Nucleic acid clamp-mediated transcriptional and translational modulation of oncogenes via the stabilization of G-quadruplexes

Taisen Hao
University of Mississippi

Follow this and additional works at: https://egrove.olemiss.edu/etd

Part of the Molecular Biology Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Graduate School at eGrove. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of eGrove. For more information, please contact egrove@olemiss.edu.
NUCLEIC ACID CLAMP-MEDIATED TRANSCRIPTIONAL AND TRANSLATIONAL MODULATION OF ONCOGENES VIA THE STABILIZATION OF G-QUADRUPLEXES

A Dissertation

Presented in fulfillment of requirements

for the degree of Doctorate of Philosophy

in the Department of BioMolecular Sciences

Division of Pharmacology

The University of Mississippi

by

TAISEN HAO

August 2017
G-quadruplexes (G4s) are non-canonical secondary nucleic acid structures that exist in various biologically important regions within a cell, such as the promoters of DNA, 5’-untranslated regions (UTR) of mRNA, and telomeres. In DNA, potential G4-forming sequences tend to cluster within 500 bp from transcriptional start sites (TSS) across all chromosomes in an evolutionarily conserved manner in addition to at telomeric regions. Extensive efforts have been dedicated to targeting G4s in oncogenic promoters, telomeres, and 5’-UTRs with traditional small molecules. However, most small molecules recognize multiple G4s, raising concerns about selectivity and specificity. Therefore, we developed a novel nucleic acid clamp (NA-clamp) based strategy, which takes advantage of both the highly specific complementary base pairing capability of nucleic acids and the three-dimensional variation in size of G4 structures, as compared to linear DNA, in order to stabilize individual structures with high specificity. In this work, we explored the possibility of applying this NA-clamp approach at both the transcriptional and the translational levels. Specifically, we used this NA-clamp approach on the MYC promoter and the NRAS 5’-UTR G4s. Silencing MYC and NRAS expression is an effective approach to induce apoptosis of lymphoma and breast cancer, and melanoma, respectively. We examined the binding specificity and G4 stabilizing capability of the designed clamps to their corresponding targets using electromobility shift assay, electronic circular dichroism and DMS footprinting. Cellular studies including cell viability assays, luciferase assays, fluorescent microscopy, ChIP-PCR, qPCR and western blot were used to evaluate the pharmacological effects of the clamps. This research provided insights
on the applicability of NA-clamps as specific G4 stabilizers that modulate gene expression in cancer cells, providing the foundation for the future development of NA-clamps into effective therapeutics, or as diagnostic companions, against many cancer types associated with MYC or NRAS.
DEDICATION

This dissertation is dedicated to my beloved family, friends, the department of BioMolecular Sciences and all the people who helped me selflessly.
LIST OF ABBREVIATIONS AND SYMBOLS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’-UTR</td>
<td>5’-Untranslated region</td>
</tr>
<tr>
<td>6-FAM</td>
<td>6-Carboxyfluorescein</td>
</tr>
<tr>
<td>ACN</td>
<td>Acetonitrile</td>
</tr>
<tr>
<td>AKTIP</td>
<td>AKT interacting protein</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium persulfate</td>
</tr>
<tr>
<td>Bcl-2</td>
<td>B-cell lymphoma 2</td>
</tr>
<tr>
<td>CDK</td>
<td>Cyclin dependent kinases</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin immunoprecipitation</td>
</tr>
<tr>
<td>CNBP</td>
<td>Cellular nucleic acid-binding protein</td>
</tr>
<tr>
<td>CTD</td>
<td>C-terminal domain</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s Modified Eagle Medium DMS Dimethyl sulfate</td>
</tr>
<tr>
<td>DMS</td>
<td>Dimethyl sulfate</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double-stranded DNA</td>
</tr>
<tr>
<td>E2</td>
<td>17β-Estradiol</td>
</tr>
<tr>
<td>ECD</td>
<td>Electronic circular dichroism</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>EMSA</td>
<td>Electromobility shift assay</td>
</tr>
<tr>
<td>EMT</td>
<td>Epithelial-mesenchymal transition</td>
</tr>
<tr>
<td>ER</td>
<td>Estrogen Receptor</td>
</tr>
<tr>
<td>EV</td>
<td>Empty vector plasmid</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal Bovine Serum</td>
</tr>
<tr>
<td>FBP</td>
<td>FUSE binding protein</td>
</tr>
<tr>
<td>FRI</td>
<td>FBP interacting repressor</td>
</tr>
<tr>
<td>FUSE</td>
<td>FarUpStream Element</td>
</tr>
<tr>
<td>G4</td>
<td>G-quadruplex</td>
</tr>
<tr>
<td>GAP</td>
<td>GTPase-activating protein</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GDP</td>
<td>Guanosine diphosphate</td>
</tr>
<tr>
<td>GEF</td>
<td>Guanine nucleotide exchange factor</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanosine-5’-triphosphate</td>
</tr>
<tr>
<td>HLH-/Leu-zipper domain</td>
<td>Helix-loop-helix-leucine zipper domain</td>
</tr>
<tr>
<td>hnRNPk</td>
<td>Heterogeneous nuclear ribonucleoprotein K</td>
</tr>
<tr>
<td>HRAS</td>
<td>Harvey rat sarcoma viral oncogene homolog</td>
</tr>
<tr>
<td>hTERT</td>
<td>Telomerase reverse transcriptase</td>
</tr>
<tr>
<td>IC&lt;sub&gt;50&lt;/sub&gt;</td>
<td>Half maximal inhibitory concentration</td>
</tr>
<tr>
<td>IGF-II</td>
<td>Insulin-like growth factor II</td>
</tr>
<tr>
<td>kRAS</td>
<td>Kirsten rat sarcoma viral oncogene</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MCF</td>
<td>Michigan cancer foundation</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MT</td>
<td>Mutant</td>
</tr>
<tr>
<td>MTS</td>
<td>3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium</td>
</tr>
<tr>
<td>NHE</td>
<td>Nuclease hypersensitivity element</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance spectroscopy</td>
</tr>
<tr>
<td>NRAS</td>
<td>Neuroblastoma rat sarcoma viral oncogene homolog</td>
</tr>
<tr>
<td>NSC</td>
<td>New synthetic compound</td>
</tr>
<tr>
<td>p53</td>
<td>Tumor protein 53</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene glycol</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PMS</td>
<td>Phenazine methosulfate</td>
</tr>
<tr>
<td>PNA</td>
<td>Peptide nucleic acid</td>
</tr>
<tr>
<td>pRB</td>
<td>Retinoblastoma protein</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative real-time polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNA Pol II</td>
<td>RNA polymerase II</td>
</tr>
<tr>
<td>RPMI</td>
<td>Roswell Park Memorial Institute</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering RNA</td>
</tr>
<tr>
<td>Sp1</td>
<td>Specificity protein 1</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single-stranded DNA</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris/Borate/EDTA buffer</td>
</tr>
<tr>
<td>TEMED</td>
<td>Tetramethylethylenediamine</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>TERRA</td>
<td>Telomeric repeat-containing RNA</td>
</tr>
<tr>
<td>TM</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>TSS</td>
<td>Transcriptional start site</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular endothelial growth factor</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my advisor, Dr. Tracy Brooks, for her continuous support of my Ph.D. studies and my research work for the past several years. Her patient guidance helped me immensely in all the time of learning, research, and writing of this dissertation. I can never imagine having a better mentor than Dr. Brooks for my Ph.D. studies.

Next, I would like to thank the rest of my committee members: Dr. Kristie Willett, Dr. Shabana Khan, Dr. David Colby, and Dr. Joshua Sharp, for their time, effort and insightful comments and encouragement, but also the hard questions that motivated me to broaden my research from various perspectives.

My sincere thanks to my colleagues: Harshul Batra, Rhianna Morgan and Disha Prabhu, for their help, companionship, and all the team work on research over the past years. I will never forget all the great times spent with them.

I greatly appreciate the generous and kind help from the faculty, staff and graduate students in the Department of BioMolecular Sciences. I am fortunate to join this program and onboard my research career with all of these wonderful people that left me great memories that I will cherish in my heart for life.

Last but not the least, I would like to thank to my family, friends for their unconditional love and support during all the up and downs in my life. I could have never imagined accomplish for the work without their precious encouragement and support.

This work was supported by Dr. Tracy Brooks’ startup funds from the University of
Mississippi, National Institute of Health (NIH) grant [1P50CA13080501A1 sub-awarded to T.A.B], and Taisen Hao’s University of Mississippi Graduate Student Council (GSC) Research Grant in 2015.
# TABLE OF CONTENTS

ABSTRACT ............................................................................................................................................... II
DEDICATION ........................................................................................................................................ IV
LIST OF ABBREVIATIONS AND SYMBOLS .................................................................................. V
ACKNOWLEDGMENTS .................................................................................................................... IX
TABLE OF CONTENTS ..................................................................................................................... XI
LIST OF FIGURES .......................................................................................................................... XIII
LIST OF TABLES ............................................................................................................................. XV

## CHAPTER I. INTRODUCTION .............................................................................................................. 1

## CHAPTER II. MATERIALS AND METHODS ...................................................................................... 12

2.1. REAGENTS .................................................................................................................................... 12
2.2. DNA DISSOLUTION, CONCENTRATION DETERMINATION AND PURIFICATION ......................................................... 12
2.3 ELECTROPHORETIC MOBILITY SHIFT ASSAY (EMSA) ......................................................................................... 13
2.4. ELECTRONIC CIRCULAR DICHRHOISM (ECD) .............................................................................................. 14
2.5. DIMETHYL SULFATE (DMS) FOOTPRINTING ............................................................................................ 14
2.6. PULL DOWN ASSAY ......................................................................................................................... 15
2.7. CELL CULTURE ............................................................................................................................... 16
2.8. TRANSFECTION AND LUCIFERASE ASSAY ........................................................................................... 16
2.9. CELLULAR VIABILITY ASSAY (MTS ASSAY) ......................................................................................... 17
2.10. CELLULAR UPTAKE ASSAY ............................................................................................................ 17
2.11. CHIP-PCR ASSAY ......................................................................................................................... 18
2.12. QUANTITATIVE REAL-TIME PCR (qRT-PCR) .................................................................................... 19
2.13. G4 INDUCTION ASSAY ................................................................................................................ 20
2.14. WESTERN BLOTTING .................................................................................................................. 21

## CHAPTER III. TARGETING MYC PROMOTER G4 USING NA-CLAMPS .............................................. 25

3.1. INTRODUCTION ........................................................................................................................... 25
3.2. THE DEVELOPMENT OF CLAMP A AS A STABILIZER FOR THE PHYSIOLOGICALLY RELEVANT MYC G4 ............................. 41
   3.2.1. The design, optimization and concept demonstration of MYC clamps ......................................................... 41
   3.2.2. NA-clamp binding and its specific G4 recognition ...................................................................................... 46
   3.2.3. Clamp A G4 affinity and MYC G4 stabilization ....................................................................................... 52
   3.2.4. Clamp A-mediated inhibition of MYC promoter activity and differential cytotoxicity .............................. 56
3.3. DEVELOPMENT OF CLAMP A 5T AS AN ALTERNATIVE FOR CLAMP A ......................................................... 60
3.4. CELLULAR UPTAKE AND DELIVERY OF CLAMP A 5T .............................................................................. 67
3.5. INTRACELLULAR ACTIVITY OF CLAMP A 5T .................................................................................... 80
3.6. CLAMP A 5T AS A DIAGNOSTIC COMPANION .................................................................................. 84
LIST OF FIGURES

Figure 1-1. The building blocks for G4s. .................................................................3
Figure 3-1. MYC promoter structure and the NHE III1 region capable of forming G4s. .................................................................37
Figure 3-1. NA-clamp binding to the MYC wild type Pu46 sequence. ...............43
Figure 3-3. Computational modeling of the physiologically relevant MYC promoter G4 and clamp-mediated recognition. .................................................................44
Figure 3-4. Designed NA-clamps for multiple G4 isoforms formed in the MYC promoter. .................................................................45
Figure 3-5. Specific binding and recognition of clamps to MYC G4s. ................50
Figure 3-6. Clamp-DNA complexes separated by streptavidin-mediated pull down. ....51
Figure 3-7. Preferential G4 affinity of clamp A for G41-4. ..................................54
Figure 3-8. DMS footprinting of the DNA species of MYC WT DNA post incubation with clamp A. .................................................................55
Figure 3-9. Clamp A significantly decreases MYC promoter activity. ...........58
Figure 3-10. Clamp A’s cytotoxic profile on cancerous and non-cancerous cells and its ability on MYC expression. .................................................................59
Figure 3-3-1. DNA and G4 recognition comparison between clamp A, clamp A 5T and clamp fragments. .................................................................64
Figure 3-3-2. Specific recognition of clamp A 5T to the physiological MYC G41-4. ....65
Figure 3-3-3. Structure elucidation of the DNA-clamp complexes by DMS footprinting. .................................................................66
Figure 3-4-1. Time course cellular uptake of clamp A 5T in live MCF-7 cells. ....68
Figure 3-4-2. Time course cellular uptake of clamp A 5T in live RAJI lymphoma cells. .................................................................69
Figure 3-4-3. Time course cellular uptake of clamp A 5T in live CA46 lymphoma cells. .................................................................70
Figure 3-4-4. Block copolymers evaluated for the delivery of NA-clamps into lymphoma cells. .................................................................74
Figure 3-4-5. Delivery of clamp A 5T into RAJI lymphoma cells by block copolymers. .................................................................75
Figure 3-4-6. Delivery of clamp A 5T into CA46 lymphoma cells by block copolymers. .................................................................76
Figure 3-4-7. Cytotoxicity of clamp A 5T on HEK-293 and MCF-7 cells. ..........................78
Figure 3-4-8. Lymphoma cells cytotoxicity of clamp A 5T transfected by ethyl series copolymers. .............................................................................................................79
Figure 3-5-1. Clamp A 5T binding to the promoter region of MYC. .................................82
Figure 3-5-2. The effect of clamp A 5T on MYC expression in MCF-7 and RAJI cells. ................................................................................................................................83
Figure 3-6-1. Detecting the physiologically relevant G4 formation induced by selective MYC G4 stabilizing compound NSC338258. .........................................................87
Figure 3-6-2. G4 induction assay on RAJI and CA46 cells. .............................................88
Figure 4-1. Human NRAS 5'-Untranslated Region (5'-UTR) and the mRNA G-Quadruplex. .......................................................................................................................103
Figure 4-2. The central hypothesis of NA-clamps mediated stabilization of the NRAS 5'-UTR RNA G-quadruplex. ..................................................................................104
Figure 4-3. Binding of linked clamps 1 and 2 to the NRAS G4. .................................108
Figure 4-4. Binding of clamps to NRAS G4 under physiologically relevant condition. .................................................................................................................................109
Figure 4-5. NRAS G4 induction by annealing with clamp 1. ...............................112
Figure 4-6. Clamp 1-mediated stabilization of NRAS G4. .................................113
Figure 4-7. Cytotoxicity of clamp 1 on cancerous melanoma WM 3682 and non-cancerous HEK-293 cell lines. .................................................................116
Figure 4-8. Effect of clamps on NRAS translation. .............................................117
LIST OF TABLES

Table 1-1. G4s in the Hallmarks of Cancer. .................................................................5
Table 1-2. Summary of selected G4-interacting small molecules and their proposed biological activities. ........................................................................................................10
Table 2-1. MYC Oligonucleotide sequences (5'-3’). .....................................................23
Table 2-2. NRAS Oligonucleotide sequences (5'-3’). .....................................................24
CHAPTER I. INTRODUCTION

The history of nucleic acid research dates back to 1869, when the Swiss physiological chemist, Friedrich Miescher, first discovered “nuclein” (the name of which was later changed into nucleic acids and eventually DNA) inside the nuclei of white blood cells. This milestone discovery paved the way for genetic and molecular biology research (Dahm 2008). In 1920, Rosalind Franklin, a physical chemist, produced high-resolution photographs of DNA fibres, based on which she calculated the dimensions of the strands and proposed a possible helical structure with phosphates locate on the outside of the structure. In 1953, the modern double stranded helical DNA structure was proposed by Thomas Watson and Frances Crick (Watson and Crick 1953). This landmark work granted both scientists, together with Maurice Wilkins, a Nobel Prize for Physiology or Medicine in 1961.

With the fast advancement of technologies and biophysical techniques, other forms of secondary DNA structures, including left-handed double helical Z-form DNA, right-handed A-form DNA with more compact helical structure compared to B-form DNA, cruciform DNA, three stranded triplex DNA, stem and loop hairpins, and G-quadruplex (G4) DNA have been described. These non-B form secondary DNA structures exist in cells and play essential biological functional roles (Bochman et al. 2012). Our work focuses on G4 DNA (Figure 1-1), whose structure, brief summary the biological functions, and the therapeutic potential of this non-canonical secondary nucleic acid structure will be expounded below.

Guanine-rich DNA studies started with a big “Bang” in 1910 when a Norwegian physician
scientist named Ivar Bang elucidated the structure of Guanylic acid and discovered the phenomena where concentrated solutions of guanylic acids were able to form gels (Bang et al. 1910). In 1962, the quartet structure of the gel formed by guanylic acid was proposed by Drs. Martin Gellert, Marie N. Lipsett and David Davis, which marks the beginning of modern G-quartet research (Gellert et al. 1962). In 1988, Dr. Walter Gilbert, the Nobel prize winner of chemistry in 1980, first discovered that single stranded guanine rich sequences are able to self-assembly into parallel four stranded DNA complexes, later named G4s (Sen and Gilbert 1988). To date, both DNA and RNA G4s have been demonstrated to exist in human cells (Biffi et al. 2013; Biffi et al. 2014). Many research groups have utilized strategies to target promoter, 5’-UTR, and telomeric G4s to treat a variety of human diseases, especially cancer.

G4s are formed in guanine rich DNA and RNA sequences with four runs of guanines that generally contain at least three continuous guanines in a row. Instead of guanine: cytosine base-pairing in these sequences, four guanines from each continuous segment bond with each other through Hoogsteen base pairing to form a tetrad, and the tetrads stack on top of each other with nucleotides among guanine rows forming loop structures connecting the tetrads (Figure 1-1). G4 structures exhibit high diversity topologically based on multiple factors, such as the number of strands involved (either intra or intermolecular structures), relative orientation of the loops (parallel vs antiparallel), number of tetrads comprising the structure, and the nucleotide sequences. In addition, the tetrads are stabilized when a monovalent cation, such as potassium, interacts with the tetrads, due to the electronegative nature of the negatively charged carbonyl oxygen atoms located in the center of the tetrads (Guschlbauer et al. 1990; Williamson 1994). Based on the folding pattern and loop length of G4 structures, they can be categorized into four
Figure 1-1. The building blocks for G4s. In guanine-rich sequences, four guanines are able to Hoogsteen base pair with each to form a tetrad. Three or more tetrads then stack on top of each through π-π interaction to form a G4 with interconnecting nucleotides forming loops surrounding the structure. The structure could be stabilized by a monovalent cation located in the center of the tetrad, interacting with a negatively charged carbonyl oxygen pocket. The structure shown in this figure is a parallel intramolecular G-quadruplex, with the loops of 1, 2, and 1 nucleotides interconnecting the tetrads, formed in the MYC promoter (Green=A, red=G, blue=T). Underlined guanines segments are the participating guanines for the physiologically relevant MYC G4 structure (Adapted from Brooks & Hurley 2009).
classes (I, II, III, and IV) (Brooks et al. 2010).

Human telomeres are the chromosome caps protecting the ends of the DNA. They are usually 5-8 kb long with a 150-250 nucleotide 3’ single-stranded overhang. During DNA replication, around 50-200 bases in the end of the 3’ sequence cannot be copied due to the participation of Okazaki fragments. Thus, telomeric non-coding sequences get shorter with each cycle of proliferation (Moyzis et al. 1988; Stewart and Weinberg 2006). After the telomeres reach a critical short length, the apoptosis signaling pathway is activated, resulting in cellular senescence or death (Herbig et al. 2004). Cancer cells manage to avoid this fate by re-activating a reverse transcriptase called telomerase, which functions to re-synthesize the lost telomeres to elongate chromosomes, enabling cancer cell immortality. Even though telomerase is only active during embryonic development, little telomerase expression can be detected in cells after birth; however almost 90% of cancer cells re-express telomerase and the other 10% use alternative pathways to elongate telomeres (Kim et al. 1994). It is this unique trait of cancer cells, and the lack of telomerase activity in normal cells, that makes disrupting telomere elongation an effective approach to induce cancer cell apoptosis (d'Adda di Fagagna et al. 2004).

