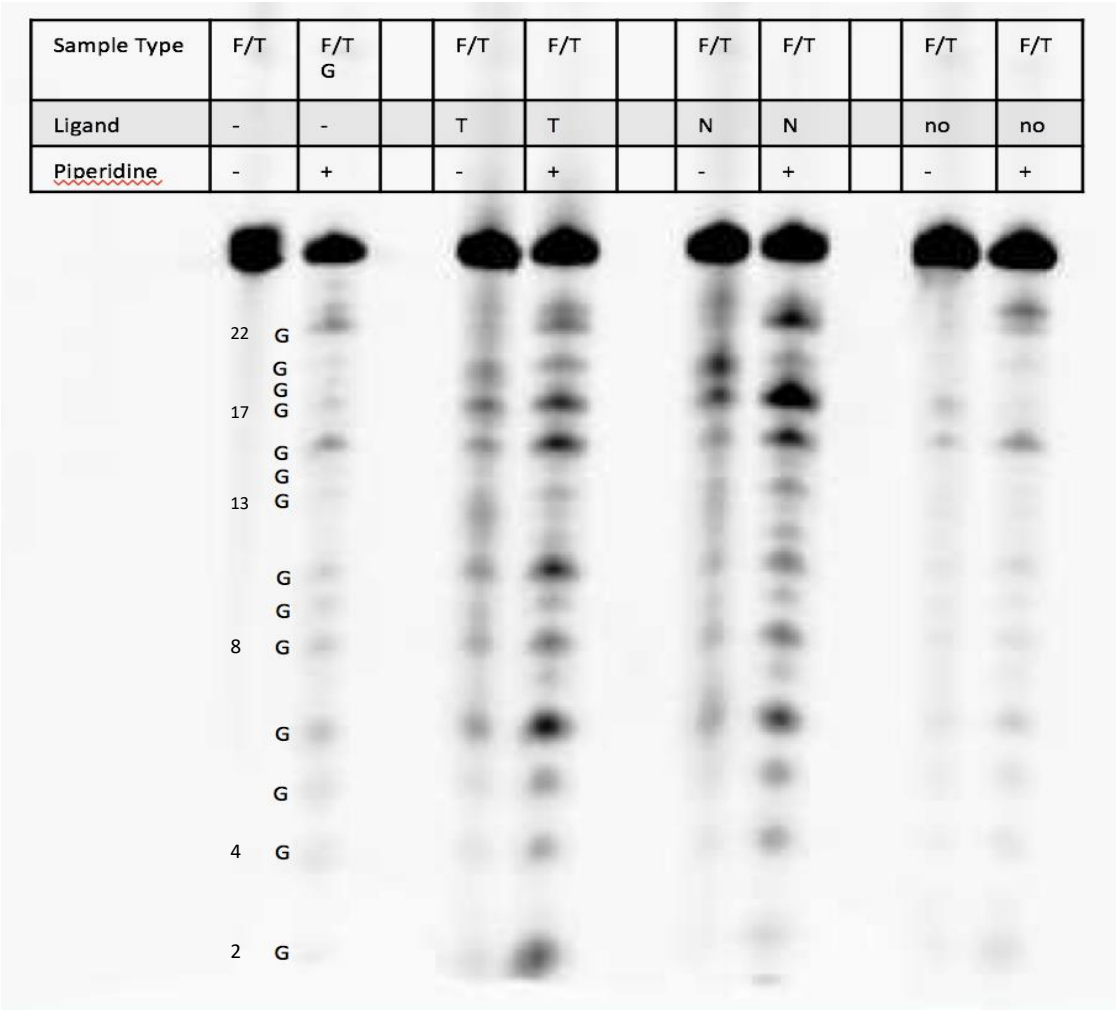

Figure 3.8. Photocleavage of F-cMYC22m-T by NMM and TMPyP4 with Guanine Ladder. Sample types include F-cMYC22m-T (F/T), F-cMYC22-T G-ladder (F/T G), BoDIPY-cMYC22m (Bo), and BoDIPY-cMYC22m G-ladder (Bo G). Ligands are TMPyP4 (T), NMM (N), or no ligand (no).