Telomeric G4 structures were the first to be described and have been pursued as therapeutic targets (S. Neidle 2010; Sekaran et al. 2014; Williamson et al. 1989). The single stranded overhang of the telomeres is comprised of repeated sequences of d(TTAGGG), which are able to form G4s. Studies on G4 structures within telomeres revealed that there were multiples isoforms with varying loop lengths and folding patterns, including both intra- and inter-molecular structures (Bryan and Baumann 2011; Lin and Yang 2017). Studies have shown that the formation of G4s in telomeres can hamper the extension process by blocking telomerase access. Moreover, stabilization of telomere G4s by small molecules induce cancer cell senescence/apoptosis.
Cancer Hallmarks | Oncogenes with DNA G4 | Oncogenes with RNA G4
--- | --- | ---
Angiogenesis | VEGF, VEGFR2 | VEGF, FGF2
Invade And Metastasize | c-KIT, ADAM-15 | MT3-MMP, ADAM10, MST1R
Tumor Promoting Inflammation | FGF2 | CXCL14
Replicative Immortality | Telomere, hTERT | hTERT, TRF2, TERRA
Avoid Immune Destruction | TGF-β | EBNA1
Evade Growth Suppressors | pRb | TP53, PIM1, MLL1, MLL4
Sustain Proliferative Signaling | MYC, kRAS | NRAS, ESR1, TGFB2, IGF2
Deregulating Cellular Energetics | MYC, HIF-1α | CSBII
Resist Cell Death | Bcl-2 | Bcl-2, AKTIP
Genome Instability And Mutation | YY1 | PRX15

Table 1-1. G4s in the Hallmarks of Cancer (Morgan & Brooks, 2016; Cammas & Millevoi, 2016 (Hanahan and Weinberg 2011)
In addition to blocking telomerase access, small molecules that are able to induce and stabilize G4 formation in telomeric regions can facilitate the exposure of these 3’-end of the DNA, making them susceptible to modulation by multiple DNA damage response pathways, which in turn inhibits cancer cell proliferation. (d'Adda di Fagagna et al. 2003; Stephen Neidle and Parkinson 2003; S. Neidle 2010) A natural product, called telomestatin, isolated from Streptomyces anulatus, has been identified as a potent stabilizer of telomeric G4s, which in turn inhibits telomerase activity and triggers DNA damage responses (Shin-ya et al. 2001). Other major known telomeric G4-stabilizing compounds include BRACO-19, 12459, Quarfloxin, RHPS4, 360 A, and Pyridostatin, each of which demonstrate a similar mechanism of action as telomestain (Muller and Rodriguez 2014). It is worth noting that the majority of these compounds are non-selective, and that they also interact with non-telomeric G4s.

Promoter G4s have also been characterized, such as MYC, KRAS, VEGF, c-Kit, hTERT, pRb, RET, PDGF-A and Hif-1α (Brooks et al. 2010). Structures of many promoter G4s have been elucidated by NMR and X-ray crystallography (Ohnmacht and Neidle 2014). During transcription, DNA double strands are separated transiently due to the negative superhelicity generated by the upstream transcriptional machineries (Brooks and Hurley 2009). As a result, single stranded guanine and cytosine-rich sequences are able to adopt higher order DNA structures called G-quadruplexes and i-motifs, respectively (Verma et al. 2008). Recently, DNA G4s have been visualized in human cancer cells with an engineered antibody (BG4) specific to G4 structures, demonstrating their bona fide intracellular formation (Biffi et al. 2013; Biffi et al. 2014). Research using a novel ChIP-seq protocol with the same BG4 antibody revealed that G4 structures mark human regulatory chromatin with around 10,000 G4 formation sites, primarily in regulatory and
nucleosome depleted regions, where transcription is highly active. Within these nucleosome-depleted regions, the chromatin structures are open, which allows for better access of transcriptional machineries, including RNA Pol II and other transcriptional factors (Hansel-Hertsch et al. 2016; S. K. Wang et al. 2015a). Notably, more than 40% of human gene promoters, especially many oncogenes of the hallmarks of cancer (Table 1-1), contain potential G4-forming regions that regulate gene expression (Huppert and Balasubramanian 2007; Jayaraj et al. 2012).

Biologically, the majority of promoter G4s are silencers of transcription by hindering the binding of regulatory protein complexes. However in some cases, G4s, such as one formed in Bcl-2 promoter, can function as transcription activators (Brooks et al. 2010).

G4s also play an important role in RNA. These naturally single stranded structures more readily adopt more complexed secondary structures and are more stable thermodynamically than DNA G4s (Agarwala et al. 2015a; Halder and Hartig 2011). The first described RNA G4 was in the 3’-end of 5s rRNA in escherichia coli (Jin et al. 1992). RNA G4s have also been described in 5’ and 3’-UTRs and coding regions of mRNA, highlighting its biologically functional importance (Agarwala et al. 2015a; Jayaraj et al. 2012; Kumari et al. 2007; Millevoi et al. 2012). RNA G4s have also been visualized in human cells using BG4 antibody, demonstrating their intracellular existence as well (Biffi et al. 2014). Recent research employing a novel methodology combining whole cell DMS treatment with a reverse transcription stop assay and next generation sequencing, revealed that RNA G4s are globally unfolded by RNA binding proteins in mouse embryonic stem cells and evolutionally depleted in bacteria. However, this study did not probe the folding state of RNA G4s in human cells (Guo and Bartel 2016). Other studies have shown that RNA G4s are involved in various cellular activities relating to both DNA and RNA. At the DNA level, RNA G4s participate the transcription, recombination, and telomere homeostasis. At the RNA level,
they are involved in the processing of pre- and micro-RNAs, the splicing of introns, the transportation and translation of mRNAs (Rhodes and Lipps 2015). RNA G4s regulate the expression of many oncogenes that belong to the hallmarks of cancer (Table 1-1), including NRAS, Bcl-2, VEGF, TRF2 (Cammas and Millevoi 2017).

In addition to targeting telomeric quadruplexes as anti-cancer therapies, work has been conducted to target promoter and RNA G4s to downregulate oncogene expression at the transcriptional and translation level, respectively (Balasubramanian et al. 2011; Cammas and Millevoi 2017; Cimino-Reale et al. 2016). Stabilizing G4s in promoters or 5’-UTRs generally sequesters binding sites for activating factors and polymerases. (Balasubramanian et al. 2011; Cimino-Reale et al. 2016). Small molecules have been the primary focus of G4 targeting and many G4-interacting compounds have been discovered and studied (Ohnmacht and Neidle 2014). The major mode of interaction between small molecules and G4s are either tetrad binding (stacked on the tetrads) or loop/groove binding (vertical to the tetrads) (Muller and Rodriguez 2014). Studies on many small molecules targeting promoter and RNA G4s; a summary of select G4 stabilizing small molecules and their proposed mechanisms of action is shown in (Table 1-2).

To date, two G4 stabilizing compounds have made it to clinical trials – quarfloxin and CX-5461. Quarfloxin indirectly stabilized the G4 in MYC promoter by the displacement of nucleolin from rRNA G4s, decreased transcription and mediated cell death in pancreatic cancer cells (Denis Drygin et al. 2008). Unfortunately, quarfloxin failed in Phase II trials due to high albumin binding (Balasubramanian et al. 2011). The other compound, CX-5461 is still under investigation in clinical trials for triple negative breast cancer, has been found to be a DNA G4 stabilizer (H. Xu et al. 2017).

The small molecule approach has been the primary strategy for stabilizing G4s, however,
high specificity across structurally homologous G4s is hard to achieve, mediating possible off-target toxicity and affecting the utility of small molecules as G4 modulating anti-cancer agents (Chen et al. 2014). Other than a small molecule strategy, a nucleic acid approach has also been taken to target G4s. Guanine rich peptide nucleic acids (PNAs) was used to form a DNA: PNA hybrid G4 in the MYC promoter, and has been demonstrated to downregulate MYC promoter activity. Another PNA sequence was found to cause scissions in the genome by forming DNA: PNA hybrid G4s (Gupta et al. 2013; Ishizuka et al. 2012). Pu27, the G4-forming region of MYC promoter has been directly applied to leukemia cells, resulting in profound DNA damages in both telomeric and non-telomeric regions triggering cellular apoptosis (Islam et al. 2014; Sedoris et al. 2012). An unmodified DNA oligonucleotide, PNT 2258, was used to target the distal region of the Bcl-2 promoter, resulting in decreased mRNA expression and cellular apoptosis (Ebrahim et al. 2016; Tolcher et al. 2014). While PNT 2258 is a non-G4-targeted therapy, it is the first-in-class DNA interference therapy that has progressed into clinical trials, highlighting the potential of targeting DNA with DNA as a therapeutic strategy. Exploring novel G4-targeting strategies, differing from the conventional small molecule approach, has the potential to enhance the pharmacological specificity of G4 modulation, thus benefitting the development of novel therapeutics.
<table>
<thead>
<tr>
<th>G4-interacting small molecule</th>
<th>Proposed biological activities</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMPYP4</td>
<td>Inhibiting MYC and hTERT by stabilizing promoter G4s, telomerase inhibition, disturbing protein binding to RNA G4s (c9 or f72)</td>
<td>(Grand et al. 2002; Zamiri et al. 2014)</td>
</tr>
<tr>
<td>Telomestatin</td>
<td>Telomerase inhibition, triggering DNA damage response, c-Myb inhibition</td>
<td>(Miyazaki et al. 2012; Shin-ya et al. 2001)</td>
</tr>
<tr>
<td>BRACO19</td>
<td>Telomere uncapping, telomerase inhibition, chromosomal end fusion</td>
<td>(Burger et al. 2005; Taetz et al. 2006)</td>
</tr>
<tr>
<td>RHPS4</td>
<td>Telomere uncapping, telomerase inhibition, hTERT displacement</td>
<td>(Phatak et al. 2007; Salvati et al. 2007)</td>
</tr>
<tr>
<td>Quarfoxin</td>
<td>Inhibiting MYC by the displacement of nucleolin from nucleoli into nucleus</td>
<td>(D. Drygin et al. 2009; Duan et al. 2001)</td>
</tr>
<tr>
<td>Pyridostatin</td>
<td>Altering shelterin binding, DNA damage response, downregulating SRC</td>
<td>(Muller et al. 2010; Muller et al. 2012)</td>
</tr>
<tr>
<td>307A and 360A</td>
<td>Telomere damage, Inducing DNA damage response, downregulating MYC</td>
<td>(Lemarteleur et al. 2004; Pennarun et al. 2008)</td>
</tr>
<tr>
<td>Phen-DC(3) and Phen-DC(6)</td>
<td>Causing genome instability, telomere uncapping, inhibiting the translation of TRF2 by stabilizing its mRNA G4</td>
<td>(Chung et al. 2014; Piazza et al. 2010)</td>
</tr>
<tr>
<td>NSC 338258</td>
<td>Inhibiting MYC by stabilizing its promoter G4</td>
<td>(Brown et al. 2011)</td>
</tr>
<tr>
<td>CX-5461</td>
<td>MYC, c-KIT Promoter G4 stabilization, inducing DNA damage response and genome instability</td>
<td>(H. C. Lee et al. 2017; H. Xu et al. 2017)</td>
</tr>
<tr>
<td>12459</td>
<td>Telomerase inhibition, DNA damage response induction, hTERT RNA G4 stabilization</td>
<td>(Gomez et al. 2003; Gomez et al. 2004)</td>
</tr>
<tr>
<td>CarboxyPDS (RNA G4 selective)</td>
<td>TERRA inhibition, RNA G4 stabilization</td>
<td>(Biffi et al. 2014)</td>
</tr>
<tr>
<td>RGB1 (RNA G4 selective)</td>
<td>TERRA inhibition, inhibiting NRAS translation by stabilizing its 5’-UTR G4</td>
<td>(Katsuda et al. 2016)</td>
</tr>
</tbody>
</table>

Table 1-2. Summary of selected G4-interacting small molecules and their proposed biological activities.
In the departure from the small molecule based stabilizing strategy and from PNAs, we developed a novel nucleic acid (NA) clamp approach that can specifically bind to, and stabilize, individual G4s. It is our central hypothesis that NA-clamps possess highly specific binding and stabilization of the G4s to which they are designed. We explored NA-clamps as specific G4 stabilizing agents at both the transcriptional and translational level. Crucial oncogenes possessing regulatory G4 in the promoter or 5’-UTR, MYC and NRAS, were selected as proof of principle targets to demonstrate our central hypothesis.
CHAPTER II. MATERIALS AND METHODS

2.1. Reagents

Most oligonucleotides and NA-Clamps (Tables 2-1 and 2-2) were synthesized by Eurofins MWG Operon, LLC (Louisville, KY); NRAS WT<sub>ext</sub> and NRAS 6-FAM-WT<sub>ext</sub> sequences (Table 2-2) were purchased from Midland Certified Reagent Co., Inc. (Midland, TX, USA). Most of the chemical reagents and cell culture materials were purchased from either Thermo Fisher Scientific (Waltham, MA), or Sigma-Aldrich (St. Louis, MO) with the exception of Acrylamide/bisacrylamide (29:1) being purchased from Bio-Rad Laboratories (Hercules, CA).

2.2. DNA Dissolution, Concentration Determination and Purification

Upon arrival, all DNA was suspended in nuclease free water and vigorously vortexed until no precipitates were observed. G4-forming sequences were heated at 95 °C for 5 min before concentrations were determined (while the DNA was still hot) utilizing a NanoDrop 2000 spectrometer to measure the absorbance at 260 nm (A<sub>260</sub>) (BioRad Laboratories, CA). The final concentration of the DNA was calculated by the nearest neighbor technique, dividing the A<sub>260</sub> by the extinction coefficient ε<sub>260</sub> (obtained from IDT oligo analyzer) x 10<sup>6</sup>. All fluorescently labeled oligonucleotides were purified by formamide treatment (100% for 15 min at 95 °C); snap cooled on ice and isolated on a 10% polyacrylamide gel as described above and extracted before concentration and determination.
2.3 Electrophoretic mobility shift assay (EMSA)

All DNA samples were prepared by mixing 6-FAM labeled oligonucleotides (1 μM) with Tris-HCl buffer (50 mM, PH7.4), and KCl (25 mM for MYC and 100 mM for NRAS), and acetonitrile (ACN, 40% for NRAS only) before annealing as described above; nucleic acid clamps (varying concentration as denoted in text) were added either pre- or post- annealing. Samples were heated immediately if the NA-clamps were added pre-annealing. If the NA-clamps were added post-annealing, the mixture solution was incubated at room temperature for 30 min in order to allow sufficient binding between oligonucleotides and clamps. For the NA-clamp competition assay, non-fluorescent oligonucleotides were added to the mixture solution simultaneously with the clamps, before the 30 min incubation. dsDNA of the MYC G4-forming Pu46 sequence was constructed by annealing (heating to 95 °C and slow cool to room temperature) the G-rich strand with the C-rich complementary sequences. For all EMSAs, the polyacrylamide gels for DNA isolation were freshly prepared. Solutions were mixed to reach final concentration of 10% acrylamide/bisacrylamide (29:1), 1X Tris/Borate/EDTA (TBE), 0.5% ammonium persulfate (APS), and 0.1% tetramethylethylenediamine (TEMED). Then the mixture was evenly cast into a glass sandwich to allow gel formation. All gels were pre-electrophoresed at 150V for 30 min before DNA samples were loaded; all DNA samples were supplemented with 10% glycerol and 1X DNA loading dye before loading onto the gels. DNA samples were then electrophoresed at 150V until the dye front reached the lower 1/5 of the gel. DNA separation was visualized in a FotoDyne Analyst Investigator FX documentation system equipped with Blue LED epifluorescence. Densitometry analysis was performed by using Image J software (Bethesda, MD, USA) to determine the optical density of each isolated DNA band.
2.4. **Electronic Circular Dichroism (ECD)**

Non-FAM labeled oligonucleotides (5 μM) were dissolved in 50 mM Tris-HCl buffer (pH 7.4) with KCl (25 mM for MYC and 100 mM for NRAS) and 40% ACN (NRAS samples only) either with or without clamps (5 μM). Before annealing, the concentration of all samples were determined as describe. Samples were transferred into a quartz cell of 1 mm optical path length, and before spectra were recorded in millidegrees, with response time as a function of high volts from 225-350 nm. ECD spectra with an Olis DSM-20 spectropolarimeter equipped with a Peltier cell holder (Huntsville, AL). ECD units were converted from millidegrees to molar ellipticity using this olis DSM-20 software to empirically determined concentration.

2.5. **Dimethyl sulfate (DMS) footprinting**

FAM labeled oligonucleotides (10 μM) were dissolved in 50 mM Tris-HCl buffer together with 25 or 100 mM KCl and 40% ACN as described above, in a final volume of 25 μl, and annealed as described above; clamps (10 μM) were added either pre- or post- annealing. The solutions were then treated with DMS by adding 1.3 μL of 10% DMS (DMS : EtOH = 1: 9) to reach a final concentration of 0.5%; 1.3 μl of 90% EtOH was added to a no DMS sample as a control. Samples were then incubated at room temperature for 15 min before the DMS reaction was quenched with 7 μl of DMS stop solution (1 M β-mercaptoethanol, 1 M Tris-HCl, 1.5 M sodium acetate) together with 1 μl of 10 mg/ml calf thymus DNA. Samples were supplemented with 10% glycerol and 1X DNA loading dye before loading onto the pre-run 10% acrylamide gel. DNA species were separated by electrophoresis at 150 V for 1 hr. DNA banding was visualized with a FotoDyne imager. Bands of interest were excised with a scapula, transferred into 1.5 mL amber Eppendorf tubes and crushed with a sterile pipette tip before 250 μl of gel elution buffer (0.4 M ammonium acetate, 1mM MgCl₂, 0.2% sodium dodecyl sulfate) was added. All samples were vortexed for 10
seconds before transferring into a 37 °C shaking incubator overnight. Crushed gel particles were removed by 0.22 μm centrifuge filter tubes and the liquid was transferred into new 1.5 ml amber tubes. 100% EtOH and sodium acetate solution was added to a final concentration of 80% EtOH and 0.3 M sodium acetate. Samples were then vortexed for 10 seconds and stored at -20 °C overnight. Tubes were centrifuged at 4 °C for 15 min at 13,000 rpm to collect the DNA pellets, which were air-dried and resuspended in 50 μl of 1 M (10%) freshly diluted piperidine solution. Samples were heated at 90 °C for 30 min and snap cooled on ice. All samples were ethanol precipitated with sodium acetate as described above, to purify the cleaved DNA. Finally, a 16% denaturing polyacrylamide gel with 8 M urea in 1X TBE was cast in 20x40 cm glass plates and pre-run for at least 30 min at 1600 V. DNA pellets were resuspended in 20 μl of water plus 2 μl of 6X DNA loading dye, heated to 85 °C for 3 min and snap cooled on ice before loading on the gel. Ten μl of each sample was loaded onto the gel and electrophoresed at 2 W, overnight at 4 °C. The gel was then visualized with a FotoDyne FX imager using blue LED epifluorescence. Densitometry analysis was performed with Image J software, and the resulting histograms were overlaid using Photoshop software (Adobe systems, CA).

2.6. Pull down assay

Non-FAM labeled oligonucleotides (5 μM) were dissolved in 50 mM Tris-HCl buffer (pH 7.4) with KCl (25 mM) with 5 μM of biotinylated clamp A 5T, and annealed as described. The solutions were subjected to EMSA, as described above, bands of interests were excised and DNA was isolated and purified by centrifugation. Magnetic streptavidin beads (50 μl) was added to each sample and incubated at room temperature for 1 hr. The beads were concentrated by the application of a magnetic field with a magnetic bar; DNA complexes were released from the beads by boiling with SDS-PAGE reducing buffer at 100°C for 5 min and snap cooling on ice. The recaptured DNA
was subjected to EMSA again for visualization under FotoDyne FX imager. Densitometry analysis was performed using Image J software.

2.7. Cell Culture

All cell lines described in this work were obtained from ATCC (Frederick, MD). Human embryonic kidney (HEK-293) cells and breast cancer (MCF-7) cells were maintained in Dulbecco’s modified eagle’s medium (DMEM) in tissue culture treated flasks. Burkitt’s lymphoma cells, RAJI and CA46, were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium in non-tissue culture treated flasks. All media solutions were supplemented with 10% fetal bovine serum and 1x penicillin/streptomycin. Culture flasks were incubated at 37 °C in an incubator supplied with 5% CO₂; cells were maintained in exponential growth until use.