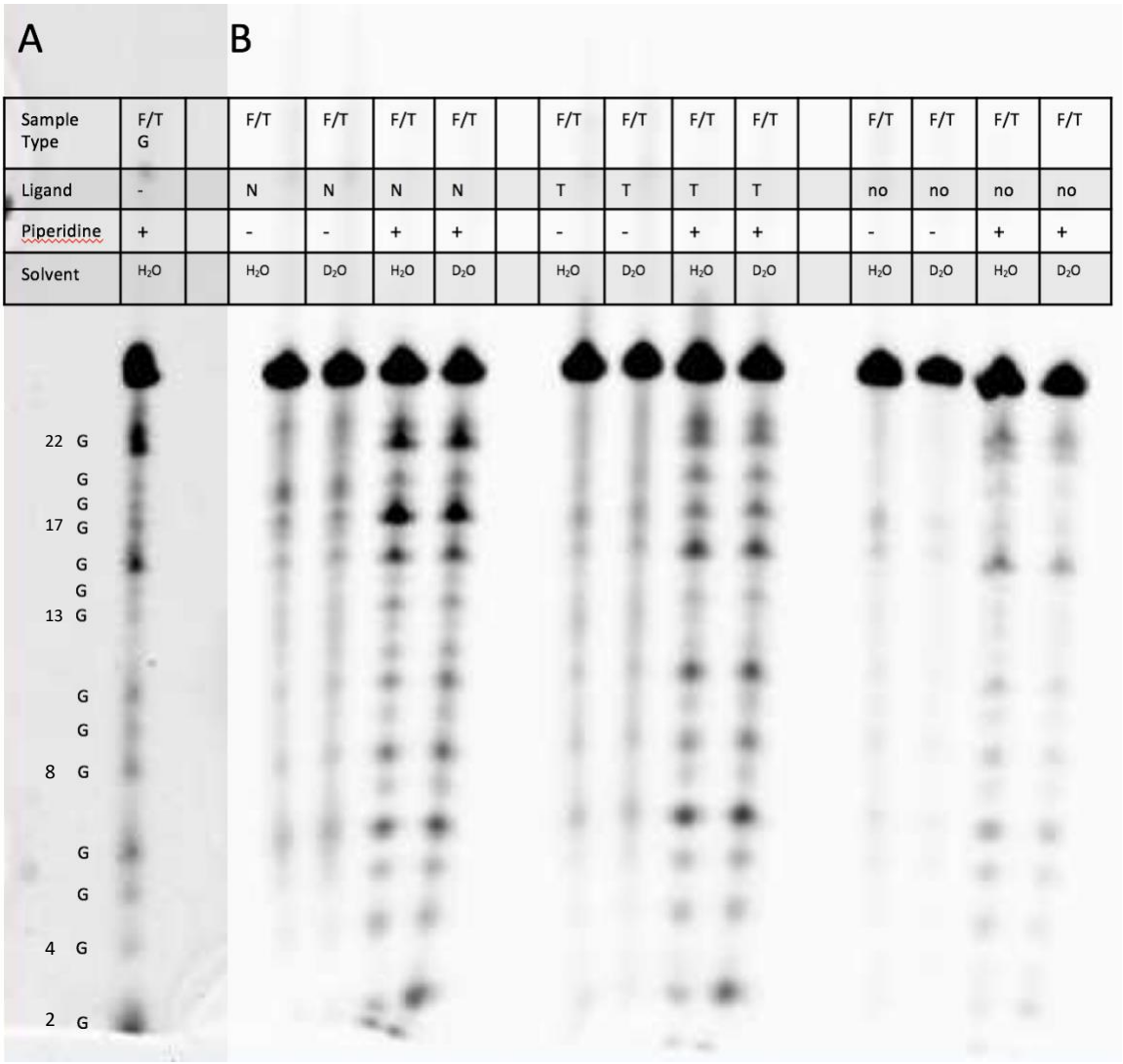

Figure 3.9. 120min Photocleavage of cMYC22m with NMM and TMPyP4, with G-ladder. Sample types include F-cMYC22m-T (F/T) and F-cMYC22-T G-ladder (F/T G). Ligands are TMPyP4 (T), NMM (N), or no ligand (no). (A) Adjusted image of G-ladder lane (B) Image taken at low intensity setting.

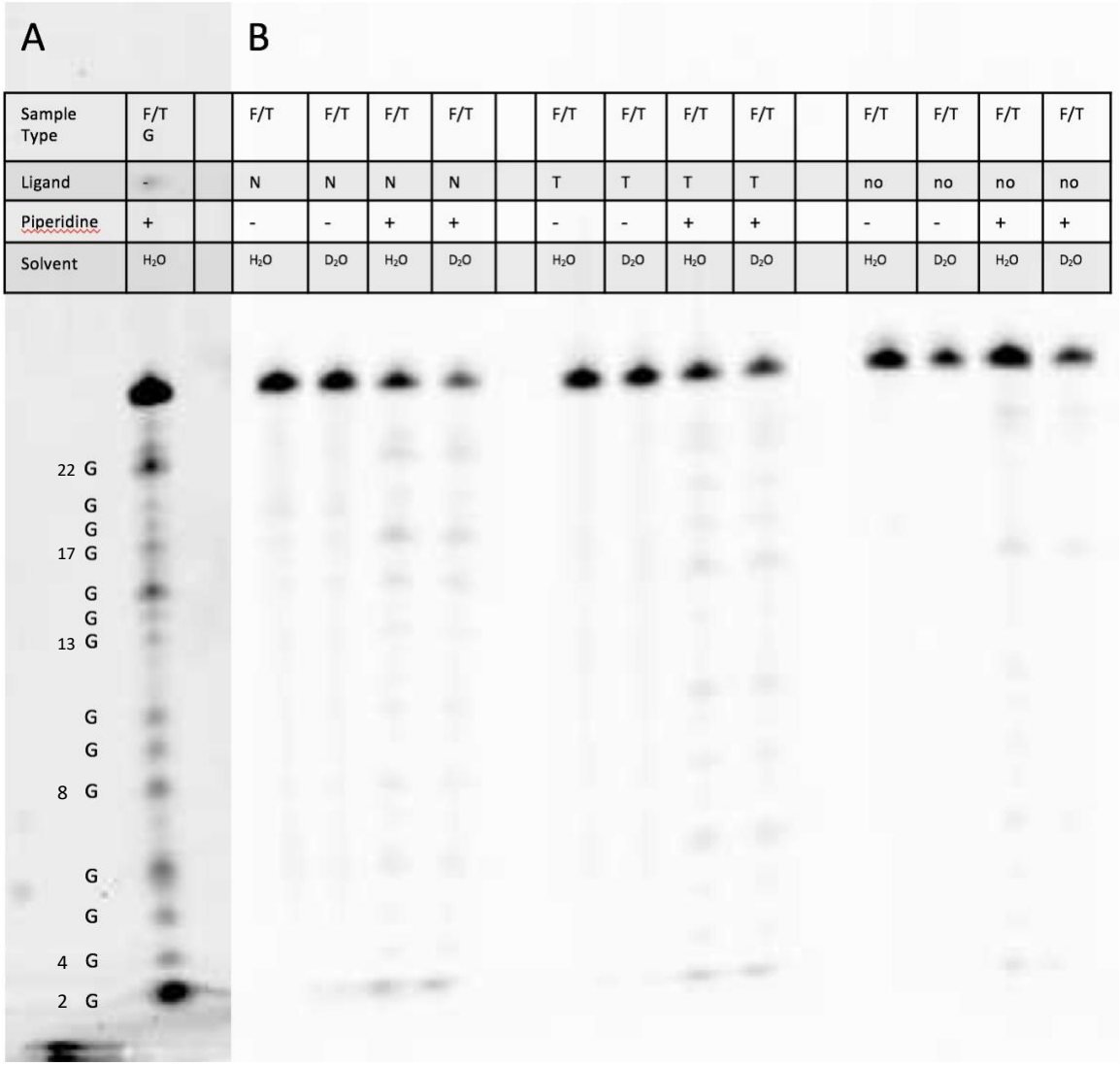

Figure 3.10. 90min Photocleavage of cMYC22m with NMM and TMPyP4, with G-ladder. Sample types include F-cMYC22m-T (F/T) and F-cMYC22-T G-ladder (F/T G). Ligands are TMPyP4 (T), NMM (N), or no ligand (no). (A) Adjusted image of G-ladder lane (B) Image taken at low intensity setting.