2.8. Transfection and Luciferase Assay

HEK-293 cells were transfected with the Del4 plasmid (Addgene, Cambridge, MA), which contains the MYC promoter +/- 400 base pairs from the TSS driving luciferase expression. Cells were seeded at 8 x 10⁴ cells/well in 1 ml of medium in 24-well plates and allowed to attach overnight. Cells were co-transfected with either the Del4 plasmid (250 ng) or the promoterless empty vector plasmid pGL4.17 (250 ng) (Promega, WI) and 125 ng of renilla plasmid (pRL-SV40) (Promega, WI) either with or without the indicated NA-Clamps in micelles formed with FuGene HD (Promega, WI) at a 3:1 ratio. Cells and the DNA were incubated for 24 or 48 hr and lysed in 1X passive lysis buffer (Promega, WI) and taken through two freeze/thaw cycles to promote cell lysis. Luciferase was measured with the Dual Luciferase Assay kit (Promega, Madison, WI) on a Lumat LB9507 luminometer; firefly: renilla ratios were recorded. All luciferase experiments were repeated three times, with internal biological duplicates. Luciferase expression changes were normalized to those with the empty vector plasmid; statistical significance was determined by one-
way ANOVA with post-hoc Tukey’s test.

2.9. Cellular Viability Assay (MTS Assay)

For all cytotoxicity assays, cells were plated in 96 well plates with 90 µl of media in each well 24 hr before treatment. MCF-7, HEK-293, and human melanoma WM3682 cells were plated at 2,000, 5,000, 20,000, and 40,000 cells per well for 144, 72, 48, and 24 hr time points respectively. RAJI and CA46 cells were plated at density of 7,000, 50,000, 100,000, 200,000 cells per well for 144, 72, 48, 24 hr time points, respectively. For viability assays, with transfected clamps, Fugene HD or block co-polymers were added and incubated with the DNA for 30 min before cellular transfection. Transfected or just clamps (10 µl) were added to each well. At the end of the incubation times, cellular viability changes were measured using the MTS assay (Mosmann 1983). Methane thiosulfonate (MTS) (2 mg/ml) in Phosphate-buffered saline (PBS) was mixed with 5% phenazine methosulfate (PMS), and 20 µl of the solution was added to each well; plates were incubated at 37 ºC for 2-4 hr. Absorbance at 490 nm was recorded using a BioTek Synergy 2 spectrophotometer (BioTek Instruments, VT, USA), which was normalized to untreated control cells. Non-linear regression for a sigmoidal inhibition dose response was used to determine the IC50 using GraphPad Prism software.

2.10. Cellular Uptake Assay

MCF-7 breast cancer cells or Burkitt’s lymphoma (RAJI and CA46) cells were plated in 96 well plates one day before transfection. MCF-7 cells were plated at a density of 5x10⁴ cells/well in 100 µl of medium and lymphoma cells were plated at a density of 2x10⁵ cells/well in 100 µl of medium 24 hr prior to transfection. FAM labeled clamp A 5T (500 ng) were either transfected with FuGene HD or block co-polymers, or were directly introduced into wells. Five hour post-transfection, Hoechst 33342 (Invitrogen, SF, CA) was added to the wells at a final concentration
of 5 µg/ml and incubated for 1 hr at 37 °C, before media containing clamps was captured and stored in a clean 96-well plate. Cells were washed with PBS for 3 times before being supplied with fresh media. Images were collected in a Cytation 5 live cell imaging station equipped with an imaging chamber maintaining 37 °C supplemented with 5% CO2 (Bio-Rad, CA, USA). The first removed media with clamps was reintroduced into the original wells post imaging. After 24 and 48 hr, media was captured again and cells stained with Hoechst 33342 again as previously before being supplemented with fresh warm media before repeating the imaging process. Scale bars were added to all collected images using the original Bio-Rad cytation 5 processing software.

2.11. ChIP-PCR Assay

MCF-7 cells were plated into a 6 well plate at the density of 5x10^5 cells/well with 2 ml of media one day before transfection. Biotinylated clamp A 5T (10 µg or 20 µg) was transfected, as indicated, using Fugene HD at 3:1 ratio. Cell cultures were then maintained in a 37°C incubator supplemented with 5% CO2 for 24 hrs. One hour before cell lysis, test cells were treated with 2 nM estradiol (E2) to stimulate MYC transcription and cells were placed back into the incubator. Media in the wells was removed 1 hr later and cells were rinsed with sterile PBS three times before adding 300 µl of passive lysis buffer (Promega, WI). Genomic DNA was then purified with a purification kit (Promega), rehydrated and centrifuged at 13,000 rpm for 30 min. The supernatant was discarded and the DNA pellets were air dried for 30 min at room temperature. An enzymatic ChIP kit (Active motif, San Diego, CA) was used to shear the genomic DNA. Digestion buffer (50 µl, supplemented with 1.75 µl protease inhibitor cocktail and 1.75 µl phenylmethane sulfonil fluoride) (Active motif, San Diego, CA) was added to each sample and incubated for 5 minutes at 37 °C. Each sample was treated with 5 µl of enzymatic shearing cocktail (200 U/mL) and incubated at 37°C for 20 min; tubes were vortexed every 5 min during the incubation process. Ice cold 0.5
M EDTA (1.5 μl) was added to each tube, and all samples were chilled on ice for 10 min to stop the shearing reaction. Magnetic streptavidin beads (50 μl) were added to each tube and the samples were incubated at room temperature for 1 hr to allow sufficient biotin-streptavidin interaction. No beads were added to the positive controls (input); and 50 μl of beads were added to the no clamp control group. A magnetic field was applied to all tubes to pull down the DNA complex associated with biotinylated clamps. The precipitates were washed with nucleus free water three times and then suspended in 10 mM EDTA (pH 8.2) with 95% formamide. All samples were heated at 90 °C for 10 min and snap cooled on ice. DNA was then precipitated with 80% EtOH supplemented with 0.3 M sodium acetate at -20 °C, overnight. DNA pellets were collected by centrifugation at 13,000 rpm for 30 min and resuspended in 30 μl of nuclease free water. The DNA concentration was determined by using a Nanodrop 2000 as described; 500 ng of template DNA and MYC primers (RB 26F and RB 26R) were used for PCR amplification with hot start Taq DNA polymerase kit (Phenix, Candler, NC). Cycling conditions were 94 °C for 3 min followed by 30 cycles of 94 °C for 30 s, 55°C for 30 s, and 72°C for 60 s. Post PCR amplification, DNA samples were subjected to agarose gel (supplemented with GelGreen (Phenix)) electrophoresis at 100 V for 40 min. Bands were visualized under a FotoDyne FX imager. Densitometry analysis was performed using Image J software.

2.12. Quantitative Real-Time PCR (qRT-PCR)

MCF-7 were plated in 12 well plates at the density of 2.5x10^5 cells/well in 1 ml of DMEM; RAJI and CA46 cells were plated in 12 well plates at the density of 1x10^6 cells/ml in 1ml of RPMI 1640 media. All cells were cultured in a 37 °C incubator supplemented with 5% CO₂ for 24 hr before transfection. Clamp A 5T (2.5 μg) was transfected into each well with Fugene HD at 3:1 ratio. Estradiol (2 nm) was added to MCF-7 cells 1 hr before cell lysis. After 24 hr transfection,
cell pellets collected from each plate were lysed with lysis buffer (Thermo Scientific, SF, CA) supplemented with 2% β-mercaptoethanol. RNA samples were extracted with a GeneJET RNA purification kit (Thermo Scientific). RNA concentrations were determined using a Nanodrop 2000; the purity of RNA was monitored with the A260/A230 ratio and only samples with values >2 were used in the study. cDNA synthesis was conducted using an Applied Biosystem cDNA synthesis kit (ABI, SF, CA). 500 ng (MCF-7) and 2 μg (RAJI and CA46) of mRNA were reverse transcribed. qRT-PCR was performed on a BioRad CFX Connect™ Real-Time PCR Detection System (BioRad, CA) using Taqman primers from ABI (MYC Exon 1: Hs01562521_m1, MYC Exon 2: Hs00153408, GAPDH: Hs02758991_g1). mRNA expression level was normalized to GAPDH and again vehicle controls, using the ΔΔCq method. All experiments were run in biological triplicates with qPCR run with technical duplicates. One-way ANOVA with post hoc-Tukey’s test was used to determine the statistical significance using GraphPad Prism software.

2.13. G4 Induction assay

MCF-7, RAJI, and CA46 cells were plated into 96-well plates one day before treatment with the MYC selective G4-stabilizing compound NSC 338258 (Brown, et al, JBC, 2011) at 0.5X, 1X and 1.5X IC50 of 24 hr for each cell line. Six hr post treatment, cells were fixed with 4% formaldehyde for 15 min and permeabilized with 0.2% Triton-X 100 for 10 min at room temperature. Cells were then incubated with 0.1 μM of FAM labeled clamp A 5T or control ODN (negative control) and 5 μg/ml Hoechst 33342 at room temperature for 1 hr. Cells were washed with PBS for three times before imaging on a Cytation 5 imaging station. The fluorescence was normalized to the control cell (no treatment group) by deducting the auto fluorescence of control cells.
2.14. Western Blotting

Melanoma WM3682 cells were plated in 12 well plates at a density of $3 \times 10^5$ cells/well in 1 ml of Leibovitz's L-15 medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin 24 hr before transfection. Varying concentrations of NRAS clamp 1 and clamp 2 (500, 1000, 2500, 5000 ng) were transfected into cells using Fugene HD at a 3:1 ratio. Cells were lysed 24 hr post-transfection using radio immunoprecipitation RIPA buffer (Thermofisher Scientific). Plates were shaken for 30 min before transferring the lysates into microcentrifuge tubes. The lysates were centrifuged for 20 min at 12,000 rpm at 4 °C before the supernatants were collected. The bicinchoninic acid (BCA) assay was used to determine the protein concentrations of all lysates. Briefly, BCA working reagents were prepared first by mixing BCA reagents A and reagent B at a 50:1 ratio. Bovine serum albumin (BSA, 30 µl of 2 mg/ml) were transferred into the first well for the BSA standard followed by series dilution (10 times) in RIPA buffer; the mixture working reagents (200 µl) were then added into each well of a 96-well plate. For sample measurements, samples collected from lysis buffer were diluted 10 fold then 30 µl of each diluted cell lysate was added to wells containing 200 µl of working reagents. Plates were incubated at 37 °C for 30 minutes prior to absorbance reading at 562 nm on a Bio-Tek Synergy II spectrophotometer. Protein concentrations were calculated based on the standard BSA curve by linear regression. After obtaining the total protein concentration, 30 µg of total protein from each sample denatured by heating to 95 °C for 5 minutes, and loaded onto an 8% pre-cast bis-polyacrylamide gel (BioRad) along with a molecular weight ladder. The gel was electrophoresed at 100 V for 1 hr. For protein transfer, a polyvinylidene difluorid (PVDF) membrane pre-activated with methanol for 1 minute and rinsed with transfer buffer (25 mM Tris, 190 mM glycine, 20% methanol) was used. Proteins were transferred from the gel to the membrane
at 25 V overnight. For immunostaining, the membranes were blocked by soaking at with 5% BSA in TBST overnight at 4 °C. Primary purified mouse monoclonal NRAS antibody (Abgent, SD, CA) diluted in blocking buffer at a 1:5000 ratio was incubated with the membranes overnight at room temperature on a shaker. Membranes were then washed with TBST (Tris-buffered saline, 0.1% Tween 20) three times for 5 minutes each, then incubated with fluorescent-labeled (LiCOR IRDye 800CW) secondary anti-mouse antibody (LICOR Biosciences, NE, USA) in blocking buffer for 1 hour. The membranes were washed again with TBST three times. Finally, images were acquired on a Li-COR Odyssey CFX imaging system. Densitometry analysis was performed on the Li-COR Odyssey software.
<table>
<thead>
<tr>
<th></th>
<th>Oligonucleotide Sequences (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pu46</td>
<td>GCGCTTATGGGGAGGTTGGGAGGTTGGGAGGGAGGAC</td>
</tr>
<tr>
<td>mt5</td>
<td>GCGCTTATGGGGAGGTTGGGAGGTTGGGAGGGAGGAC</td>
</tr>
<tr>
<td>mt3</td>
<td>GCGCTTATGGGGAGGGTTTTAGTGGGAGGTTGGGAGGAGGAC</td>
</tr>
<tr>
<td>mt1,6</td>
<td>GCGCTTATGGGTAGGTGGGAGGTTGGGAGGTTGGGAGGAGGAGGAC</td>
</tr>
<tr>
<td>mt2,3</td>
<td>GCGCTTATGGGAGGTGGTGGGAGGTTGGGAGGTTGGGAGGAGGAC</td>
</tr>
<tr>
<td>mt1,2,3,4</td>
<td>GCGCTTATGGGTAGGTGGTGGGAGGTTGGGAGGTTGGGAGGAGGAC</td>
</tr>
<tr>
<td>[6-FAM] Pu46</td>
<td>[6-FAM] GCGCTTATGGGGAGGTTGGGAGGTTGGGAGGAGGAC</td>
</tr>
<tr>
<td>[6-FAM] mt5</td>
<td>[6-FAM] GCGCTTATGGGGAGGTTGGGAGGTTGGGAGGAGGAC</td>
</tr>
<tr>
<td>[6-FAM] mt3</td>
<td>[6-FAM] GCGCTTATGGGGAGGTTTTAGTGGGAGGTTGGGAGGAGGAC</td>
</tr>
<tr>
<td>[6-FAM] mt1,6</td>
<td>[6-FAM] GCGCTTATGGGTAGGTGGGAGGTTGGGAGGTTGGGAGGAGGAC</td>
</tr>
<tr>
<td>[6-FAM] mt2,3</td>
<td>[6-FAM] GCGCTTATGGGGAGGTGGTGAGGTGGGAGGTTGGGAGGAGGAC</td>
</tr>
<tr>
<td>[6-FAM] mt1,2,3,4</td>
<td>[6-FAM] GCGCTTATGGGTAGGTGGTGAGGTGGGAGGTTGGGAGGAGGAC</td>
</tr>
<tr>
<td>Clamp A</td>
<td>CTCTCCCCACCCTCCCC[C-18]ATAAG</td>
</tr>
<tr>
<td>Clamp A 5’</td>
<td>CTCTCCCCACCCTCCCC</td>
</tr>
<tr>
<td>Clamp A 5’ link</td>
<td>CTCTCCCCACCCTCCC[C-18]</td>
</tr>
<tr>
<td>Clamp A 3’</td>
<td>ATAAG</td>
</tr>
<tr>
<td>Clamp A 3’ link</td>
<td>[C-18]ATAAG</td>
</tr>
<tr>
<td>Clamp B</td>
<td>CTGAGTCTCCTCCCCACCTT[C-18]TCCCC</td>
</tr>
<tr>
<td>Clamp C</td>
<td>CTGAGTCTCCTCCCCACCTT[C-18]ATAAG</td>
</tr>
<tr>
<td>Clamp A 5T</td>
<td>CTCTCCCCACCTTCCCCCCTTTTATAAG</td>
</tr>
<tr>
<td>[6-FAM] Clamp A 5T</td>
<td>[6-FAM] CTCTCCCCACCTTCCCCCCTTTTATAAG</td>
</tr>
<tr>
<td>[Biotin] Clamp A 5T</td>
<td>[Biotin] CTCTCCCCACCTTCCCCCCTTTTATAAG</td>
</tr>
<tr>
<td>5T</td>
<td>TTTTT</td>
</tr>
<tr>
<td>Clamp A 5’ 5T</td>
<td>CTCTCCCCACCTTCCCCCCTTT</td>
</tr>
<tr>
<td>Clamp A 3’ 5T</td>
<td>TTTTTATAAG</td>
</tr>
<tr>
<td>Clamp A 4T</td>
<td>CTCTCCCCACCTTCCCCCCTTTTATAAG</td>
</tr>
<tr>
<td>Clamp A 3T</td>
<td>CTCTCCCCACCTTCCCCCCTTTTATAAG</td>
</tr>
<tr>
<td>RB 26F</td>
<td>CGATGATTATATCTCACAGGACAAG</td>
</tr>
<tr>
<td>RB 26R</td>
<td>GCGAGAGAGAGATTTTTCTTCTTCC</td>
</tr>
<tr>
<td>[6-FAM] Control ODN</td>
<td>[6-FAM] AGTCCTCCTCCTCGCAAA</td>
</tr>
</tbody>
</table>

Table 2-1. MYC oligonucleotide sequences (5’-3’).
| NRAS WT | TGTGGAGGGGCGGCTCTGGGTGC |
| NRAS MT | TGTGGGATTTTCCGGTCTGGGTGC |
| NRAS WT<sub>ext</sub> | GAAAAGTCCCCGTGTTGGAGGGCGGGCTCTGGGTGCGCCTGCGCA |
| NRAS MT<sub>ext</sub> | GAAAAGTCCCCGTGTTGGGGATTTCGGGTCTGGGTGCCTGCGCA |
| NRAS clamp 1 | TGGCGCAGGCGCA(C-18)ACACGGGACGTTC |
| NRAS clamp 2 | GCAGGCCGCA(C-18)ACACGGGACG |
| [6-FAM] NRAS WT<sub>ext</sub> | [6-FAM] GAAAAGTCCCCGTGTTGGAGGGCGGGCTCTGGGTGCGCCTGCGCA |
| [6-FAM] NRAS MT<sub>ext</sub> | [6-FAM] GAAAAGTCCCCGTGTTGGGGATTTCGGGTCTGGGTGCCTGCGCA |

Table 2-2. NRAS oligonucleotide sequences (5’-3’).
CHAPTER III. TARGETING MYC PROMOTER G4 USING NA-CLAMPS

3.1. Introduction

G-quadruplexes (G4s) are non-canonical secondary DNA structures that are formed by guanine-guanine, instead of guanine-cytosine, base pairing. A G4 structure can occur if there are more than four consecutive guanine runs with at least three guanines in a row in any nucleic acid sequence. Genome wide sequencing analysis, together with visualization studies in human cancer cells have demonstrated that G4 structures form in human cells, generally clustering around the regulatory, nucleosome-depleted, regions of DNA and RNA, indicating that they play important functional roles in the regulation of oncogene expression at both transcriptional and translational levels (Biffi et al. 2013; Biffi et al. 2014; Verma et al. 2008). As discussed previously, potential G4-forming sequences were discovered to exist in over 50% of human gene promoters, especially these associated with the cancer hallmarks (Huppert and Balasubramanian 2007; Jayaraj et al. 2012). As shown in Table 1-1, many DNA and RNA G4s form in critical cancer associated genes across the ten hallmarks of cancer. These G4s formed in various locations within the genome can modulate transcription, translation, direct splicing of introns, and control telomere associated signaling pathways that are closely correlated with tumorigenesis (Balasubramanian et al. 2011; Brooks and Hurley 2010; Brown et al. 2011; Cammas and Millevoi 2017; Ivancich et al. 2017; Kumari et al. 2007; Marcel et al. 2011; Phatak et al. 2007; Siddiqui-Jain et al. 2002)

Among the oncogenic genes that contain biologically active G4s, the promoter of proto-oncogene MYC was the first identified, and has been the most, studied. The discovery of the
relationship between MYC overexpression and chromosomal translocation in Burkitt’s lymphoma cells marked MYC as a major oncogene (Dalla-Favera et al. 1982). MYC is a basic helix-loop-helix/leucine zipper transcription factor that regulates up to 15% of gene expression genome-wide with over 30,000 potential binding sites (Patel et al. 2004). MYC has been described as a symphony conductor that regulates a wide range of cellular functions from cellular proliferation, differentiation, metabolism, apoptosis, and metastasis to angiogenesis. However, the best-known function of MYC is driving cancer cell proliferation, and alterations of MYC are commonly found to relate to tumorigenesis (Dang 2010).

Initially, the ability of MYC to drive cellular proliferation was thought to be associated with its ability to directly regulate DNA replication, which was supported by the early evidence in several studies including: (1) In the Xenopus laevis oogenesis process, maternal cytoplasmic MYC was translocated to the nucleus during a developing period with no transcriptional activity and activates rapid replication of these non-dividing cells (Gusse et al. 1989); and (2) polyclonal and monoclonal MYC antibodies can inhibit DNA synthesis and polymerase activity in human leukemia cell lines that contain a high concentration of MYC (Studzinski et al. 1986). Their studies pointed out that MYC is regulating and activating DNA synthesis and driving cellular replication; however, research using MYC antibodies produced in other laboratories did not recapture a similar phenomenon. Gutierrez’s research pointed out the replication inhibition of MYC antibody from watt lab’s study might result from the contamination of their antibody preparation process by an unknown DNA polymerase inhibitor (Dominguez-Sola et al. 2007; Gutierrez et al. 1988).