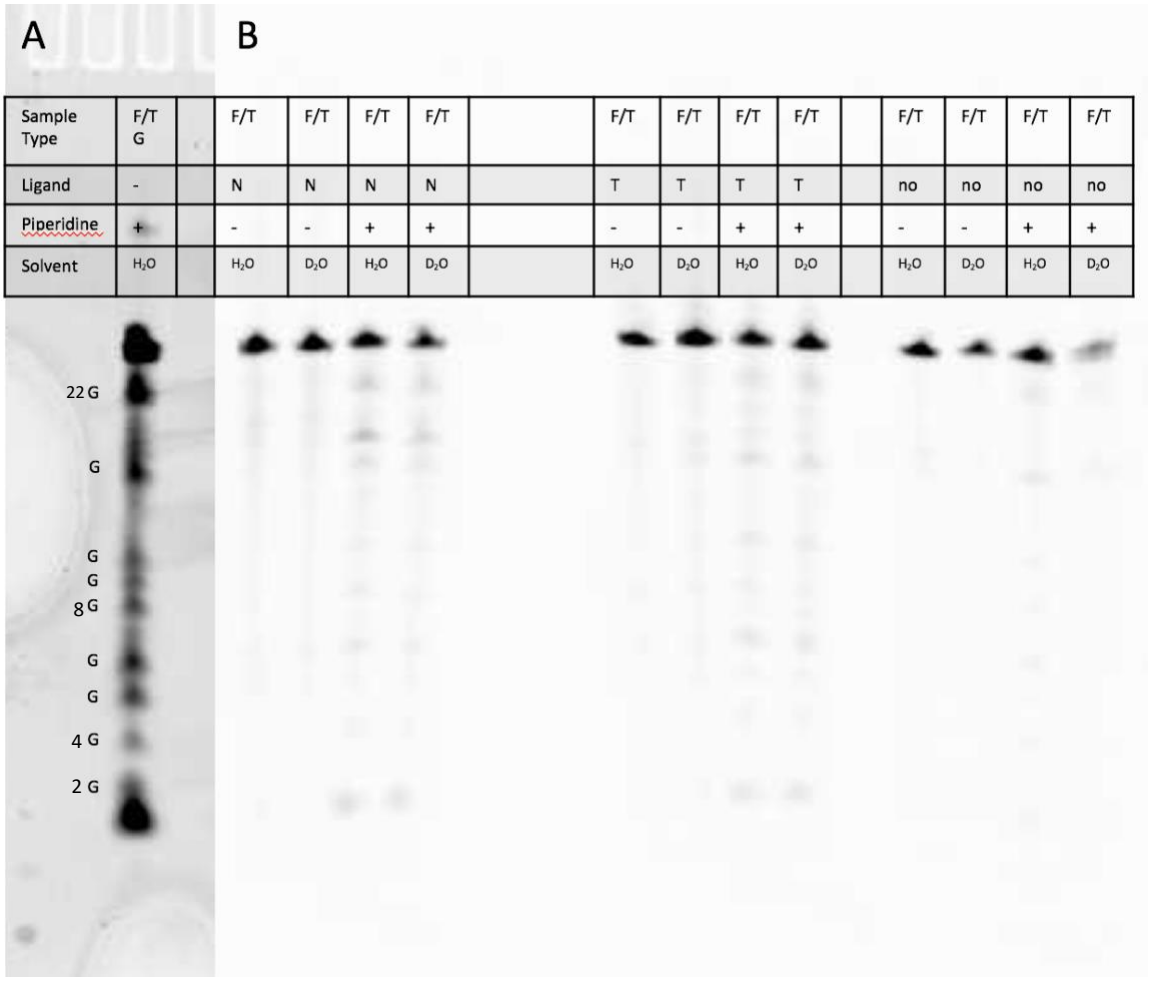

Figure 3.11. 60min Photocleavage of cMYC22m with NMM and TMPyP4, with G-ladder. Sample types include F-cMYC22m-T (F/T) and F-cMYC22-T G-ladder (F/T G). Ligands are TMPyP4 (T), NMM (N), or no ligand (no). (A) Adjusted image of G-ladder lane (B) Image taken at low intensity setting. All guanines are not shown on G-ladder due to poor resolution

3.4 Discussion

In order to deduce the involvement of fluorophores in the photocleavage reactions attachment of an alternative fluorophore to FAM, BoDIPY, was examined. BoDIPY strained alkynes were attached to an azide-labeled cMYC22m oligonucleotide using click chemistry. These click reactions were expected to result in a single band on polyacrylamide gels, indicating attachment of the fluorophore to the full-length oligonucleotide. However, multiple bands were visible for these reactions. Samples were run on denaturing polyacrylamide gels; therefore, these bands likely do not represent folded species, diffuse faster-moving bands are likely a decomposing product of the BoDIPY. The click reaction products were also subjected to DMS and piperidine heat treatment to create G-ladders. These reactions, however, proved unsuccessful and therefore BoDIPY-cMYC22m was not used for the photocleavage studies. As a result, no comparison may be made against F-cMYC22m-T photocleavage reactions to understand the involvement of fluorophores in the reaction. Future experiments may use fluorophores such as Alexa in click reactions for these experiments.

Photocleavage reactions were performed on F-cMYC22m-T with TMPyP4 was a control photocleavage agent for comparison with photocleavage by NMM. The photocleavage pattern of F-cMYC22m-T by both of these photocleavage agents is very similar. However, NMM shows a more cleavage at guanines 17 and 22 compared to TMPyP4. Upon comparison of the NMM photocleavage pattern with a G-ladder it is clear that NMM has a preference for guanines composing the external G-tetrads (Figure 3.12). These cleavage agents also showed poorly resolved and less intense bands for samples which were not heat treated with piperidine. This information, along with the

preference for guanine, indicates that these cleavage agents likely do not directly cleave the sugar-phosphate backbone.

In attempt to understand the photocleavage mechanism, photocleavage reactions with NMM and TMPyP4 were run in both H_2O and D_2O. The use of D2O as a solvent has been shown to increase the lifetime of singlet oxygen species by approximately ten-fold. No difference in cleavage efficiency between these two solvent types was visible. This suggests that singlet oxygen is not an active part of the photocleavage. However, this does not exclude the possibility of singlet oxygen production by the porphyrin compounds during photolysis.[53] While not conclusive, these data suggest that NMM acts on G-quadruplex structures through a mechanism other than singlet oxygen. NMM continues to show potential as a molecular probe for G-quadruplexes, however, further investigation is required to understand its specificity and mechanism of action.

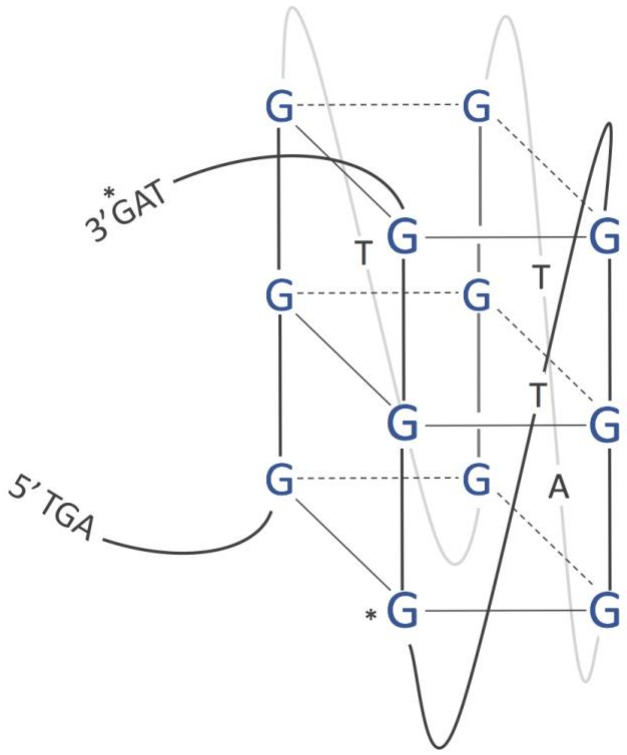

Figure 3.12. cMYC22m Parallel Stranded G-quadruplex. Parallel stranded cMYC22m with asterisks (*) showing external guanines 17 and 22. NMM shows higher cleavage rates at these guanines.

IV. CONCLUSION

This project focused on the characterization and photocleavage of G-quadruplexes from the *c-MYC* promoter region. The first goal of the project was to characterize a G-quadruplex from the G-rich hotspot region of *c-MYC*. This sequence, cMYCHS, in the presence of its complementary sequence is suspected to form an H-DNA structure in addition to the G-quadruplex. Therefore, it was believed that these two secondary structures compete for formation and exist in some equilibrium based on thermodynamic stability. Spectroscopic methods were used to investigate the cMYCHS G-quadruplex. UV-melts and FRET analysis yielded results which were not consistent with intramolecular G-quadruplex formation, but these results alone were inconclusive. However, when UV and fluorescent data was analyzed in conjunction with CD data it became evident that the G-quadruplex forms through intermolecular interactions. Intermolecular formation of the G-quadruplex is assumed to be biologically irrelevant in this context. Therefore, it is assumed that this G-quadruplex is not in competition for formation with H-DNA from this region. SPR studies in a collaborating lab, not shown in this document, corroborate these findings. Further investigation of the H-DNA from the *c-MYC* hotspot region are ongoing.

The second goal of this project was to investigate the G-quadruplex-specific photocleavage mechanism of a porphyrin, NMM. This compound had previously been shown to preferentially cleave the parallel topology of the F-cMYC22m-T G-quadruplex. Other porphyrins, such as TMPyP4, have shown the ability to cleave G-quadruplex DNA, but without specificity toward any type of fold, or for G-quadruplex over duplex DNA. The specificity seen in NMM photocleavage made it of interest in the search for G-

quadruplex molecular probes. However, the mechanism of NMM photocleavage had not yet been investigated. Therefore, this project focused on understanding how the photocleavage reaction occurs, as well as investigating if the FAM and TAMRA fluorophores attached to the oligonucleotide play any role in the cleavage process. Addition of a BoDIPY tag to cMYC22m was attempted in order to compare photocleavage patterns with F-cMYC22m-T. However, this addition of BoDIPY to cMYC22m was successful, however, it degraded and could not be used for comparison. A Maxam-Gilbert G-ladder and the control photocleavage agent, TMPyP4, were used for comparison against NMM photocleavage reactions with F-cMYC22m-T. PAGE analysis revealed similar cleavage patterns with preference for guanines between TMPyP4 and NMM. NMM, however, did show more specificity for guanines not involved in the middle G-tetrad layer. The involvement of singlet oxygen in both NMM and TMPyP4 photocleavage is assumed to be unlikely due to samples in D_2O showing no increased efficiency in photocleavage.