In the late 1980s, interest in further studying the relationship between MYC and DNA replication was overshadowed by studies on MYC’s role as a transcriptional regulator (Luscher and Eisenman 1990; Marcu et al. 1992; Meyer and Penn 2008). The MYC protein contains an N-
terminal transcriptional regulatory domain followed by a nuclear localization sequence and a C-terminal domain comprised of a basic helix-loop-helix-leucine zipper (HLH-/Leu-zipper) dimerization motif. As a transcription factor, MYC dimerizes with another protein, MAX, through the HLH/Leu-zipper domain to facilitate DNA binding. The MYC-MAX complex can bind E-boxes with a consensus sequence (5’-CACGTG-3’) (Amati et al. 1993; Kerkhoff et al. 1991). It can either activate or repress its target gene expression by different mechanisms. For activation, MYC binds to E-boxes in the promoter of the target gene and recruits multiple proteins, including RNA Pol II, transcription cofactors, and chromatin remodeling complexes to activate transcription (Eberhardy and Farnham 2002; Gomez-Roman et al. 2006; Kanazawa et al. 2003; Park et al. 2001). On the other side, MYC represses gene expression either epigenetically, by recruiting proteins like histone methyltransferases to reduce chromatin accessibility, or by binding to other transcriptional factors, such as Miz1 and Sp1, to repress their transactivating capability on the expression of genes that are normally cell cycle inhibitors or differentiation signaling (Amati et al. 2001; Herkert and Eilers 2010).

In addition to being a transcription initiation, MYC affects transcriptional elongation globally throughout the nucleus by affecting the phosphorylation of the C-terminal domain (CTD) of RNA Pol II (Cowling and Cole 2007). MYC binds to and recruits CTD kinases and other transcriptional factors to the promoters of target genes to induce further phosphorylation of RNA Pol II and thus transcriptional elongation (Bouchard et al. 2004; Cowling and Cole 2007; Eberhardy and Farnham 2002).

As a key regulator of cellular proliferation, differentiation, apoptosis and cellular growth, MYC also participates in significant transcription-independent biological functions (Cole and Cowling 2008). Previous research has shown that a MYC mutant protein, devoid of DNA binding
capacity, promotes global phosphorylation of RNA Pol II, thus activating transcription. Also, the transcription activation of MYC is associated with MYC’s function as a facilitator for the 5’ capping and methylation of mRNA, and protein translation by the recruitment of transcription factor IIH (TFIIH) and other co-factors (Cowling et al. 2006; Cowling and Cole 2007). In addition, MYC directly interacts with many components of the pre-replicative complex, such as mini chromosome maintenance complex (MCM) proteins and Carbohydrate deficient transferrin 1 (CDT1), and initiates replication by binding to DNA replication origins (Dominguez-Sola et al. 2007).

As a master commander of cellular activities, especially proliferation and metabolism, the overexpression of MYC can play an essential role in oncogenesis. MYC expression is strictly controlled and usually expressed at low levels, both transcriptionally and translationally, under normal cellular proliferation circumstances. In these non-transformed cells, MYC expression is balanced by positive or negative signal transduction pathways such as WNT, MAPK, and Notch, or TGF-β, respectively (Oster et al. 2002). MYC is overexpressed in over 80% of cancers via various mechanisms, including (1) insertional mutagenesis, (2) chromosomal translocation, and (3) gene amplification, all of which enable MYC signaling independent from the extracellular context (Meyer and Penn 2008). As MYC participates in many key cellular functions of a cell, such as DNA replication, RNA processing, differentiation, cell cycle regulation, transcriptional factor and tumor suppressor interaction, and cellular apoptosis, dysregulation and aberrant expression of the oncogene is keenly involved in tumorigenesis (Dang 2010).

The first mechanism that enables MYC deregulation is retroviral insertional mutagenesis. The discovery of the MC29 v-gag-myc chimeric protein in leukemia in 1979 was the first demonstration that retroviral promoter insertion into a non-mutated cellular gene can account for
tumorigenesis and MYC became the first oncogene known to be activated through this mechanism (Payne et al. 1982; Varmus 1984). Since the discovery of v-myc, it has been shown that the overexpression MCY can readily transform various types of cells, including chicken embryonic fibroblasts, murine bone marrow, and hematopoietic cells with high efficiency (Chisholm et al. 1992; Pirami et al. 1991; Ramsay et al. 1990).

Similar to the well-known BCR-ABL1 fusion protein that resulted from a chromosomal translocation (Gabert et al. 1989), the second mechanism that deregulates MYC is also based on chromosomal translocation that was first discovered in Burkitt’s lymphoma cells. In these cells, the immunoglobulin (Ig) light or heavy chain on chromosome 2, 14, or 22 are translocated to chromosome 8, proximal to the MYC locus. As a result, Ig chains are situated by the MYC gene and function as transcriptional enhancers for MYC in these immune cells (Dalla-Favera et al. 1982). This juxtaposition enables MYC with unrestrained expression and transform these cells (L. M. Boxer and Dang 2001).

The third mechanism that deregulates MYC is gene amplification where greater copy numbers of the MYC gene increase expression, thus transforming the cells. This amplification phenomenon was reported in the early 1980s from different groups on several human cancer cell types. The first copy number alteration of MYC was discovered in human promyelocytic cell line HL-60 and primary leukemia cells from the same patient (Dalla Favera et al. 1982). Later, amplified gene copies of MYC was noted in colon carcinoma, neuroblastoma, breast cancer, small cell lung cancer and diffuse large B-cell lymphoma, which further demonstrated the importance of amplification of MYC on tumorigenesis (Alitalo et al. 1983; Escot et al. 1986; Nau et al. 1985; Slavc et al. 1990; Stasik et al. 2010).

In addition to viral promoter insertion, chromosomal translocation, and gene amplification,
other mechanisms that also activate MYC expression include the activation of hormones, growth factors and their receptors, activated upstream signaling pathways, and mechanisms that either directly or indirectly stabilize MYC mRNA or protein, such as disrupted ubiquitination of mRNA in cells (Dani et al. 1984; Dubik and Shiu 1988; Lacy et al. 1989; Meyer and Penn 2008; Spencer and Groudine 1991).

In contrast to normal cells, MYC overexpressing cancer cells can become obligated to the fast-paced and non-restricted cellular growth and proliferation (Dang 2012). MYC-induced cancers are addicted to the downstream benefits of MYC overexpression, such as nutrients and growth signals; however, they are addicted to the MYC protein itself in many human tumor types, including breast cancer, lymphomas, nasopharyngeal carcinoma, and colon cancer (Brown et al. 2011; Niu et al. 2015; Y. H. Wang et al. 2005; X. Zhang et al. 2009). Turning off ectopic MYC expression in these cancer cells triggers an imbalance between biogenesis and nutrient supply and subsequent cell death. Notably, not all MYC overexpressing cancer cells are susceptible to MYC deactivation, possibly due to multiple oncogenic dysregulation, such as kRAS and VEGF, to sustain cancer cell survival (R. B. Boxer et al. 2004; Jain et al. 2002; Soucek et al. 2008).

On the path to cancer, MYC not only contributes to the sustained proliferation and heightened biogenesis, but also to genomic instability and cell metastasis (Dang 2012). Previous researchers has established a solid correlation between MYC overexpression and increased genomic instability that in turn leads to tumor initiation (Kuzyk and Mai 2014). As discussed previously, MYC regulates a large network of cellular activities including transcription activation, DNA replication, translation, RNA processing, microRNA expression profile regulation and heightened ribosomal and mitochondrial biogenesis. All of these hyperactive states, which are linked to increased mutational rates, boosted reactive oxygen species (ROS) generation, telomere
damage and DNA repair, set the stage for the genomic alterations in MYC overexpressing cancer cells (Egler et al. 2005; Louis et al. 2005; Prochownik 2008; Vafa et al. 2002; Wade and Wahl 2006). In addition to inducing genomic instability, MYC facilitates epithelial-mesenchymal transition (EMT) and subsequent invasion, migration, and metastasis, by inhibiting genes regulating cell to cell contact and attachment, such as E-cadherin (Rapp et al. 2009; Wolfer and Ramaswamy 2011). With all of these capabilities of MYC transforming cell, targeting MYC and deactivating MYC expression is a viable therapeutic pursuit.

It is clear that MYC is a key oncogene that plays an essential role in the genesis and progression of many cancer types. Many studies have demonstrated that inhibiting MYC can trigger cellular apoptosis in a tissue- and cell-type-dependent manner, as discussed above. Inhibiting MYC expression is well tolerated in a mouse model, where whole body knockout decimated cancer development but only reduced proliferation rates in the skin, testes, gastrointestinal tract, and hematopoietic cells with no overt phenotypic effects (Soucek et al. 2008). Moreover, a threshold of MYC expression was required to maintain the tumor phenotype; decreasing MYC expression below this threshold was sufficient to induce the apoptosis of T-cell lymphocytes (Shachaf et al. 2008). This research indicates that MYC inhibition can be a practical strategy for treating sensitive human cancers.

A variety of strategies have been utilized to target this well-established driver of cancer. However, there are hurdles that need to be overcome to successfully inhibit the MYC protein or its expression. As a transcription factor, MYC is localized in the nucleus, which erased the possibility of using antibody based strategies to block its activity due to the size of antibodies. As a protein lacking enzymatic function and small molecule binding pockets on the surface, it has been described as “undruggable” (Posternak and Cole 2016). So far, several methods have been
employed to regulate MYC expression and function, mainly by the following five strategies: (1) blocking the interaction between MYC and MAX; (2) inhibition of bromodomain-containing protein BRD4; (3) targeting cyclin-dependent kinase 7 (CDK7); (4) synthetic lethality of MYC overexpressing cells; and (5) targeting MYC expression by stabilizing its promoter G4. Among these strategies, only number one and five inhibits MYC function or expression directly (Brooks and Hurley 2010; Posternak and Cole 2016).

MYC functions primarily through dimerization with MAX using the HLH/Leu-zipper motif to achieve DNA binding at a consensus sequence (5’-CACGTG-3’) of its targets genes’ E-boxes, oncogenic activity of MYC is dependent on this interaction with MAX (Amati et al. 1993; Blackwood and Eisenman 1991; Kerkhoff et al. 1991). Significant effort has been devoted to disrupting the dimerization between MYC and MAX in order to inhibit MYC’s function as a transcription factor. Several small molecule inhibitors have been developed to achieve this goal (Prochownik and Vogt 2010). The most potent small molecular developed so far that inhibits MYC-MAX dimerization is KJ-Pyr-9, which has shown efficacy on inhibiting the expression of MYC target genes and retarding cell cycle progression, as well as inhibiting cell proliferation in vitro and in vivo (Hart et al. 2014). However, this class of compounds functions through targeting the HLH-LZ domain, they can non-specifically bind to any other proteins also containing a HLH-LZ domain in the cell, such as HIF. Other researchers have tried to inhibit the binding of MYC-MAX complex to DNA by utilizing a mutant MYC HLH-LZ domain (Omomyc), which forms a homodimer itself, as a competing binder to MYC/MAX complex, thus inhibiting its function (Savino et al. 2011; Soucek et al. 1998). This mutant domain triggers apoptosis of MYC overexpressing tumor cells in vitro and in vivo (Jung et al. 2017; Soucek et al. 2002; Soucek et al. 2008).
Indirectly, MYC has been targeted by bromodomain (BRD) and extra terminal (BET) inhibitors, specifically inhibitors for BRD4 that belongs to the mammalian BET protein family together with BRD2, BRD3, and BRDT. BET proteins comprise two tandem bromodomains and an extra-terminal domain. Functionally, BRD4 binds to acetylated lysines (Ac-K) on histones, together with the binding to P-TEFb. BRD4 can bind to and activate P-TEFb, which in turn phosphorylates the CTD of RNA Pol II, thus sustaining transcriptional elongation (Itzen et al. 2014). As discussed previously, MYC not only initiates transcription, but also promotes transcription elongation globally when overexpressed by phosphorylating the lysine on the CTD of RNA Pol II (Cowling and Cole 2007). Even though the mechanism of MYC downregulation by BET inhibitors is still unknown, it is plausible that by inhibiting BRD4 and P-TEFb function, more MYC is recruited as a general kinase for CTD, thus leaving less for transcription initiation of its target genes. Small molecules have been developed to inhibit the binding between BRD4, and Ac-K. The most well studied compound is JQ1, which binds to the Ac-K binding site of BRD4 thus preventing transcriptional elongation (Filippakopoulos et al. 2010). Intriguingly, studies have found that treatment with JQ1 results in over 80% MYC downregulation within 1 hour in multiple myeloma cells (Aird et al. 2017; Delmore et al. 2011). Multiple studies have evaluated the efficacy of this compound against a variety of leukemias, lymphomas, ovarian, gastric, and triple negative breast cancers in vitro and in vivo, which revealed that JQ1 leads to tumor regression by downregulating MYC and its downstream targets (Bid et al. 2016; Delmore et al. 2011; Mertz et al. 2011; Montenegro et al. 2016; Z. Zhang et al. 2016). In terms of toxicity, JQ1 produces a dose-dependent decrease of MYC expression in immune and lymphoid cell compartments and lead to a decrease in B- and T- cell counts in the peripheral system at an efficacious dose-range (D. U. Lee et al. 2016). Currently, seven bromodomain inhibitors, including GSK525762 (NCT01587703),
BAY 1238097 (NCT02369029), MK-8628 (NCT02698176), RO6870810/TEN-010 (NCT02308761), CPI-0610 (NCT01949883), BMS-986158 (NCT02419417), and INCB057643 (NCT02711137) are in clinical trials at varying stages for multiple hematopoietic and solid tumors (clinicaltrials.gov). With current limited trial results that suggest promise for BET inhibitors on several cancer types, further investigations and trial data are needed to determine the clinical impact for these small molecules and BET inhibition approach for cancer treatment (Aird et al. 2017).

CDK7 is the catalytic subunit of transcription factor II H (TFIIH), which phosphorylates the CTD of RNA Pol II for transcription initiation (Egly and Coin 2011). Small molecules, including benzamide THZ1 and THZ2, have been developed for inhibiting CDK7 activity. Initial research has shown that THZ analogues are able to reduce tumor size, halt proliferation, and induce apoptosis in a variety of MYC and MYC-N driven tumors, including neuroblastoma, small cell lung cancer, and triple negative breast cancer. These effects have been attributed to the significant downregulation of MYC due to the inhibition of super enhancers upstream of the gene that are responsible for transcription (Chipumuro et al. 2014; Christensen et al. 2014; Y. Wang et al. 2015b).

Other efforts have been made to target the synthetic lethality of cancer cells that are overexpressing MYC. Previous research has revealed that MYC-driven cancer cells are susceptible to the manipulation of certain pathways, including aurora B kinase, AMPK related kinase 5, spliceosome core factors, RAS phosphorylation pathways and glucose uptake pathways that are critical for cancer cell growth or survival (Hsu et al. 2015; L. Liu et al. 2012; Sabnis et al. 2017; Sears et al. 2000; Yang et al. 2010). Further investigations towards the therapeutic opportunity of
these pathways or a combination of multiple pathways that are lethal to cancer cells should be explored in future to facilitate MYC targeted cancer therapy.

Lastly, but most importantly for the current work, targeting MYC expression directly through the stabilization of the promoter G4 has been well characterized (Brooks and Hurley 2010). The MYC G4 in the NHE III_1 region upstream of the P1 promoter is a well-characterized biologically active promoter structure that represses transcription (Simonsson et al. 1998). The MYC gene possesses four promoters – P0, P1, P2, and P3 (Levens 2008; Wierstra and Alves 2008). Among these four promoters, P1 and P2 are the major promoters that initiate most of the transcription of MYC. Previous studies have shown that these two start sites, and most notably P2 (75-90%), initiate 90% of MYC transcription, while P0 and P3 each initiate less than 5% of transcription (Albert et al. 2001; Levens 2008). Just upstream of the P1/P2 sites, there is a nucleus hypersensitivity element (NHE III_1) region (Figure 3-1) that is guanine- and cytosine-rich, with the ability to form higher order structures, including G4s and i-motifs, during transcription. Early studies demonstrated that this NHE III_1 region is required for the transcription initiation of MYC regardless of which promoter (P1 or P2) is being used (Berberich and Postel 1995; Simonsson et al. 1998). During transcription, the negative superhelicity generated by transcriptional machineries enables the dynamic equilibrium between single stranded, double stranded, and secondary DNA structures within the NHE III_1 region (Brooks and Hurley 2009). Instead of being relieved from the negative superhelicity generated by transcription immediately through the action of topoisomerase I and II, the supercoiling in the MYC promoter is maintained by the melting of the Far Up Stream Element (FUSE) 1.7 kb upstream of the P2 promoter site (J. Liu et al. 2006). This FUSE element can bind to FUSE binding protein (FBP) and FBP interacting repressor (FIR) protein to positively and negatively, respectively, control the rate of promoter firing as a cruise
control system (Kouzine et al. 2008; Lavelle 2008). Due to this maintained torsional stress, the NHE III_1 region is able to adopt G4s and i-Motifs. When the DNA is in duplex form, Sp1 can bind and activate transcription. During transcription; in single stranded form, CNBP and hnRNP K bind to the G- and C-rich strands, respectively, and activate the transcription. However, when there is G4 or i-motif formation, transcription is deactivated (Gonzalez and Hurley 2010). Proteins nucleolin and NM23-H2 have been identified as proteins that assist the folding and unraveling of the G4s formed in this region, respectively (Dexheimer et al. 2009; Gonzalez et al. 2009).

Within the NHE III_1, there are six guanine runs that facilitate the formation of multiple isoforms of G4s. Studies on the linear DNA oligonucleotide Pu27, which contains the first five guanine runs of the NHE III_1 region, highlighted the participation of guanine runs II, III, IV, and IV with the loops of 1:2:1 (5’-3’), that adopts a parallel G4 conformation (Siddiqui-Jain et al. 2002). However, DMS footprinting studies using plasmid DNA that contained the MYC G4-forming region revealed that the predominant physiologically relevant G4 isoform formed under supercoiled conditions is G4_1-4, a parallel structure with loops of 1:2:1 in the 5’ to 3’ direction, which involves the participation of guanine runs I, II, III, and IV, despite its lower thermal stability compared to G4_2-5 in a later study (Mathad et al. 2011; Sun and Hurley 2009).
Figure 3-1. MYC promoter structure and the NHE III₁ region capable of forming G4s. Within the MYC promoter, there are four transcription start sites from P0 through P3, with P1 and P2 initiate majority of the transcription. –142 to –115 base pairs upstream of the P1 promoter, there is a NHE III₁ region that regulates 90% of the transcription. This region is guanine and cytosine rich and contains six consecutive guanine runs capable of forming multiple G4s. As shown in the figure, the major isoform of G4 isoform in single stranded (ssDNA) is G₄₂₅ while the physiologically relevant G4 isoform is G₄₁₄ formed in supercoiled DNA (scDNA), both structure share loops of 1:2:1 configuration from 5’ to 3’ (Adapted from Brook & Hurley, 2009; González, Verónica & Hurley, 2010).
MYC G4s have been demonstrated to be transcriptional silencers by the sequestration of transcriptional binding sites (Brooks and Hurley 2009, 2010; Gonzalez and Hurley 2010). By far, the most marked efforts have been exerted to discover small molecules that stabilize the MYC G4 to downregulate MYC transcription. The most successful G4 stabilizing compound (Quarfloxin, NCT00780663), as discussed above, went to Phase II clinical trials in patients with low to intermediate grade neuroendocrine carcinoma. However, this compound stopped at phase II trials due to high albumin binding (clinicaltrials.gov). Interestingly, a recent study revealed that one analogue of Quarfloxin, CX-5461, which is already on active phase I/II clinical trials (NCT02719977) for multiple solid tumor types, is a G-quadruplex stabilizer that induces structure formation and chromosomal instability in vivo (H. Xu et al. 2017). CX-5461 is an inhibitor of rRNA synthesis that selectively inhibits Pol I-driven transcription of rRNA with nanomolar range of IC\textsubscript{50} on a variety of cancer cell lines, it showed no effect on Pol II, and selectivity inhibits rRNA transcription over DNA replication and protein translation. Also, this compound is demonstrated to be orally bioavailable and present in vivo antitumor activity against murine xenograft human tumor models, especially MIA PaCa-2. Currently, it has been tested on several patients showing promising tumor regression, further investigation is needed to determine if it possess better efficacy than current standard care (clinicaltrials.gov).