These results support the use of NMM as a molecular probe. However, the mechanism of this reaction is still a puzzle, now with only a few pieces in place. Therefore, future studies will be conducted in order to better understand this mechanism. Future experiments will include the use of a cMYC22m-F system in order to investigate any involvement of the TAMRA fluorophore. Also, photocleavage reactions will be performed in the presence of a singlet oxygen sensor in order to detect if singlet oxygen is being produced, despite it not being involved in the DNA cleavage.

REFERENCES

(1) Du, Y. and Zhou, X. (2013). Targeting Non-B-Form DNA in Living Cells. *Chem. Rec.* **13:** 371-384

(2) Bacolla, A., Wang, G. and Vasquez, K.M. (2015). New Perspectives on DNA and RNA Triplexes as Effectors of Biological Activity. *PLoS Genet.* **11:** e1005696

(3) Ou, T.M., Lu, Y.J., Tan, J.H., Huang, Z.S., Wong, K.Y. and Gu, L.Q. (2008) G-quadruplexes: Targets in Anticancer Drug Design. *Chem. Med. Chem.* **3:**690-713.

(4) Rocca, R., Moraca, F., Costa G., Nadai, M., Scalabrin, M., Talarico, C., Distinto, S., Maccioni, E., Ortuso, F., Artese, A., Alcaro, S. and Richter, S.N. (2016) Identification of G-quadruplex DNA/RNA Binders: Structure-based Virtual Screening and Biophysical Characterization. *Biochim. Biophys. Acta.* **12:** 1329-1340

(5) Đapić, V., Abdomerović, V., Marrington, R., Peberdy, J., Rodger, A., Trent, J.O., and Batesb, P.J. (2003) Biophysical and Biological Properties of Quadruplex Oligodeoxyribonucleotides. *Nucleic Acids Res* **31:**2097-107

(6) Del Mundo, I.M.A., Zewail-Foote, M., Kerwin, S.M. and Vasquez, K.M. (2017) Alternative DNA Structure Formation in the Mutagenic Human *c-MYC* Promoter. *Nucleic Acids Res* **45:** 4929-4943.

(7) Keniry, M. A. (2001). Quadruplex Structures in Nucleic Acids. *Biopolymers* **56:** 123-146

(8) Hardin, C.C., Perry, A.G. and White, K. (2000) Thermodynamic and Kinetic Characterization of the Dissociation and Assembly of Quadruplex Nucleic Acids. *Biopolymers* **56:** 147–194.

(9) Hazel, P., Huppert, J., Balasubramanian, S. and Neidle, S. (2004). Loop-Length-Dependent Folding of G-Quadruplexes. *J. Am. Chem. Soc.* **126:** 16405-16415

(10) Yang, D. and Okamoto, K. (2010) Structural Insights into G-quadruplexex: Towards New Anticancer Drugs. *Future Med Chem* **2:** 619-46

(11) Chaires, J. B., and Graves, D., editor. Quadruplex Nucleic Acids. Louisville, KY: Springer; 2013

(12) Mergny, J., Phan, A. and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett.* **11:** 74-8.

(13) Schoonover, M. and Kerwin, S. M. (2012). G-quadruplex DNA Cleavage Preference and Identification of a Perylene Diamine G-quadruplex Photocleavage Agent Using Rapid Fluorescent Assay. *Bioorganic & Med. Chem.* **20:** 6904-6918

(14) Sun, D. and Hurley, L. H. (2009). The Importance of Negative Superhelicity in Introducing the Formation of G-Quadruplex and i-Motif Structures in the c-Myc Promoter: Implications for Drug Targeting and Control of Gene Expression. *J. Med. Chem.* **52:** 2863-2874.

(15) Dalla-Favera, R., Bregni, M., Erikson, J., Patterson, D., Gallo, R. and Croce, C.M. (1982) Human c-MYC Onc Gene is Located on the Region of Chromosome 8 that is Translocated in Burkitt Lymphoma Cells. *Proc. Natl. Acad. Sci.* **79:**7824-7.

(16) Ma, J. and Wang, M. D. (2016). DNA Supercoiling During Transcription. *Biophys. Rev.* **8:** 75-87

(17) Huppert J L, and Balasubramanian S. G-Quadruplexes in Promoters Throughout the Human Genome. (2007) *Nucleic Acids Research* **35:** 406–413.

(18) Maizels N, and Gray LT. The G4 Genome. (2013) *PLoS Genetics* **9:** e1003468

(19) Henderson, A., Wu, Y., Huang, Y.C., Chavez, E.A., Platt, J., Johnson, F.B., Brosh, R.M., Sen, D. and Lansdorp, P.M. (2014). Detection of G-quadruplex DNA in Mammalian Cells. *Nucleic Acids Res.* **42:** 860-869

(20) Azad, R. N., Zafiropoulos, D., Ober, D., Jiang, Y., Chiu, T., Sagendorf, J. M., Rohs, R. and Tullius, T. D. (2018). Experimental Maps of DNA Structures at Nucleotide Resolution Distinguish Intrinsic from Protein-induced DNA Deformations. *Nucleic Acids Res.* **46:** 2636-2647

(21) Balasubramanian, S., Hurley, L.H. and Neidle, S. (2011). Targeting G-quadruplexes in Gene Promoters: A Novel Anticancer Strategy. *Nat. Rev. Drug Discov.* **10:** 261-275

(22) Qin, Y. and Hurley, L.H. (2008) Structures, Folding Patterns, and Functions of Intramolecular DNA G-quadruplexes Found in Eukaryotic Promoter Regions. *Biochimie* **90:** 1149-1171

(23) Ambrus, A., Chen, D., Dai, j., Jones, R.A. and Yang, D. (2005). Solution Structure of the Biologically Relevant G-Quadruplex Element in the Human c-MYC Promoter. Implications for G-Quadruplex Stabilization. *Biochemistry* **44:** 2048-2058

(24) Wang, G., Carbajal, S., Vijg, J., DiGiovanni, J. and Vasquez, K.M. (2008) DNA Structure-induced Genomic Instability in vivo. *J. Natl. Cancer Inst.* **100:** 1815-7

(25) Nambiar, M. and Raghavan, S. C. (2011). How Does DNA Break During Chromosomal Translocations? *Nucleic Acids Res.* **39:** 5813-5825

(26) Kang, Z.J., Liu, Y.F., Xu, L.Z., Long, Z.J., Huang, D., Yang, Y., Liu, B., Feng, J.X., Pan, Y.J., Yan, J.S. and Liu, Q. (2016). The Philadelphia Chromosome in Leukemogenesis. *Chin. J. Cancer* **35:** DOI 10.1186/s40880-016-0108-0

(27) Rhodes, D. and Lipps, H.J. (2015). G-quadruplexes and Their Regulatory Roles in Biology. *Nucleic Acids Res.* **43:** 8627-8637

(28) Hurley, L.H. (2001) DNA and its Associated Processes as Targets for Cancer Therapy. *Nature* **2:** 188-200.

(29) Nguyen, L., Papenhausen, P. and Shao, H. (2017) The Role of c-MYC in B-Cell Lymphomas: Diagnostic and Molecular Aspects. *Genes* **8:** 116.

(30) Siddiqui-Jain, A., Grand, C. L., Bearss, D. J. and 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 Transcription. *Proc. Natl. Acad. Sci. U.S.A.* **99:** 11593-11598.

(31) Phan, A.T. and Patel, D.J. (2003). Two-repeat Human Telomeric d(TAGGGTTAGGGT) Sequence Forms Interconverting Parallel and Antiparallel G-quadruplexes in Solution: Distinct Topologies, Thermodynamic Properties, and Folding/Unfolding Kinetics. *J. Am. Chem. Soc.* **125:** 15021-7.

(32) Lane, A.N., Chaires, J.B., Gray, R.D. and Trent, J.O. (2008) Stability and Kinetics of G-quadruplex Structures. *Nucleic Acids Res.* **36:**5482-5515.

(33) Armitage, B. (1998). Photocleavage of Nucleic Acids. *Chemical Reviews* **98:** 1171-1200

(34) Maxam, A. and Gilbert, W. (1976). A New Method for Sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.* **74:** 560-564

(35) Baptista, M.S., Cadet, J., Mascio, P.D., Ghogare, A.A., Greer, A., Hamblin, M.R., Lorente, C., Nunez, S.C., Ribeiro, M.S., Thomas, A.H., Vignoni, M. and Yoshimura, T.M. (2017). Type I and Type II Photosensitized Oxidation Reactions: Guidelines and Mechanistic Pathways. *Photocleavage and Photobiology* **93:** 912-919

(36) Ros, T.D., Spalluto, G., Boutorine, A.S., Bensasson, R.V. and Prato, M. (2001). DNA-Photocleavage Agents. *Current Pharmaceutical Design* **7:** 1781-1821

(37) McBrayer, D. N. (2015). Synthesis and Evaluation of Rationally Designed Small Molecular Libraries for G-quadruplex Selective DNA Photocleavage (Dissertation). In press, The University of Texas at Austin.