These data strongly suggest that G4-interacting small molecules can be effective as cancer therapeutics. Specifically, for MYC, many G4-interacting compounds ranging from perylene derivatives, cationic porphyrins, quindolines, DNA dyes, alkaloids, metal complexes, carbamide and its analogues, and many other ligands, stabilize the MYC promoter G4. Notably, an ellipticine compound called NSC 338258 has been identified as an active and moderately selective MYC G4 stabilizer, with nanomolar affinity. In vitro studies on Burkitt’s lymphoma cell lines RAJI and
CA46 demonstrated that the antiproliferation effect of this compound is mediated by the stabilization of MYC promoter G4. This compound showed G4 class-related effects on other promoter quadruplexes, such as VEGF and HIF-1α, which dampens its development as an anti-cancer therapeutic (Brown et al. 2011). To date, however, it is the only compound to have demonstrated a mechanism of action in cells of stabilizing the higher order structure.

All these studies demonstrate that targeting the MYC promoter G4 is a practical and effective strategy for the treatment of MYC-related or driven cancer types. However, the limitation of selectivity of small molecules for an individual G4, like MYC, can be a challenging obstacle to overcome due to the high homology of G4s in the cell. Therefore, a novel G4-targeting approach possessing higher specificity for an individual gene promoter, or even a specific isoform of G4 within the promoter of a gene should be developed.

To address the issue of specificity, some nucleic acid-based approaches have also been explored. An oligonucleotide, Pu27, from the G4-forming region of MYC promoter was directly applied to leukemia cells, wherein it induced profound cellular apoptosis by damaging both telomeric and non-telomeric regions (Islam et al. 2014). Other works using complementary locked nucleic acids targeting the G4-forming region of MYC have downregulated MYC promoter activity by targeting this G4-forming region (Gupta et al. 2013; Kumar et al. 2008).

In this work, we took advantage of the nucleic acid approach and developed a novel NA-clamp-based strategy to target specific G4s. NA-clamps are composed of two pieces of DNA complementing the G4-flanking 5’ and 3’ regions linked by a polyethylene glycol phosphate spacer or sequential thymines to span the three-dimensional space occupied by a particular structure. Our working hypothesis is that an NA-clamp will recognize and stabilize the biological relevant G4 in the promoter region of MYC, thus down-regulating its transcription. Although we
are testing the transcription regulation ability of NA-clamps only on MYC in our proposed research, the success of this aim will indicate the potential of NA-clamps as applied to more promoter G4s.
3.2. The development of Clamp A as a stabilizer for the physiologically relevant MYC G4

3.2.1. The design, optimization and concept demonstration of MYC clamps

Our novel approach began with the design of the MYC G4-targeting NA-clamp. Briefly, two factors were taken into consideration for the design of NA-clamps: (1) when the G4 is formed in the MYC sequence, the distance between the proximal nucleotides on the 5’ and 3’-side of the G4 was shortened from 58 Å to 16.7 Å, (2) when the G4 is formed within the MYC promoter, the flanking nucleotides became single-stranded, which offers the opportunity for complementary binding. We took advantage of both factors, and designed NA-clamps that complementarily bind to the flanking region of the G4, and restrict these regions to the shortened distance of <20 Å, thus stabilizing or facilitating the formation of the G4 (Figure 3-2).

Previous research has identified the predominant physiologically relevant G4 isoform under supercoiled DNA condition to be G4_{1-4} (Sun and Hurley 2009). With computational modeling, we were able to determine the three-dimensional space occupied by G4 to be 16.7 – 30.7 Å. In particular, the distance between the nucleotides (nt) on the 5’ and 3’- side of the G4 is 16.7 Å, to the second most proximal nt on either side is 20.1 Å, and to the third most proximal nt measured 30.7 Å, as compared to 58 Å for the linear sequence of the 16 nt (Figure 3-3A). Three distinct clamps were engineered to match these distances and the remaining flanking regions, as detailed in Table 3-1; the region connecting the two complementary regions consisted of varying lengths of polyethylene linkers. Clamp A1 contained 5 matching nt on the 3’ side of the G-quadruplex and 19 nt on the 5’ side connected by a 17.7 Å linker, clamp A2 used a 21.3 Å linker to connect 4 and 18 nt on the 3’ and 5’ sides, respectively, and clamp A3 used a 31 Å linker to connect 3 and 17 nt, respectively (Figure 3-3B and C). The binding affinity was examined by using EMSA as shown in Figure 3-3D. The MYC G4 was induced from the Pu46 sequence with 25 mM
KCl, all clamps retarded that migration of Pu46 sequence on the polyacrylamide gel in a dose-dependent manner (from 0.01 to 10 µM). Clamps A1 and A2 were able to isolate distinct higher bands, with clamp A1 demonstrating the greatest affinity and most distinct binding to the wild type MYC sequence. Based on this data, all future linkers were maintained at 17.7 Å corresponding to clamp A1 (simplified as clamp A hereafter).

In addition to the physiologically relevant MYC G4\textsubscript{1-4}, two other major isoforms have been reported. The most well-known G4 isoform is the parallel G4\textsubscript{2-5} with loop lengths of 1:2:1 from 5’ to 3’, formed from short linear DNA sequences, which was the first discovered and most extensively used G4 isoform for the screening of MYC G4 stabilizing compounds based on early studies. (Siddiqui-Jain et al. 2002). The other, less studied, isoform (G4\textsubscript{1-5}) is also a parallel structure with loop lengths of 1:6:1 in the 5’ to 3’ direction, formed by guanine runs I, II, IV, and V (Dettler et al. 2011). To enable studies on all G4 isoforms in this region, G to T mutations were made to the Pu46 sequence in different runs of guanines (as shown in table 2-1) to either isolate individual G4s formed in this region or knock out all G4 formation. In addition to clamp A, two other clamps, clamp B and clamp C were designed to recognize the reported G4 isoforms, G4\textsubscript{2-5} and G4\textsubscript{1-5}, respectively, with linker lengths maintained at 17.7 Å and overall flanking regions lengths maintained. The correspondence of the G4 isoforms to their recognizing NA-clamps is described in Figure 3-4B.
Figure 3-2. NA-clamp binding to the MYC wild type Pu46 sequence. During MYC transcription, the negative superhelicity enables the separation of the dsDNA and allows for G4_1,4 (red) formation within the NHE III region. The unpaired single stranded overhang on both 5’ and 3’ ends were complemented by the designed MYC clamp A (as shown in green) to form a G4-Clamp complex as shown on the bottom of the figure. The 17.7 Å polyethylene glycol linker connecting the 5’ and 3’ ends of the clamp restricts the G4 within its computationally modeled space thus stabilizing the structure.
Figure 3-3. Computational modeling of the physiologically relevant MYC promoter G4 and clamp-mediated recognition. (A) 3D modeling of the G4-forming region in the NHE III1 of MYC promoter, distance between the first (16.69 Å, yellow), second (20.05 Å, white) and the third (30.7 Å, blue) proximal bases were calculated. (B) The offset of G4 and iM in supercoiled DNA allows for the binding of (C) designed NA-clamps to the unpaired flanking 5’ and 3’-end of the G4 structure. (D) EMSA analysis shows the binding and recognition pattern of each designed clamp (0.01 to 10 µM) to the MYC sequence. Clamp A1 showed the greatest affinity for WT MYC. (Hao, Gaerig & Brooks, 2016)
Figure 3-4. Designed NA-clamps for multiple G4 isoforms formed in the MYC promoter. (A) Within the NHE III₁ region of the MYC promoter, there are six consecutive guanine runs that are capable of forming multiple G4 isoforms. Previous research have identified the isoform under supercoiled condition (bolded), G₄₁⁻⁴; and two other isoforms formed in linear DNA (Dettler et al. 2011; Siddiqui-Jain et al. 2002; Sun and Hurley 2009). (B) Mutations were made to isolate individual G4 isoforms and clamps recognizing each isolated G4 were designed with constant linker length of 17.7 Å. (Hao, Gaerig & Brooks, 2016)
3.2.2. NA-clamp binding and its specific G4 recognition

The designed clamps A, B, and C were subjected to EMSA and ECD studies to evaluate their ability to recognize and bind to their corresponding target G4s either in the wild type or the mutant sequences. All three clamps, at a concentration gradient of 0.1, 1 and 10 µM, were incubated with either 1 µM of MYC WT or the complete G4 knockout mt 1,2,3,4 sequences at room temperature for 1 hour after the sequences were annealed as described. The DNA species were isolated by Polyacrylamide gel electrophoresis (PAGE) and visualized fluorescently. As shown in Figure 3-5A, as compared to no clamp controls, all clamps produced an upward shift of DNA, indicating some binding between the clamp and the target sequence. A supra shift was produced in a dose-dependent manner only with clamp A, indicating higher order G4 structure complex formation due to the larger size of the G4 DNA complex. No supra-shifted banding was evident with the mt knockout sequence with any clamp; only the moderately retarded bands occurred. This pattern is consistent with the lower of the two shifted bands being that of the clamps binding to their complementary flanking regions and the higher supra-shift band representing binding the flanking regions with a G4 formed between them.

To demonstrate the G4 structure or DNA-clamp complex formation within each sequence, ECD spectrums were recorded to analyze the incubated DNA alone or DNA clamp mixtures as shown in figure 3-5B. With both the MYC WT sequence and the mt 5 sequence isolating the physiological G41-4 isoform, to which only clamp A should bind, signature parallel G4 formation cotton effects with minima at 240 nm and maxima at 265 nm were observed. Upon the incubation with clamp A post-annealing at one equivalence, WT and mt 5 sequences produced, more intense cotton effects at 265 nm indicating G4 formation, as well as a broadening of the cotton effect around 275, which is in accordance with dsDNA produced by the complementary binding of the
flanking region. The addition of clamp A to mt 1,2,3,4 shifted the spectrum from a nondescript signature to a typical dsDNA cotton effect at 275 nm. These ECD spectra indicated that the moderately slower migration bands presented on all mt 1,2,3,4 EMSA and with clamp B and C represent the complementary clamp binding to the flanking region of DNA, but no G4 structure formation. In contrast, the spectrum of WT and mt 5 incubated with clamp A demonstrated that there are G4-dsDNA complex formation besides the dsDNA. As presented in Figure 3-5C, supra-shifted bands were only observed in between clamp A and MYC WT or mt 5, but not mt 1, 6 and mt 3 groups, which demonstrates a specific recognition pattern for clamp A to its designed G4 target G4_{1-4}. In contrast, with mt 1,6 and mt 3, only moderately shifted bands, which also occurred during the incubation of clamps with the complete knockout sequence mt 1,2,3,4, were observed, denoting that only the complementary binding on the flanking region occurred between clamp A and mutant sequences isolating G4_{2-5} and G4_{1-6}. The post-annealing incubation experiments, in which clamps were added after the G4 annealing process, examines whether clamps can specifically bind to and recognize their corresponding G4 isoforms. We also sought to explore the possibility of clamps being able to participate in the ssDNA ⇐ G4 equilibrium and to shift the equilibrium towards G4s. Interestingly, the incubation of clamp A with MYC WT sequence pre-annealing fostered more supra-shifted banding as compared to post-annealing samples as shown in Figure 3-5D. Image J was employed to quantify the band density; band i represents the G4-clamp complex, band ii represents the dsDNA complex, and band iii represents the linear DNA sequences. The incubation with clamp A produced a dose-dependent induction of the supra-shifted band i regardless of pre- or post-annealing. Noticeably at the presence of 1 µM clamp A, the pre-annealing group (lane 3) produced more G4-clamp complex (black bar) as compared to the post-
annealing group (lane 6); annealing MYC WT DNA in the presence of clamp A increased the formation of a G4 approximately 2.5-fold.

Further studies were carried out to elucidate that the DNA complex formed between clamp A and mt 5 DNA isolates G4_{1-4}. Both mt 5 and mt 1,2,3,4 sequences were incubated with biotinylated clamp A, post-annealing, and the samples were subjected to EMSA (Figure 3-6A). The migration pattern of the DNAs post incubation with biotinylated clamp A matched with the previously observed pattern using clamp A, demonstrating the equal efficacy of a biotinylated clamp. For lanes 3,4, 7, and 8, samples post-incubation were subjected to streptavidin pulldown, the recaptured DNA complexes were released from the streptavidin beads and were then subjected to EMSA, together with samples without the pulldown process. The bandings post-pulldown in lanes 4 and 8 were enhanced by adjusting contrast (Figure 3-6B). In lane 4, the only evident banding was the supra shift, compared to lane 3, indicating that the majority of the biotinylated clamp A was involved in the apparent clamp:G4 complex. This is compared to lane 8 where only the moderately shifted band was observed, versus lane 7, which is consistent with complementary DNA formation only on the flanking regions of mt 1,2,3,4 with clamp A. The structural characterization of the mt 5-clamp A complex was evaluated by ECD (Figure 3-6C). Without clamp A, the mt 5 sequence (blue) showed a signature parallel G4 cotton effect at 265 nm indicating the formation of a G4. Upon incubation with biotinylated clamp A, the complete DNA mixture (initial) was separated by pulldown process into the final pulldown products (final) and supernatant (left over DNA post pulldown). Upon the addition of biotinylated clamp A, mt 5 (initial) maintained the signature G4 cotton effect and presented a broadening maxim around 275 nm, indicating the formation of dsDNA. Post-pulldown with streptavidin beads, both the broadening and signature parallel G4 cotton effects were decreased in mt 5 supernatant (green)
compared to mt 5 initial (red), which was expected based on the biotin-streptavidin and clamp-mt 5 interaction. Excitingly, in the final pulldown DNA released from streptavidin beads, the spectrum of mt 5 final (purple) presented both the signature parallel G4 and the dsDNA cotton effects demonstrating that the DNA complex possesses not only the complementary binding on the flanking regions, but also the G4_1-4 structure locked in between. This observation further supports that the supra shift shown in MYC WT and mt 5 sequences post-annealing with clamp A detected a G4-clamp A complex.
Figure 3-5. Specific binding and recognition of clamps to MYC G4s. (A) The binding specificity of clamps A, B, and C to MYC WT and the complete knockout mt 1,2,3,4 were examined by EMSA in a dose-dependent manner. (B) ECD further demonstrated the formation of G4 structure (black line) by both WT and mt 5 but not mt 1,2,3,4 sequence. After incubation with clamp A (blue line), the signature parallel G4 cotton effect was maintained in WT and mt 5 group, while all three groups produced a dsDNA broadening cotton effects around 275 nm denoting complementary binding on the flanking region. (C) Clamp A (designed against the G41,4 isoform isolated by mt 5) mediated a supra-shifted band with WT and mt 5, and a moderately shifted band indicating dsDNA of the flanking region with mt 1,6 (isolating G42,5) and mt 3 (isolating G41,6). (D) Incubation of clamp A with MYC WT, pre-annealing, promotes more of the supra-shifted band (band i) formation compared to post-annealing incubation, most noticeable at 1 equivalence of clamp to DNA (lane 3 vs 6). Densitometry analysis by Image J revealed that there is an ~2.5-fold increase of G4 formation with pre-annealing at 1 µM. (Hao, Gaerig & Brooks, 2016)
Figure 3-6. Clamp-DNA complexes separated by streptavidin-mediated pull down. (A) The incubation of biotinylated clamp A with mt 5 (isolating G4<sub>1-4</sub>) (lanes 1&2) and the complete knockout mt 1,2,3,4 (lanes 5&6) induced the same pattern of DNA migration as in figure 3-5A. Post pull down using streptavidin beads, the recaptured DNA species were redistributed in lanes 3,4,7,8. (B) The clarity of lanes 4 and 8 were enhanced by adjusting contrast using image J software. In lane 4, only the supra shift was evident indicating majority of the clamp A was bound in this DNA complex. In contrast, lane 8 showed only the moderately slower migrating band which is consistent with the flanking region dsDNA band. (C) ECD analysis of the pull-down DNA species with mt 5. Notably, compared to the mt 5 supernatant (green) which is the residue DNA after pulling down, the final pull-down product mt 5 final (purple) showed not only a signature parallel G4 cotton effect at 265 nm, but also a maximum around 275-280 nm indicating dsDNA formation on the flanking region of the G4.
3.2.3. Clamp A G4 affinity and MYC G4₁₋₄ stabilization

A series of competing EMSAs were conducted to examine the recognition and binding preference of clamp A to its designed target, G4₁₋₄. The MYC WT sequence and clamp A ratio was maintained at 1:1, while a panel of non-FAM labeled competing oligonucleotides were added and incubated with the annealed WT G4 post-annealing (Figure 3-7). dsDNA of the G-rich region of the MYC promoter, at up to a 5-fold excess, could not compete for the binding of clamp A. With the MYC mutant sequences mt 1,2,3,4 and mt 1,6, which knockout or isolate other G4 isoforms, respectively, the recognition of clamp A to MYC WT sequence was disrupted only at 5-fold excess of competing oligonucleotides, indicated by the disappearance of supra-shifted band. Only with mt 5, which isolates the target structure G4₁₋₄, the binding between WT and clamp was significantly disrupted at one equivalence of competing oligonucleotide (Figure 3-7A). A narrower range of competing concentrations was then performed with mt 5 to probe the general affinity of clamp A for G4₁₋₄ from a mixture of all the G4 structures in equilibrium within the MYC WT sequence. As shown in Figure 3-7B, the mt 5 formed G4 is able to compete off the binding of clamp A to MYC WT at ratio as low as 0.5, and completely displace the binding at ratio of 1.2. This observation denotes the high binding specificity of clamp A for its target structure G4₁₋₄.

There are three fragments that compose clamp A: a 5-nucleotide portion complementing the 5’-end flanking region, an 18-nucleotide portion that complements the 3’-end flanking region, and a 17.7 Å polyethylene glycol linker. We next probed which fragments are necessary or responsible for G4 recognition and induction. Clamp A was dissected into fragments (Table 3-2) and EMSAs were done to demonstrate the interaction of each partial clamp to the MYC WT sequence (Figure 3-7C). The longer fragment (5’), complementing the 3’-end flank of the G4, either with or without the polyethylene glycol linker, showed more binding as compared to the
shorter fragment (3’), complementing the 5’-end flank of the G4, regardless of linker attachment; this is expected due to strength of complementary bases where the longer fragment shows enhanced binding. When both 5’ and 3’ fragments are added independently, the band migration is slower than either fragment presented alone, but does not facilitate G4 recognition or formation that would be evidenced by a supra-shifted band. Importantly, when we compare clamp A (lane 2) to the presence of all disconnected fragments either as 5’link + 3’ or 3’link + 5’, only clamp A is able to facilitate the supra-shifted band but not the disconnected fragments. This observation made it clear that the complementary binding of the flanking regions, and the spatial restriction are both necessary for the recognition and induction of a G4 structure.

In order to confirm that the recognized and induced G4 formation by clamp A is indeed G4_{1-4}, DMS footprinting study was conducted on species separated by EMSA from MYC WT pre-annealed with clamp A (Figure 3-8A); each band separated by EMSA was excised and extracted to recapture the DNA species before all the DNA was subjected to DMS footprinting (Figure 3-8B). MYC WT without clamp A incubation (lane i) is a mixture of G4 isoforms, denoted by clear protection from guanine runs 2-4, and less distinctive protection from runs 1 and 5. After incubation with clamp A, there is a proportion of the species with no change in migration pattern (lane ii), a band of moderate migration retardation (lane iii), and a supra-shifted band (land iv). Band ii contains linear DNA, as evidenced by clear cleavage after each guanine. Band iii is a mixture of linear DNA and multiple G4 isoforms as evidenced by a broad pattern of partial guanine protection in multiple runs. Most notably, lane iv showed a clear protection pattern from guanine runs 1-4 which is consistent with the target G4 structure (G4_{1-4}).
Figure 3-7. Preferential G4 affinity of clamp A for G41-4. (A) Clamp A binding was competed for by various unlabeled oligonucleotides from MYC WT. dsDNA was not able to compete for any clamp A binding even at 5-fold ratio while mt 1,6 and 1,2,3,4 reduced the binding of clamp A to WT at 5-equivalence. In contrast, competition starts by one equivalence with mt 5 that isolates the target structure of clamp A. (B) A narrower range of competing mt 5 showed as low as 0.5 equivalence of mt 5 disrupts the binding of clamp A to WT sequence, indicating clamp A’s preference for G41-4. (C) A comparison of the dissected fragments of clamp A to the intact species demonstrated that all fragments and physical connection among the three fragments are essential for the G4 recognition and induction. (Hao, Gaerig & Brooks, 2016)
Figure 3-8. DMS footprinting of the DNA species of MYC WT DNA post incubation with clamp A. (A) MYC WT was annealed at the presence of clamp A and 25 mM KCl, the samples were then subjected to EMSA. (B) DNA species were extracted from (A) and subjected to DMS footprinting. (C) Histograms were generated by Image J based on the banding density of (B). Band i represents a mixture of linear and G4s as evidenced by the protection throughout guanine runs. Band ii represents linear DNA reflected by the disappearance of protection patterns. Band iii represents a mixture of linear and non-G41-4 isoforms supported by the partial protection in various runs of guanines. In lane iv, guanine runs 1-4 presented clear protection pattern, in accordance with G41-4, which is the target structure of clamp A.
3.2.4. Clamp A-mediated inhibition of MYC promoter activity and differential cytotoxicity

As clamp A is able to recognize and stabilize the physiologically relevant MYC G4 with high specificity, we investigated if clamp A was able to downregulate MYC promoter activity within the cell, whether clamp A possesses cytotoxicity and towards what cells, and whether clamp A was able to downregulate MYC expression in cancer cells.