(38) Postel, E.H., Mango, S.E. and Flint, S.J. (1989). A Nuclease-Hypersensitive Element of the Human *c-MYC* Promoter Interacts with a Transcription Initiation Factor. *Mol. Cell Biol.* **9:** 5123-5133

(39) Le, H., Miller, C., Buscaglia, R., Dean, W., Holt, P. A., Chaires, J. B. and Trent, J. (2012). Not All G-quadruplexes are Created Equally: An Investigation of the Structural Polymorphism of the c-MYC G-quadruplex-Forming Sequence and its Interactions with the Porphyrin TMPYP4. *Organic and Biomolecular Chemistry* **10:** 9393-9404

(40) Olsen, C. M., Lee, H. and Marky L. A. (2009). Unfolding Thermodynamics of Intramolecular G-quadruplexes: Base Sequence Contributions of the Loops. *J. Phys. Chem.* **113:** 2587-2595

(41) Sun, Y., Rombola, C., Jyothikumar, V. and Periasamy, A. (2013). Förster Resonance Energy Transfer Microscopy and Spectroscopy for Localizing Protein-Protein Interactions in Living Cells. *Cytometry A.* **83:** 780-793

(42) FAM (6-fluorescein amidite (6-FAM)) [internet]; c2016 [cited 2016 10/31]. Available from: http://www.genelink.com/newsite/products/mod_detail.asp?modid=33

(43) TAMRA-3' (Carboxytetramethylrhodamine) [internet]; c2016 [cited 2016 10/31]. Available from: http://www.genelink.com/newsite/products/mod_details.asp?modid=37

(44) Randazzo, A., Spada, G.P. and Silva, M.W. (2012). Circular Dichroism of Quadruplex Structures. *Topics in Current Chemistry* **330:** 67-86

(45) Vorlickova, M., Keinovska, I., Sagi, J., Renciuk, D., Bendnarova, K., Motlova, J. and Kypr, J. (2012). Circular Dichroism and Guanine Quadruplexes. *Methods* **57:** 64-75

(46) Bishop, G.R. and Chaires, J.B. (2003). Characterization of DNA Structures by Circular Dichroism. *Current Protocols in Nucleic Acid Chemistry* **11:** 7.11.1-7.11.8

(47) Brown, N.M., Rachwal, P.A., Brown, T. and Fox, K.R. (2005). Exceptionally Slow Kinetics of the Intramolecular Quadruplex Formed by the Oxytricha Telomeric Reapeat. *Org. Biomol. Chem.* **3:** 4153-4157

(48) Hatzakis, E., Okamoto, K. and Yang, D. (2010). Thermodynamic Stability and Folding Kinetics of the Major G-Quadruplex and its Loop-Isomers Formed in the Nuclease Hypersensitive Element in the Human c-*MYC* Promoter-Effect of Loops and Flanking Segments on the Stability of Parallel-Stranded Intramolecular G-Quadruplexes. *Biochemistry* **49:** 9152-9160

(49) Beniaminov, A., Novikov, R.A., Mamaeva, O.K., Mitkevich, V.A., Smirnov, I.P., Livshits, M.A., Shchyolkina, A.K. and Kaluzhny, D.N. (2016). Light-Induced Oxidation of the Telomeric G4 DNA in Complex with Zn(II) Tetracarboxymethyl Porphyrin. *Nucleic Acids Research* **44:** 10031-10041

(50) Dai, J., Carver, M. and Yang, D. (2008). Polymorphism of Human Telomeric Quadruplex Structures. *Biochimie.* **90:** 1172-1183

(51) Jewett, J.C. and Bertozzi, C.R. (2010). Cu-Free Cycloaddition Reactions in Chemical Biology. *Chem. Soc. Rev.* **39:** 1272-1279

(52) Agnez-Lima, L.F., Melo, J.T.A., Silva, A.E., Oliveira, A.H.S., Timoteo, A.R.S., Lima-Bessa, K.M., Martinez, G.R., Medeiros, M.H.G., Mascio, P.D., Galhardo, R.S. and Menck, C.F.M. (2012). DNA Damage by Singlet Oxygen and Cellular Progective Mechanisms. *Mutation Research* **751:** 15-28

(53) Kopkalli, Yasemin & Biegel, Craig & Khemraj, Ryan & Davenport, Lesley. (2015). Binding Studies of Small Molecules to Telomeric Quadruplex DNA for Targeted Singlet Oxygen Production. *Biophysical Journal* **108**: 395a. 10.1016/j.bpj.2014.11.2166.