To examine if any of the clamps were able to shift the equilibrium of the DNA structures within the NHE III1 region, thus altering promoter activity, we used the MYC promoter-containing Del4 plasmid (He et al. 1998; Siddiqui-Jain et al. 2002), and the promoterless empty vector (EV) plasmid as a control, to probe if MYC promoter activity could be altered by clamps. None of the clamps significantly downregulated promoter activity 24 hr post-transfection. Forty-eight hr post-transfection, only clamp A significantly reduced the expression of firefly luciferase indicating decreased MYC promoter activity from the Del4 plasmid. In contrast, clamp B significantly increased MYC promoter activity, while clamp C showed no effect (Figure 3-9A). A dose-response of clamp A 48 hr post-transfection was performed (Figure 3-9B). MYC promoter activity decreased significantly from 0.5 to 5 equivalents of clamp A. These observations further support the finding that the MYC G41-4 isoform is the predominant physiologically relevant G4 isoform that functions as a silencer in the MYC NHE III1 region; disruption of the formation of this G4 might result in increased MYC promoter activity as evidenced by the increased promoter activity induced by clamp B.

The inherent cytotoxicity of clamp A on both cancerous (MCF-7) and non-cancerous cells (HEK-293) was examined. Previous research has shown that knocking down MYC protein expression through RNAi in MCF-7 cells inhibits tumor cell growth both in vitro and in vivo (Y. H. Wang et al. 2005). It has also been demonstrated that naked DNA oligonucleotides can enter

In this work, varying concentrations of clamp A were directly applied to HEK-293 and MCF-7 cells, and the MTS assay was used to determine the cellular viability 48 hr later. As shown in Figure 3-10A, the viability of cancerous MCF-7 cells was significantly decreased at doses, as low as 30 μM of clamp A, and through 120 μM, which showed an over 80% viability reduction. The calculated IC<sub>50</sub> of clamp A on MCF-7 breast cancer cells is 44.26 ± 1.77 μM. In contrast, the cellular viability of non-cancerous HEK-293 cells were not affected at all by the treatment of clamp A up to 120 μM. This observation indicated that clamp A may possess a selective cytotoxic profile and has the potential to be developed as a cancer therapeutic.

MYC mRNA expression level was examined in MCF-7 breast cancer cells by qPCR (figure 3-10B). MCF-7 cells were transfected with 0, 250, or 500 ng of clamp A one day after plating; cells were lysed 48 hr post-transfection and total mRNA was extracted and qPCR was performed. GAPDH was used as internal control and MYC exon 2 expression level was measured. With 250 and 500 ng of clamp A transfection, we can observe a slight, but not statistically significant, downregulation of MYC expression. Further investigations are required to examine clamp A’s effect on MYC expression.
Figure 3-9. Clamp A significantly decreases MYC promoter activity. (A) One equivalence of clamps A, B or C were transfected together with either MYC promoter-containing Del4 plasmid or promoterless EV plasmid plus the control renilla plasmid into HEK-293 cells. Twenty-four or 48 hr post-transfection, luciferase was measured to determine the promoter activity. At 24 hr, MYC promoter activity was not altered by any of the three clamps. After 48 hr incubation, clamp A significantly downregulated MYC promoter activity. In contrast, clamp B significantly upregulated MYC promoter activity while clamp C showed no effect. (B) A dose response of clamp A was examined 48 hr post-transfection. Starting from as low as 0.5 equivalence of clamp A, an over 50% decrease of MYC promoter activity was observed, which was maintained through 5 equivalences. *p<0.05 as compared to control (0 eq clamp), as determined by one way ANOVA with post-hoc Tukey’s test. (Hao, Gaerig & Brooks, 2016)
Figure 3-10. Clamp A’s cytotoxic profile on cancerous and non-cancerous cells and its ability on MYC expression. (A) Varying doses of clamp A were directly applied to HEK-293 and MCF-7 cell one day post plating, cell viability was measured by MTS assay 48 hr post incubation. Clamp A significantly decreased the cell viability of MCF-7 breast cancer cells starting from 30 µM but not non-cancerous HEK-293 cells. The IC$_{50}$ was reported to be 44.6 µM in MCF-7 cells as calculated by GraphPad Prism software (*p<0.05, as compared to untreated control, determined by one-way ANOVA). (B) MCF-7 cells were transfected with either 250 or 500 ng of clamp A, and MYC mRNA expression level was measured through real time qPCR 48 hr post-transfection. MYC mRNA level was not significantly changed (*p<0.05, as compared to untreated control, determined by one-way ANOVA).
3.3. Development of Clamp A 5T as an alternative for clamp A

So far, we have demonstrated that clamp A is able to recognize and stabilize the physiologically relevant MYC G4_{1-4} isoform and downregulate MYC promoter activity in a luciferase reporter gene assay. Further investigations to determine the intracellular localization, and the binding target of clamp A to chromosomal DNA, require modifications to the clamp. However, the current clamp A with the polyethylene glycol linker is costly and the modification potential of this clamp A is restricted. In order to reduce the cost of the clamp and to reach more flexible oligonucleotides modification potential, we replaced the polyethylene glycol phosphate linker (clamp A) with five thymines (clamp A 5T), which maintain the linker length of ~18 Å. Clamp A 5T can be fluorescently tagged with FAM, or internally modified with biotin, which enable the possibility for microscopy and immunohistochemistry studies, respectively. All the oligonucleotides and clamps used in this section of the study are listed in table 3-3-1. To further this research with clamp A 5T, its G4 recognizing ability had to be compared to clamp A.

MYC WT, or the knockout mt 1,2,3,4, were annealed in the presence of clamp A, 5’ link, clamp A 5T, 5T, or 5’ 5T. With the mt 1,2,3,4 sequence, all full or partial clamps induced one slower migration band, indicating complementary binding of the flanking region. All inclusions of 5 thymines, as compared to the polyethylene glycol linker, resulted in slower band migration due to the increased size and weight of the thymines. With the MYC WT sequence, only the intact clamp A or clamp A 5T fostered a supra-shifted band, indicating the formation of G4-clamp complex (Figure 3-3-1A).

As with the original clamp A, clamp A 5T was dissected into its fragments and the migration patterns of FAM labeled MYC WT were examined by EMSA, including the linker 5
thymines alone. Clamp A 5T fragmented parts included the 5’ complementing sequence plus 5 thymines (5’ 5T), the 3’ complementing sequence plus 5 thymines (3’ 5T) and five thymines (5T) alone; each fragment was incubated with MYC WT (Figure 3-3-1B). The longer 5’ 5T complemented better than the shorter 3’ 5T at equal molarity, as expected. Any fragments containing the 5’-end produced moderately slower migrating bands compared to the control. The 5T alone did not induced any band migration at all. In contrast, only the complete clamp constructs, either clamp A and clamp A 5T introduced supra-shifted bands indicative of G4-Clamp complex formation. This data is consistent with previous findings with clamp A, which further demonstrates that the spatial restriction of G4 structure is essential for the recognition and stabilization of MYC G41-4.

The binding specificity of clamp A 5T was examined by incubating FAM-labeled clamp A 5T with non-fluorescently labeled MYC sequences, including the WT and all mutants, and other random non-MYC G4-forming sequences (Figure 3-3-2A). All MYC sequences incubated with clamp A 5T, including WT; mt 5; mt 2,3; mt 1,6; and even the complete knockout sequence mt 1,2,3,4, fostered moderately slower migrating bands as compared to either labeled clamp A 5T alone or any of the other G4-forming sequences (VEGF, KRAS, Bcl-2 and NRAS). Moreover, none of the other G4-forming sequences induced any changes to the migration of clamp A, confirming the binding specificity of clamp A 5T to the flanking region of the MYC promoter G4. Only with MYC WT and mt 5, both of which are able to form the MYC G41-4 isoform, were supra-shifted bands induced by clamp A 5T. This demonstrates the specific recognition of clamp A 5T to the target gene and to the physiologically relevant MYC G41-4 isoform.

The structure of the binding complex formed by clamp A 5T and MYC WT was examined by ECD (Figure 3-3-2B). MYC WT alone (orange) showed a signature 240 nm minimum, and a
265 nm maximum, cotton effect, indicating the formation of a parallel G4 structure. Upon the addition of 25 mM KCl, MYC WT (cyan) formed more parallel G4 structure, reflected by a more intense 265 nm cotton effect. After incubation with clamp A 5T, more G4 formation was noted, as compared to either WT alone or WT plus 25 mM KCl, as well as a broadening of the cotton effect from 270-280 nm, indicating the formation of dsDNA between clamp A 5T and MYC WT on the flanking regions of the G41-4.

To confirm the structure of the fostered DNA complexes, all DNA species introduced by the binding of clamp A 5T on MYC WT separated by EMSA were DMS footprinted (Figure 3-3-3). 6-FAM labeled MYC WT sequence was annealed at the presence of clamp A 5T at equal molarity, samples were subjected to chemical (DMS) footprinting and loaded on a polyacrylamide gel for species separation. Four DNA bands (lanes i – iv) (Figure 3-3-3A) were excised and extracted; DNA complexes were then recaptured and cleaved by piperidine and the cleaved samples were separated on a DNA sequencing gel (Figure 3-3-3B). Histograms of all banding patterns from lanes i-iv were generated using Image J software and aligned by Photoshop software (Figure 3-3-3C). MYC WT without clamp A 5T incubation in lane i (black) is a mixture of G4 isoforms, denoted by partial protection throughout the guanine runs. In band ii (blue), linear DNA is evident by the loss of protection pattern throughout the guanines runs. Band iii (green) is mixed by linear DNA and other non-G41-4 isoforms evidenced by partial guanine protection in multiple runs. Lane iv (red), in a manner consistent with the original clamp A, showed a clear protection pattern specifically from guanine runs 1-4, which is consistent with the target G4 structure (G41-4) of clamp A 5T.

Based on the results above, clamp A 5T with five thymines as the interconnecting linker showed the same G41-4 recognizing and stabilizing ability as clamp A and is suitable for further
study. These findings highlighted again that the spatial restriction is the key for the recognition and stabilization of the physiologically relevant MYC G4_{1-4}.
Figure 3-3-1. DNA and G4 recognition comparison between clamp A, clamp A 5T and clamp fragments. (A) MYC WT and mt 1,2,3,4 sequences were incubated with clamp A, clamp A 5T and their 5’ fragments at equal concentrations (1 µM); Banding pattern was examined by EMSA. Only the intact clamp constructs (clamp A and clamp A 5T) fostered the supra-shifted bands, indicating G4-clamp complex. (B) fragments of clamp A 5T were compared to the intact clamp constructs for post-annealing binding to the MYC WT sequence by EMSA. The supra-shifted band was only evident in the presence of the intact clamps, including both clamp A and clamp A 5T.
Figure 3-3-2. Specific recognition of clamp A 5T to the physiological MYC G4.
4. (A) FAM-clamp A 5T was incubated with MYC sequences or other DNA sequences, pre-annealing and was then examined by EMSA. Only WT MYC and mt 5 that are able to form G4 isoform but not any MYC mutant or other sequences produced supra-shifted bands indicating G4-clamp complex formation indicated by (B) The signature cotton effect at 260 nm plus the broadening dsDNA maximum around 270-280 nm post incubation with clamp A 5T as examined by ECD spectrum, as compared to WT alone or WT+25 mM KCl, which only showed the cotton effects indicative of G4 formation.
Figure 3-3-3. **Structure elucidation of the DNA-clamp complexes by DMS footprinting.** (A) MYC WT was annealed at the presence of clamp A 5T, and the samples were DMS footprinted before subjected to EMSA. (B) Each band on the EMSA was extracted and cleaved at methylated guanine sites, samples were then visualized on a DNA sequencing gel, and band indicators (i to iv) and colors are consistently used throughout the figure. (C) Histograms were generated by Image J based on the density of the bandings. Band i showed partial protection in all the guanine runs indicating its nature of multiple G4s in equilibrium. Band ii showed loss of protection pattern throughout the guanine runs, especially runs 1-4, which indicated its nature as linear DNA. Band iii showed weak partial protection on some runs indicating its nature as a mixture of linear and non-G4 isoforms. Notably, band iv, which corresponds to the supra-shifted band on the EMSA, showed significant protection in guanine runs 1-4, suggesting its nature as G41-4, which is consistent with the target G4 of clamp A 5T.
3.4. Cellular uptake and delivery of clamp A 5T

To study the cellular uptake and the intracellular localization of clamp A 5T, 6-FAM labeling was employed. FAM-clamp A 5T was introduced into MCF-7 breast cancer, RAJI, and CA46 lymphoma cells either directly or via transfection with Fugene HD, for up to 48 hr. After the nuclei of cells were stained with Hoescht 33342, images were collected in a Cytation 5 imaging multi-mode reader to monitor the migration of the clamp (Figures 3-4-1 to 3-4-3). For all cells at all times points, intracellular oligonucleotide was only evident post-transfection. Six hr post treatment, no FAM-clamp A 5T uptake was evident in any of the MCF-7 cells; minor uptake was noted in both the RAJI and CA46 cells. Twenty-four hr post treatment, marked FAM fluorescence signal was observed in the cytoplasm and nuclei of MCF-7, RAJI, and CA46 cells. Forty-eight hr post treatment, no FAM-clamp A 5T was evident in any cell line.
Figure 3-4-1. Time course cellular uptake of clamp A 5T in live MCF-7 cells. FAM labeled clamp A 5T was introduced into MCF-7 cells either directly or transfected with Fugene HD. Before all images were obtained, cells were incubated with Hoeschst 33342 nuclear dye (blue). Six hr post treatment, neither naked nor Fugene transfected clamp were notably accumulated in the cells. Twenty-four hr post-transfection, marked FAM fluorescence was observed in the transfected group, both in the cytoplasm and in the nucleus. Forty-eight hr post-treatment, FAM fluorescence was no longer evident.
Figure 3-4-2. Time course cellular uptake of clamp A 5T in live RAJI lymphoma cells. FAM labeled clamp A 5T was added to RAJI cells either directly or transfected with Fugene HD. Before all images were obtained, cells were incubated with Hoeschst 33342 nuclear dye (blue). Six hr post treatment, no fluorescent signal was observed in the non-transfected group. In contrast, green FAM fluorescence was detected on cellular membranes with Fugene transfected group. Twenty-four hr post-transfection, marked FAM fluorescence was observed in the transfected group, both in the cytoplasm and in the nucleus. Forty-eight hr post treatment, FAM fluorescence was no longer evident.
Figure 3-4-3. Time course cellular uptake of clamp A 5T in live CA46 lymphoma cells. FAM labeled clamp A 5T was added to CA46 cells either directly or transfected with Fugene HD. Before all images were obtained, cells were incubated with Hoeschst 33342 nuclear dye (blue). Six hr post treatment, no fluorescent signal was observed in the non-transfected group. In contrast, green FAM fluorescence was detected on cellular membranes with Fugene transfected group. Twenty-four post-transfection, marked FAM fluorescence was observed in the transfected group, both in the cytoplasm and in the nucleus. Forty-eight hr post treatment, FAM fluorescence was no longer evident.
Based on the observations above, it is clear that a delivery system that can transport clamps into cells is essential. However, current viral and non-viral mediated transduction or transfection systems have their limitations and sides effects for in vivo application. Therefore, developing and testing a system that is able to effectively introduce clamps into cancer cells, and particularly into difficult to transfect lymphoma cells, is a key step for developing NA-clamps into therapeutics. To fulfill this need, we collaborated with Dr. Adam Smith, an Assistant Professor in the Department of Chemical Engineering at the University of Mississippi, to test polymer delivery systems.

Polymers have attracted significant interest in the field of NA delivery. The efficacy of the polymers depends on their ability of transport NA across the cellular membrane and to escape from the endosomes and releasing the NA into the cytoplasm. Cationic polyamine-containing polymers have exhibited a great buffering capability around endosomal pH (5.5-6), which causes osmotic swelling and the final rupture of the endosomal compartment (Du, Wang et al. 2010; Aied, Greiser et al. 2013; Jones, Chen et al. 2013). However, the polyplexes formed from unmodified cationic polymers tend to be unstable when anionic serum proteins are present thus leading to premature NA unpackaging, degradation, aggregation, and clearance in vitro and in vivo. The incorporation of a neutral, hydrophilic segment, such as polyethylene glycol (PEG), into cationic polymers has been utilized to address the colloidal stability issue. In addition, the inclusion of a hydrophilic stabilizing segment increases the stability and blood circulation time of the polyplexes (Zhang, Satterlee et al. 2012). PEG-based polymers have been widely used to deliver NA into cells, but there are still obstacles to overcome. Previous research found that PEG can nonspecifically interact with blood components and induce antibody production, thus promoting the clearance of nanomedicine and limiting the utility of PEG-based polymeric delivery systems clinically (Knop, Hoogenboom et al. 2010). Other non-ionic, hydrophilic polymers, including poly(amino acid)s,
poly(glycerol), poly(2-oxazoline)s, and vinyl polymers have been proposed to substitute PEG to overcome this issue (Knop, Hoogenboom et al. 2010). It is still important to discover new polymer vehicles that possess efficacy without these drawbacks.

Synthetic glycopolymers have gained interest for their improved colloidal stability and possible cellular surface binding interactions (Alexandra, Orietta et al. 2013; Wang, Hong et al. 2013). Previous work by Dr. Smith has demonstrated that glycopolymplexes can efficiently deliver DNA to HeLa cells and siRNA to U-87 glioblastoma cells with low cytotoxicity (Smith, Sizovs et al. 2011). More recent work on copolymers comprised of different sugar moieties demonstrated selective cellular uptake of NAs in cells overexpressing the GLUT-5 transporter, suggested the potential utilization of glycopolymers as targeted NA delivery system for cancer cells (von der Ehe, Rinkenauer et al. 2016).

Dr. Smith’s laboratory synthesized nine block copolymers (Figure 3-4-4), each comprised of a hydrophilic block of oligo (ethylene glycol) methyl ether methacrylate (OEGMA) and a varying number of tertiary amine blocks, for the delivery of NAs. These block copolymers were able to effectively deliver anti-MYC siRNA into MCF-7 cells and to decrease MYC expression (data not shown; submitted to Macromolecular Biosciences). All nine block copolymers were evaluated for their ability to delivery clamp A 5T into RAJI and CA46 lymphoma cells. FAM labeled clamp A 5T was incubated with each block copolymers at a ratio of 3:1 for 30 minutes before being applied to plated RAJI and CA46 lymphoma cells. Six - 48 hr post-transfection, cells were stained with Hoechst 33342 and imaged with a BioRad Cytation 5 multi-mode imaging reader (Figures 3-4-5 and 3-4-6). With both RAJI and CA46 cells, no FAM fluorescence was detected 6 hr post-transfection with any of the nine polyplexes; 24 hr post-incubation, the DNA:polyplex complexes formed with the ethyl copolymer series were able to qualitatively
transfect clamp A 5T into cells, as evidenced by the green FAM fluorescence in the cytoplasm and the cyan colored co-localization with blue Hoechst 33342 nuclear stain. As with the FuGene HD transfections, 48 hr post-complex application, no FAM fluorescence signals were detected, under any condition, in either cell line.
Figure 3-4-4. Block copolymers evaluated for the delivery of NA-clamps into lymphoma cells. (A) Hydrophilic OEGMA and cationic methacrylate moieties with terminal methyl, ethyl, or isopropyl substituents linked in blocks of 25, 50, or 75 were synthesized into block copolymers and (B) characterized by gel permeation chromatography (McClellan et al., 2016).
Figure 3-4-5. Delivery of clamp A 5T into RAJI lymphoma cells by block copolymers. 6-FAM labeled clamp A 5T was transfected by nine copolymers into RAJI lymphoma cells. Twenty-four post-transfection, the nuclei were stained with Hoechst 33342 and cells were imaged. No FAM fluorescence was detected at any time point in cells transfected by methyl and propyl polymer series. With ethyl polymers, no FAM fluorescence was detected at 6 or 48 hr time points (data not shown). Only at 24 hr, and with the ethyl series, was FAM fluorescence clear.
Figure 3-4-6. Delivery of clamp A 5T into CA46 lymphoma cells by block copolymers. 6-FAM labeled clamp A 5T was transfected by nine copolymers into CA46 lymphoma cells. Twenty-four hr post-transfection, the nuclei were stained with Hoechst 33342 and cells were imaged. No FAM fluorescence was detected at any time point in cells transfected by methyl and propyl polymer series. With ethyl polymers, no FAM fluorescence was detected at 6 or 48 hr time points (data not shown). Only at 24 hr, and with the ethyl series, was FAM fluorescence clear.
The MTS assay was used to evaluate the inherent cytotoxicity of clamp A 5T on HEK-293 and MCF-7 cell lines (Figure 3-4-7) at 24-144 hr. No cellular viability changes were observed after 24 and 48 hr incubation with clamp A 5T in either cell line. Seventy-two hr after treatment with clamp A 5T, differential cytotoxicity was evident at 70 µM, with 33% vs 15% decrease in cellular viability in the MCF-7 and HEK-293 cells, respectively. 144 hr post-clamp A 5T incubation, similar patterns of decreased cell viability were evident in both cell lines, with IC_{50s} of 52.13 ± 1.8 and 68.1 ± 3.0 µM in the MCF-7 and HEK-293 cells, respectively.

As seen above, the ethyl series copolymers were able to effectively transfect clamp A 5T into both RAJI and CA46 lymphoma cells after a 24 hr incubation. Here, we evaluated the cytotoxic profile of the ethyl series block copolymers and the polyplexes formed by clamp A 5T or scramble clamp with the copolymers in both cell lines (Figure 3-4-8). All three copolymers alone, E25 (ocean blue), E50 (orange), and E75 (gray), demonstrated a dose-dependent cytotoxic profile in both RAJI and CA46 lymphoma cell lines with up to 38% decreased cellular viability at 72 hr. Among the three copolymers, E75 (grey) showed the highest inhibition of cell growth (35% with RAJI, 38% with CA46) in both cell lines, followed by E50 (orange, 24% with RAJI, 36% with CA46) and E25 (ocean blue, 0% with RAJI, 22% with CA46). Compared to copolymers alone, the polyplexes formed by scramble clamp reduced the cytotoxic profile in RAJI and CA46 cells at 72 hr. Compared to the polyplexes formed by scramble clamps, polyplexes formed by clamp A 5T demonstrated more cytotoxicity on both RAJI (5% with E50, 23% with E75) and CA46 (20% with E50, 32% with E75) cell lines, as evidenced by the lower cellular viability of E50/75 plus clamp A 5T (blue/green) vs E50/75 plus scramble clamp (brown/dark gray) at 72 hr time point, indicating that clamp A 5T likely contributes to the inhibition of cellular growth of both RAJI and CA46 lymphoma cells.
Figure 3-4-7. Cytotoxicity of clamp A 5T on HEK-293 and MCF-7 cells. MTS assay was used to measure the cytotoxicity of clamp A 5T on HEK-293 and MCF-7 cells. Twenty-four or 48 hr post treatment, no significant dose-dependent cytotoxicity was clear in either cell line. By 72 hr, up to a 20 or 40% decrease was noted in HEK-293 and MCF-7 cells, respectively. By 144 hours of treatment with clamp A 5T, cells demonstrated a dose-dependent decrease in viability; no difference were notable between the two cell lines.
Figure 3-4-8. Lymphoma cells cytotoxicity of clamp A 5T transfected by ethyl series copolymers. Clamp A 5T or scramble clamp were transfected with the ethyl series block copolymers into RAJI and CA46 lymphoma cells and the cellular viability was determined by MTS assay at various time points. E75 (green) and E50 (orange) copolymers alone demonstrated dose and time dependent cytotoxicity on both cell lines. When complexed with NAs, either clamp A 5T or scramble clamp, the cytotoxicity of all three copolymers were decreased as compared to polymers alone. Polyplexes formed by clamp A 5T, showed a more cytotoxic profile compared to polyplexes formed by scramble clamp, which is most evident at 48 or 72 hr with E75 copolymers (green vs dark gray).
3.5. Intracellular activity of clamp A 5T

The binding of clamp A 5T to its target region of the MYC promoter and the impact of clamp A binding on MYC transcription was evaluated in MCF-7 breast cancer cells. Estrogen-receptor positive MCF-7 cells transcriptionally upregulate MYC within 1 hr of stimulation by 17β-estradiol (E2, 2 nM) (C. Wang et al. 2011). We confirmed this effect in our cells, where MYC mRNA was elevated over 6-fold at one hour post-E2 stimulation in MCF-7 cells (Figure 3-5-1A). In order to probe the binding of clamp A 5T to the promoter of MYC, MCF-7 cells were transfected with biotinylated clamp A 5T one day post plating. Twenty-three hr later, as indicated, cells were treated with 2 nM of E2 for 1 hr. Cells were then washed and harvested for genomic DNA extraction. The extracted DNA was then sheared enzymatically and streptavidin beads were used to pull down the DNA bound to biotinylated clamp A 5T. The pulled-down DNA was then released from the beads and amplified by PCR reaction with MYC specific PCR primers. The final DNA products were visualized on an agarose gel (Figure 3-5-1B). As expected, the MYC promoter could be PCR amplified from extracted genomic DNA (positive control). In the absence of biotinylated clamp A 5T, and after isolation by streptavidin, no MYC promoter was able to be amplified (negative control). After treatment with biotinylated clamp A 5T (10 µg and 20 µg), the final pulldown products were enriched for MYC promoter DNA, as evidenced by the clear banding at 287 bp. After E2 stimulation, which activates the transcription activity of the MYC gene, thus allowing for more G4 formation in the NHE III region, more binding of clamp A 5T to MYC promoter was observed as evidenced by the darker bands at both 10 µg and 20 µg. These observations demonstrate that clamp A 5T is able to enter the nucleus and bind to the MYC promoter region in live cells, and that as more active transcription of MYC generates more negative
superhelicity, the double stranded DNA is separated and more G4 is formed, clamp A 5T binding to the promoter region was markedly promoted.

Next, we examined the impact of clamp A 5T on the transcription of MYC in MCF-7 breast cancer cells without and with E2 stimulation (Figure 3-5-2A). Clamp A 5T or scramble clamp were transfected with FuGene into MCF-7 cells, that were subsequently stimulated by E2. Twenty-four hr post-transfection, cells were harvested and mRNA was extracted. MYC expression in MCF-7 cells was not significantly changed either with or without E2 treatment. The effect of clamp A 5T on MYC regulation in RAJI lymphoma cells was also determined by qPCR (Figure 3-5-2B). Compared to control, clamp A 5T either alone or transfected by E25 and E50 decreased the expression of MYC by up to 40%, though no significance was found (n=2). When complexed with E75, the expression of MYC was not significantly changed, as compared to control; further experiments with more replicates and more time points are required to determine the effect of clamp A 5T on MYC expression in RAJI cells.
Figure 3-5-1. Clamp A 5T binding to the promoter region of MYC. (A) MYC mRNA expression post-E2 stimulation in MCF-7 cells was measured using qPCR. MYC transcription was activated by 2 nM E2 treatment as evidenced by the 6-fold increase of MYC mRNA expression compared to the control with no E2 treatment (p<0.05, as determined by student T-test). (B) Clamp A 5T binding to the MYC promoter was confirmed by a ChIP-PCR assay. Biotinylated clamp A 5T was transfected into MCF-7 cells one day post plating. Twenty-three hr post incubation, 2 nM E2 was added as indicated, and the extracted and sheared DNA was amplified using MYC primers; amplified DNAs were visualized on an agarose gel. Input DNA was diluted 20 times before loading. The input positive control but not the negative streptavidin beads control, regardless of E2 treatment, showed clear bands of amplified MYC promoter as evidenced by the 287 bp bands. After pulldown, from cells treated with either 10 µg or 20 µg of biotinylated clamp A 5T, amplified MYC promoter was detected. Notably, after E2 stimulation, more binding of clamp A 5T to the MYC promoter was observed, as evidenced by more promoter amplified, indicating active transcription facilitates the binding of clamp A 5T to the MYC promoter.
Figure 3-5-2. The effect of clamp A 5T on MYC expression in MCF-7 and RAJI cells. (A) Clamp A 5T and scramble clamp were transfected into MCF-7 cells and their impact on MYC transcription was determined by qPCR. For both E2 treated and non-treated groups, clamp A 5T did not significantly affect the expression of MYC. (B) The transcriptional impact of clamp A 5T, polyplexes formed by clamp A 5T and Fugene or the E series polymers on MYC transcription in RAJI lymphoma cells were evaluated by qPCR. Clamp A 5T alone, with Fugene, or with E50 decreased the expression of MYC while clamp with E75 affected no change. However, no significant difference was identified.
3.6. Clamp A 5T as a diagnostic companion

As demonstrated above, clamp A 5T is able to recognize and stabilize the physiologically relevant MYC G4_{1-4} isoform in vitro. In addition, ChIP-PCR assay demonstrated that clamp A 5T can bind to the MYC promoter in live cells. These data suggest that clamp A 5T could possibly be employed as a diagnostic tool to detect MYC G4 formation in cells in order to inform the use of small molecules that recognize and stabilize this structure. In this section, we will explore the potential of using clamp A 5T to detect induced MYC G4 formation in fixed and permeabilized MCF-7 breast cancer, RAJI and CA46 lymphoma cells.

Previous research, through fluorescence-based screening identified an ellipticine derivative, NSC 338258, as a MYC G4 stabilizer with nanomolar affinity. This compound has been demonstrated to be moderately MYC G4 selective and most notably, its mechanism of action has been validated to be MYC promoter G4-mediated (Brown et al. 2011). Here, we utilized this compound to induce MYC G4 formation, and then detected the G4 formation using fluorescent labeled clamp A 5T in fixed and permeabilized MCF-7, RAJI, and CA46 cells. The IC_{50} of NSC 338258 on MCF-7 cells was determined to be 25 µM at 24 hr (Figure 3-6-1A). For the G4 induction assay, MCF-7 cells were treated with varying concentrations of NSC 338258 (12.5, 25, and 37.5 µM) for 6 hr. Cells were then fixed and permeabilized before incubation with either FAM clamp A 5T or FAM control ODN, together with Hoechst 33342 (nuclear stain), for 1 hr and imaged on a Cytation 5 multi-mode imaging reader (Figure 3-6-1B). With no G4 induction by NSC 338258, no visible binding of clamp A 5T was observed, as compared to control. Starting from 12.5 µM (0.5X IC_{50}) of NSC 338258 treatment, a dose-dependent colocalization of FAM and Hoechst 33342 nuclear dye was observed, which was maximally noted at 25 µM (1X IC_{50}). With 37.5 µM (1.5X IC_{50}) of NSC 338258, a loss of detectable FAM fluorescence, indicating the loss of clamp
binding to the DNA was observed, this was likely due to chromatin condensation induced by the apoptosis of MCF-7 cells. As compared to clamp A 5T probed cells, FAM labeled control ODN probed cells showed only sporadic fluorescence in the 12.5 μM (1X IC₅₀) NSC 338258 treated group, but no dose-dependency.

The same G4 induction and detection assay was also conducted on RAJI and CA46 lymphoma cells (Figure 3-6-2). As it relates to the MYC promoter, RAJI and CA46 Burkitt’s lymphoma cells differ from each other by the location of the translocation between the MYC gene and the immunoglobulin heavy chain gene. Translocated RAJI cells retained the entire MYC gene including the promoter G4, and exons 1, 2, and 3. In contrast, translocated chromosomes in CA46 cells have lost the endogenous G4-forming region and exon 1, only retaining exons 2 and 3 (which are the minimal requirements for a functional MYC protein), with expression driven by the Ig gene. Therefore, the heightened MYC transcriptional activity reflected by increased MYC promoter G4 formation should be most evident with RAJI, versus CA46 lymphoma cells. The 24 hr IC₅₀s of NSC 338258 were determined for both cell lines previously to be 2.6 and 13 μM, respectively (Brown, et al, 2011). For both cell lines, incubation with FAM labeled clamp A 5T, but not the control ODN, coincided with NSC 338258 dose-dependent G4 induction, as evidenced by the cyan colored colocalization of FAM fluorescence and Hoechst 33342 nuclear dye. When we compare (Figure 3-6-2A) RAJI to (Figure 3-6-2B) CA46 cells incubated with FAM-clamp A 5T, we can observe binding of clamp A 5T in the nucleus of RAJI, but not CA46, cells at low dose (0.5X IC₅₀) of NSC 338258 treatment. With RAJI cells incubated with high dose (1.5X IC₅₀) of NSC 338258, we could see marked colocalization of FAM fluorescence with Hoechst dye indicating the binding of clamp A 5T to the promoter G4 of MYC. In contrast, with CA46 cells, even incubation with high dose (1.5X IC₅₀) of NSC 338258 only produced small amount of G4 detection by clamp A
5T. These data further support the hypothesis that the binding of clamp A 5T to the nuclear DNA in these fixed and permeabilized cells is through a promoter G4-mediated mechanism. These data also suggest that clamp A 5T has potential to be used as a diagnostic companion to detect the hyper-transcriptional activity, and G4 formation in certain cancer cells.
Figure 3-6-1. Detecting the physiologically relevant G4 formation induced by selective MYC G4 stabilizing compound NSC338258. (A) The IC$_{50}$ of NSC 338258 on MCF-7 cells was determined to be 25 µM by MTS assay. (B) MCF-7 cells were treated with varying concentrations of NSC 338258 6 hr before fixation and permeabilization. Fixed cells were incubated with either FAM clamp A 5T or control ODN with Hoechst 33342 nuclear dye for 1 hr before being subjected to imaging. Compared to control, cells incubated with clamp A 5T presented a dose-dependent binding to the DNA, as evidenced by the cyan color colocalization maximized at 25 µM NSC 338258, compared to control. With 37.5 µM NSC 338258, a loss of fluorescence signal was observed, possibly due to chromatin condensation induced by apoptosis. No dose-dependent binding was observed with FAM control ODN (negative control), further demonstrating that the dose-dependent binding is MYC G4-
Figure 3-6-2. **G4 induction assay on RAJI and CA46 cells.** (A) RAJI and (B) CA46 lymphoma cells were treated with NSC 338258 for 6 hr before fixation and permeabilization. Cells were then incubated with either FAM-clamp A 5T or control ODN, together with Hoechst 33342 nucleus blue dye for 1 hr before images were collected. With both cell lines, only incubation with clamp A 5T demonstrated dose-dependent effect of colocalization with NSC 338258 treatment.
3.7. Discussion

In this work, we described the development, optimization, delivery, in cell activity and potential diagnostic value of a novel nucleic acid clamp-based approach that targets the physiologically relevant G4$_{1-4}$ isoform in the NHE III$_1$ region of the MYC promoter. The central tenant of the clamp approach utilizes the spatial restriction generated by the complementing DNA flanking the G4 structure to constrain it within the locked distance in between the 5’- and 3’-flanking regions. Through computational modeling and a series of affinity assays, the optimal linker length was determined to be ~18 Å, and the lead clamp A designed to detect the G4$_{1-4}$ isoform showed its specific recognition. Interestingly, we observed that this clamp A can specifically recognize G4$_{1-4}$ and also can shift the equilibrium of the possible DNA structures within this region to a G4$_{1-4}$ favorable state, as evidenced by DMS-footprinting.

Within the NHE III$_1$ region of the MYC promoter, three major G4 isoforms have been described by previous research. The most well studied G4 isoform is from the linear DNA oligonucleotide Pu27, which contains the first five guanine runs of the NHE III$_1$ region. With this linear DNA sequence, research has shown that the G4 formed in linear DNA within this region was G4$_{2-5}$, with loops 1, 2, and 1 nucleotides in the 5’-3’ direction, involving the participation of guanine runs II, III, IV, and IV in a parallel conformation (Siddiqui-Jain et al., 2002). This G4 isoform has been widely used for the screening of G4-interacting compounds, but so far, no compound screened specifically against this isoform selectively interacts with MYC promoter. Most notably, a DMS footprinting study using Del4 plasmid that contains the MYC G4-forming region under supercoiled conditions revealed that the predominant G4 isoform formed under physiologically relevant torsional stress is G4$_{1-4}$, another parallel G4 conformation with the same loops of 1, 2, and 1 nucleotides in the 5’ to 3’ direction, but involving the participation of guanine
runs I, II, III, and IV (Sun & Hurley, 2009). Supercoiling has been identified to be the major driving force, at least in the MYC promoter, that opens double stranded DNA and allows for non-canonical secondary DNA structures including G4s and iMs to form (Brooks and Hurley 2009). In our study, a reporter gene assay using the same Del4 plasmid containing the MYC promoter, we demonstrated that only clamp A 5T, which specifically recognizes and stabilizes the G41-4 isoform, is able to downregulate MYC promoter activity. In contrast, clamp B, which is designed against the predominant G42-5 isoform found in linear DNA, induced a significance increase in MYC promoter activity. This could be due to clamp B disrupting the dynamics of NHEIII1 region, decreasing the formation of the physiologically relevant MYC G41-4 isoform, thus impeding the G4 regulation on MYC transcription. These observations extended further support that G41-4 is the physiologically relevant G4 isoform that regulates MYC transcription. Further research efforts on screening MYC G4 targeting compounds should be focusing on against this physiologically relevant G41-4 isoform or the entire NHE III1 region, rather than G42-5 that forms in the linear DNA with Pu27.

Within MYC promoter, the G4 and iM structures form offset from one another, which leaves a 14 bp long ssDNA overhang on the 5’ side of the G4. Also, computational modeling studies suggested that there are at least 5 bases serving as buffer regions for the transition from secondary DNA structures to dsDNA. This offset plus the buffering regions provided us with the basis for the designed NA-clamps. This clamp approach takes advantage of this over 20 bp long complementing region, which offers great selectivity for the MYC promoter. In contrast, with current small molecule G4 targeting strategies, high selectivity for one single G4 is hard to achieve due to the general homologous topology of these G4 structures. For example, G41-4 and G42-5 isoforms are both parallel G4 structures with 1:2:1 loops from 5’ to 3’ direction, the only variance
that differ these two structures is the loop components which are T: GA: T for G4₁₋₄ and A: TG: A for G4₂₋₅. Other promoter G4 structures, such as the ones from HIF-1α and VEGF, also share high structural topology with the MYC promoter G4. This topological homology introduces obstacles to drug screening processes that lead to non-selective G4-interacting small molecules.

In our qPCR studies, we did not observe significant downregulation of MYC transcription by clamp A in MCF-7 breast cancer cells 48 hr post-treatment. It is possible that clamp A is degraded or expelled from the cell, as supported by our study of cellular localization. It is well known that non-modified DNA suffers from stability issues in cells. It can be readily digested by nucleases and other intracellular enzymes when transported into cells (Watts and Corey 2012). MYC mRNA has a rapid turnover rate in the cytoplasm at approximately half an hour in both human and murine cells (Jones and Cole 1987). The rapid MYC mRNA turnover rate of MYC in MCF-7 breast cancer cells might also contribute to the rapid replenishment of the MYC mRNA pool in clamp A-treated cells. Other possible reasons for this negative data include the choice of a single concentration of the clamp A and the possible accessibility issues of clamp A into the nucleus. Further investigation is required to confirm and explain the lack of a significant change in MYC gene expression.

With the selectivity drawbacks of small molecules, other nucleic acid based approaches have been used to target the MYC G4-forming region. A 27-base pair oligonucleotide, Pu27, from the G4-forming region of MYC promoter was directly applied to leukemia cells, wherein it induced profound cellular apoptosis by damaging both telomeric and non-telomeric regions (Islam et al. 2014). Other works using complementary peptide nucleic acid (PNA) probes invading the G4-forming region of MYC to form heterologous PNA:DNA hybrid G4s have also downregulated MYC promoter activity (Gupta et al. 2013; Kumar et al. 2008; Roy et al. 2007). In addition,
sequence specific recognition and scission of human genome by PNA/DNA hybrid G4s have also been demonstrated (Ishizuka et al. 2012). On the one hand, this sequence specific scission of DNA highlights the potential of PNA in gene manipulation; on the other hand, it may also lead to irreversible toxicity in normal tissues. With respect to our clamp approach, the binding between clamp and the G4-flanking region is a reversible process, and thus will not cause permanent damage as with PNA/DNA hybrid G4s that are foreign to a cell. As the majority of unmodified DNA suffers from stability issues in cells, clamps modified to the more stable PNA or morpholino backbone will be pursued in the future to explore their potential.

Apart from the stability issues of unmodified DNA, all nucleic acids approach with few exceptions (DNA/RNA aptamers) face another challenge: in cell delivery. Two major classes of strategies have been employed to deliver nucleic acids into cells, including viral vector based transduction and liposome mediated transfection, with promising results seen from both strategies (Mali 2013). With liposomal-based strategies, FuGene and lipofectamine have been widely used for in vitro gene delivery, but neither these, nor viral vehicles are suitable for delivery of gene therapies systemically. In this work, we collaborated with a chemical engineering group that synthesized nine functionalizable blockcopolymers to deliver clamps into hard to transf ect lymphoma cells. By cellular uptake studies, we demonstrated qualitatively that clamp A 5T could not be taken up by cells alone, and that it required transfection. Once being transfected in cells, clamp A 5T was able to cross the nuclear membrane and get access to chromosomes in breast cancer and Burkitt’s lymphoma cells as evidenced by the colocalization noted by microscopy and ChIP-PCR assays. With the nine copolymers, only the ethyl series were able to effectively transf ect clamp A 5T into lymphoma cells. These findings are consistent with a recent set of studies on an unmodified DNA, PNT2258, which is an invading complementary oligonucleotide that targets the
distal region of the Bcl-2 promoter. This unmodified DNA also crosses the nucleus membrane and binds to its target region in the Bcl-2 promoter, facilitating apoptosis in lymphoma cells. This invading DNA PNT2258 is currently undergoing phase II clinical trials for diffuse B cell lymphomas, which makes it the first-in-class DNA interference (DNAi) therapy for cancer (Ebrahim et al. 2016; Tolcher et al. 2014). Both unmodified DNAs, Clamp A 5T and PNT2258, were able to readily cross the nuclear membrane and bind to their targets in the genome, which really highlights the potential of DNAi in future drug development. Further studies with the block copolymers demonstrated that the innate cytotoxicity of the ethyl series was dampened when they were complexed with DNA, suggesting the potential of these block copolymers for systemic delivery of gene therapies.

As discussed above, many MYC-targeting strategies have been employed either directly or indirectly (Mertz et al. 2011; Posternak and Cole 2016; Sabnis et al. 2017). Compare to any other strategies, targeting the MYC promoter G4 specifically using this clamp approach takes advantage of the relationship between MYC promoter G4 formation and transcription, which is significantly higher in cancer cells versus normal cells. As shown in our ChIP-PCR assay, when MYC transcription was activated by E2 in MCF-7 breast cancer cells, more binding of clamp A 5T to its target was observed. This data is consistent with the theory that higher transcription activity generates more negative superhelicity in the NHE III\textsubscript{1} region of the MYC promoter thus fostering more physiological G4 formation capable of being bound by clamp A 5T. In normal cells, MYC transcription activity is tightly under control, and less G4 formation will occur versus their transformed counterparts. This way, the presence of the in-cell G4 targets of the clamps is higher in cancer cells versus normal cells, which contributes to the selective gene regulation capability of
this clamp approach and many other future approaches that target the physiologically relevant MYC G4 specifically.

Previous research has shown that MYC upregulation in MCF-7 cells is largely due to copy number amplification, whether or not the transcription activity of an individual MYC copy was elevated is still unclear (Shadeo and Lam 2006). Based on our qPCR studies, MYC expression in MCF-7 cells was not significantly downregulated by clamp A 5T. This could possibly be a cell line specific effect due to the gene amplification mechanism of MYC in MCF-7 cells. Also, time course cellular uptake assays demonstrated that clamp A 5T is only evident between 6 and 24 hr. This potentially contributes to the lack of significant changes in MYC gene expression, especially considering the high turnover rate of MYC mRNA and increased gene copy numbers in MCF-7 cells.

The moderately MYC G4 selective compound NSC 338258 downregulates MYC expression within 6 hours in RAJI lymphoma cells (Brown et al. 2011). In the current work, we observed a notable, but not significant, decrease of MYC expression in RAJI cells when treated with clamp A 5T either alone or transfected by Fugene and E50 copolymer. As there were only duplicate sample in this study, further work is required to determine significance. Further investigations are required to define the dose- and time-dependent effects on clamp A 5T on MYC expression in both breast cancer and lymphoma cells.

Lastly, we explored the potential of using clamp A 5T as a diagnostic tool for specific MYC G4 detection in breast cancer and Burkitt’s lymphoma cells. Increasing evidence supports the bona fide G4 formation in cells. Several G4 detection methods have been developed to probe the intracellular G4 formation in human cells. Notably, a research group from the UK demonstrated the existence of endogenous DNA and RNA G4 formation in human cancer cells.
using an engineered antibody, termed BG4 (Biffi et al. 2013; Biffi et al. 2014). Other G4 detecting strategies also have been adopted by researchers to detect and characterize DNA structures in vivo. These detecting strategies can be generally classified into either internal or external-fluorescent probes based on their mode of interaction with G4 structures, or into small molecule or nucleic acids probes according to the nature of materials. These G4-detecting small molecules bind to a G4 structure non-covalently, which activates their fluorescence through fluorescence resonance energy transfer (Vummidi et al. 2013). However, these G4-interacting fluorescent probes are not without drawbacks. The majority of the small molecule probes are derived from previously discovered G4-interactive compounds, thus lacking specificity for individual G4s. For fluorescent antibodies or nucleic acid-based fluorescent probes, they mainly interact with pre-existing G4 structures, which has little effect on the DNA conformational equilibrium; and this mode of interaction is not enough to recognize distinctive features of an individual G4. Therefore, all current G4 detecting methodologies are still restricted to general G4 detection with no specificity for an individual structure.

With our clamp approach, we took advantage of the available unpaired nucleic acid bases flanking the MYC G4, which granted us the high specificity for a certain sequence thus a certain G4. As shown by the G4 induction assay, fluorescently labeled clamp A 5T, but no other fluorescent labeled DNA, was able to detect an increase in G4 formation induced by the selective MYC G4 stabilizing compound NSC 338258 in breast cancer and lymphoma cells. These observations further demonstrate that this clamp approach possesses high potential of being developed as an intracellular G4 recognition tool that evaluates the transcription activity of a certain gene. In this case, clamp A 5T can be potentially developed as a diagnostic tool to detect the MYC G4 formation in patient-derived cancer cells, correlating with hyperactivity of MYC
transcription. This information could be directive in terms of employing MYC G4-targeted therapies in the clinics.

Overall, this nucleic acid clamp-based approach described herein is a novel way to target individual promoter G4s that regulate transcription of key genes involved in cancer, neurodegenerative, infectious or other diseases. There are still issues to be addressed before this MYC clamp could be clinically advanced, and future work is likely to focus on modifying the backbone of the clamp, such as replacing the DNA backbone with PNA, to make them more resistant to cellular enzymes that degrade unmodified DNA. Also, the linker type of the clamps could be optimized in future by replacing the nonfunctional restricting segment with a G4-interacting moiety of the same length, thus presenting a dual-mode of action in terms of stabilizing the G4 structure. More importantly, the delivery of clamps into human cells should be tackled by optimizing the nano-based polymer complexation. To increase the effective concentration of the clamps in the nucleus, the clamps could also be optimized with the addition of a nuclear localization sequence, which enhances the nuclear entry as well as the cellular uptake.
CHAPTER IV. TARGETING NRAS 5’-UTR G4 WITH NA-CLAMPS

4.1. Introduction

The majority of G4 research has been focused on promoter or telomeric DNA over the past 20 years. However, G4s have been reported to form in RNA in cell as well. RNA molecules are transcribed as single stranded nucleic acids from DNA, and this nature of RNA allows for the formation of more complexed higher order structures, such as hairpins and G4s formed both intra- and inter-molecularly (Halder & Hartig, 2011). An RNA G4 was first discovered in the 3’ rRNA of Escherichia coli (Jin, Gaffney, Wang, Jones, & Breslauer, 1992). An mRNA G4 was discovered in the 3’-UTR of insulin-like growth factor II (Christiansen, Kofod, & Nielsen, 1994), and many have reported RNA G4s in both regulatory and coding regions of mRNA of many genes, such as NRAS, VEGF, and Bcl-2 (Agarwala, Pandey, & Maiti, 2015; Jayaraj et al., 2012; Kumari, Bugaut, Huppert, & Balasubramanian, 2007; Millevoi, Moine, & Vagner, 2012). Thermodynamically, RNA G4s are generally more stable than DNA G4s due to the additional interactions of ribose C2’ hydroxyl groups with the surrounding DNA structures (Agarwala et al. 2015b; Arora and Maiti 2009; Joachimi et al. 2009). Just as DNA G4s, RNA structures have been visualized in human cells using the specific BG4 antibody, demonstrating their bona fide existence (Biffi et al., 2014). One recent report utilizing mouse embryonic cells revealed that RNA G4s are globally unfolded, which introduced more confounding factors on the in-cell formation of RNA G4s. This finding was not repeated in human cells, further research is required into the prevalence of this
phenomenon, especially considering the epigenetic differences, which largely impact gene expression profiles, between human and mouse (Guo & Bartel, 2016).

Traditionally, mRNAs are translated through the cap-dependent mechanism, which involves three major steps to translate an mRNA into a protein. The first step is initiation, which is the rate limiting step that assembles ribosomes on the mRNA molecule. During initiation, a pre-initiation ribosomal complex is assembled onto the 5’-cap of the mRNA, which then scans down the mRNA until it reaches an initiation codon (AUG), which triggers the recruitment of a 60S ribosomal subunit. The second step is elongation, which synthesizes the polypeptides into a chain and the last step is termination, which disassembles ribosomes from the mRNA molecules (Gebauer and Hentze 2004).

Alternatively, the translation of mRNA could be initiated by non-cap dependent mechanisms involving internal ribosome entry sites (Baird et al. 2006). Both mechanisms of translation initiation are regulated by higher order RNA structures, including G4s (Bugaut and Balasubramanian 2012).

The RNA G4-forming sequences are highly enriched in the 5’-UTRs of the human genome (Huppert et al. 2008). Many genes containing 5’-UTR RNA G4s are associated with the hallmarks of cancer, as shown in table 1-1. NRAS, a human proto-oncogene contains one of the first identified 5’-UTR RNA G4s that functions as a translation inhibitor in vitro (Kumari et al. 2007). Various cellular activities are regulated by RNA G4s at both the DNA and RNA levels (Rhodes & Lipps, 2015). In the DNA level, RNA G4s can bind to RNA binding proteins that involves the regulation of DNA-related processes, such as recombination, or telomere elongation (Biffi et al. 2012; Zheng et al. 2015). In the RNA level, particular with 5’-UTR RNA G4s, these secondary
structures inhibit the translation of their corresponding proteins in many oncogenes, such as NRAS, Bcl-2, VEGEF, and TRF2 (Cammas & Millevoi, 2017).

The NRAS (neuroblastoma RAS viral oncogene homolog) proto-oncogene is located on the short arm of chromosome 1, at position 13.2. Together with kRAS, hRAS, mRAS and rRAS comprise the RAS oncogene family. They are intracellular signal cascade mediators, with intrinsic GTPase activity, functioning to initiate the mitogen-activated protein kinase-signaling (MAPK) pathway that is involved in cellular proliferation, differentiation and survival (Wennerberg et al. 2005). When RAS proteins are bound to GTP, they are activated and facilitate signal transduction; after the conversion of GTP into GDP, the proteins return to their inactive form, and GDP has to be exchanged for GTP to re-initiate the process. A single point mutation of the RAS genes is sufficient to convert them into oncogenes, which is locked in the GTP bound active state that relays proliferation signaling constantly. Most commonly, mutations with RAS occur at codon 12, 13, or 61 (Pylayeva-Gupta et al. 2011). For NRAS specifically, the most common mutation is at codon 61, converting glutamine to a lysine, arginine, or leucine, thus disrupting GTP hydrolysis (Hodis et al. 2012).

NRAS mutations are most commonly seen in melanoma (15-20%) and acute myeloid leukemia (AML) (43%) (Hodis et al. 2012); (Mardis et al. 2009). In melanoma, tumors tend to have either NRAS or BRAF mutation, but not both (Banerji et al. 2008). Recently, inhibition of BRAF have been clinically achieved with vemurafenib or dabrafenib; shifting the therapeutic focus to the 15-20% of melanomas with NRAS mutations (Kudchadkar et al. 2012). To date, no effective strategies yet been developed to target these malignancies, which contribute to the high mortality rate of melanoma patients. Discovering and developing therapies towards melanomas with NRAS mutations is urgent and crucial.
The RAS class proteins have long been the focus of drug discovery, with efforts focused on protein recruiting, membrane localization, and downstream signaling. RAS proteins bind to GTP with picomolar affinities while the abundance of GTP presence in cell approaches micromolar level. For this reason, nucleotide analogues that compete GTP binding from RAS proteins seem to be unpromising. Also, RAS proteins lack the active sites and binding pockets that could potentially be targeted by small molecules. For these reasons, RAS proteins have been considered as “undruggable” for the past several decades (Gysin et al. 2011). Since hydrophilic RAS proteins go through post-translational modification in order to achieve membrane localization, strategies targeting these have also been employed (Fiordalisi et al. 2003). Farnesyltransferase inhibitors (FTIs) have been demonstrated to be effective for blocking Ras farnesylation, subcellular localization and activity, thus inhibiting the growth of Ras-driven cancer cells (Brock et al. 2016). However, studies have shown that the most commonly mutated RAS isoforms, kRAS and NRAS, are still active even in the presence of FTIs, possibly due to lipid modification by geranylgeranyltransferase (Whyte et al. 1997). In addition, downstream effector signaling pathways of RAS, such as Raf-MEK-ERK and PI3K-AKT-mTOR signal transduction pathways have also been targeted. A synthetic benzyl styryl sulfone compound, called rigosertib, is currently in phase III clinical trials for chronic myeloid leukemia. This promising small molecule inhibitor can simultaneously inhibit both MAPK and PI3K signaling thus inhibiting tumor growth (Z. Liu et al. 2017; Prasad et al. 2016; F. Xu et al. 2014). Besides targeting RAS proteins directly, inhibiting the translation of RAS proteins has been explored using antisense or RNA interference strategies. Even though these strategies are promising in cells and some rodent models, the lack of safe and effective delivery vehicles to human patients is restricting their successful clinical applications (Smakman et al. 2005).
In this work, we explored inhibiting the translation of NRAS mRNA via stabilizing its 5'-UTR G4 (Figure 4-1). Previous research identified a G4-forming sequence in the 5'-UTR of the human NRAS mRNA, located 14 bps downstream of the 5'-cap and 222 bps upstream of the translation start site, which is evolutionarily conserved thorough many species including human, chimpanzee, macaque, rat, mouse and dog (Kumari et al. 2007). The study identified the NRAS RNA G4 as a stable parallel intramolecular G4, in the presence of 100 mM KCl. Mutational studies demonstrated that disruption of the NRAS RNA G4 led to a 3.6-fold increase in translation of NRAS protein in vitro. Further studies showed that the position of the G4 within the 5'-UTR of NRAS, and the number of guanines in the sequence, are key factors that affect the translational efficiency of NRAS in vitro (Kumari et al. 2008). These studies were the first to identify the NRAS RNA G4 as one that represses translation in vitro.

To date, only a small molecule approach has been employed primarily to target NRAS RNA G4 (Cammas and Millevoi 2017). General G4-interacting compounds, such as TMPYP4, that bind to both DNA and RNA G4s were shown to interact with this structure. In addition, two RNA-selective G4 stabilizers have been developed, including the polyacromatic molecule RGB1 and carboxyPDS (carboxy pyridostatin). RGB1 was shown to stabilize the NRAS RNA G4, but not DNA G4s or any other RNA structures. Also, this compound significantly downregulated NRAS expression in MCF-7 breast cancer cells at the translational level, supporting the hypothesis that RGB1 stabilized the NRAS 5’-UTR G4 (Katsuda et al. 2016). This study indicated that stabilizing the 5’-UTR G4 of NRAS is a possible strategy to downregulate NRAS expression. However, the selectivity of RGB1 for an individual RNA G4 remains to be evaluated. Most likely, it can interfere with multiple RNA G4 structures, similarly to the majority of DNA G4 stabilizers, leading to possible off-target effects.
Herein, we designed two NA-clamps (Figure 4-2) specifically for the NRAS 5’-UTR G4 in order to stabilize this translationally repressing element. Similar to the MYC clamps, NRAS clamps are composed of the DNA complementing the flanking region and a spatial restricting element (polyethylene glycol) that locks G4. The recognition and stabilizing ability of the designed clamps toward NRAS G4, their cytotoxicity profile, and their in-cell activities were evaluated.
NRAS mRNA G-Quadruplex

Figure 4-1. Human NRAS 5’-Untranslated Region (5’-UTR) and the mRNA G-Quadruplex. A parallel G4 structure (shown in red) is 14 nucleotides downstream of the 5’-cap and 222 nucleotides upstream of the translational start site. The entire sequence is the 5’-UTR of the NRAS gene; sequence highlighted is the NRAS WT<sub>ext</sub> oligonucleotide used in current research, and the G4-forming region is marked in red (Adapted from Kumari et al. 2007).
Figure 4-2. The central hypothesis of NA-clamps mediated stabilization of the NRAS 5’-UTR RNA G-quadruplex. The G4-forming region (red) of NRAS plus 14 bps on both 5’- and 3’-ends (top) was used to design clamp 1 and clamp 2 with either 14 or 10 bps flanking the 17.7 Å long polyethylene glycol linker, respectively. The expected G4 plus clamp complex (bottom) is predicted to downregulate NRAS translation.
4.2. Recognition and structural stabilization of the NA-clamps to NRAS G4

As an RNA G4, both the 5’ and 3’ flanking region of the G4 are usually unpaired and can be utilized as complementing bp for NA-clamps. Therefore, we designed two NA clamps, differing by the number of complementing bp on each side of the flanking region (14 bp for clamp 1 and 10 bp for clamp 2) to recognize and stabilize the NRAS RNA G4. Based on previous experience with the MYC clamp, the linker length was kept at 17.7 Å by using the same polyethylene glycol linker as MYC clamp A (Figure 4-2). Even though both the physiologically relevant MYC promoter G4 and the identified NRAS 5’-UTR G4 are parallel G4s, but the loop length difference of the two G4 structures might cause differential spatial distance. All other oligonucleotides used in this chapter were listed in Table 2-2.

The G4 recognition of the NA-clamps was evaluated by EMSA (Figure 4-3A). Clamp 1 and clamp 2 were incubated with NRAS WT<sub>ext</sub> and the G4 knockout MT<sub>ext</sub> sequences, in the presence of 100 mM KCl. Mixture DNAs with clamps were heated up to 95 °C and allowed to anneal slowly to room temperature. By EMSA, clamp 1 was seen to induced three upwardly shifted bands, while clamp 2 only induced one, with WT<sub>ext</sub> sequence, as compared to the control WT<sub>ext</sub>. With the G4 knockout MT<sub>ext</sub> sequence (complete G to T replacements on the second guanine run), only two faster migrating bands were observed with clamp 1, while one band was noted with clamp 2, as compared to the MT<sub>ext</sub> sequence alone, indicating the two faster migrating bands with clamp 1 and the single upwardly shifted band with clamp 2 are the complementing DNA complex introduced by the binding of clamps. Compared to the MT<sub>ext</sub> sequence, clamp 1 introduced a supra-shifted band with the WT<sub>ext</sub> sequence, which may be a G4: clamp 1 complex. Further, ECD studies (Figure 4-3B) confirmed that the WT<sub>ext</sub> sequence alone (solid line) showed no observable G4 structure formation. Upon the addition of clamp 1 (dashed line) and clamp 2 (dotted line), more
pronounced G4 signature cotton effects at 265 nm were observed, indicating G4 formation was induced, more notably by clamp 1. Also, a broadening of the spectrum around 280 nm was observed with both clamp 1 and clamp 2, indicating the formation of dsDNA between the flanking region of WT\textsubscript{ext} and clamp 1. As compared to clamp 2, clamp 1 produced a more intense G4 signature, as well as the broadening of the spectrum around 280 nm. As compared to EMSA, only clamp 1 showed the supra-shifted band, but not clamp 2. However, in ECD, both clamps were able to induce more G4 formation. Possibly, clamp 2 can bind to one side of the flanking region but not both sides at the same time, thus no G4 is locked in between. Our previous study has shown that truncated NRAS sequence containing only the G4-forming region can form a more stable G4 structure in the same condition, as compared to NRAS WT\textsubscript{ext} (data not shown). By binding to either side of the flanking region, clamp 2 could potentially change the structural equilibrium of the NRAS WT\textsubscript{ext} sequence, which allowed more G4 formation. Further studies with fragmented clamps and WT\textsubscript{ext} need to be conducted to confirm this hypothesis.

Previous research by the group discovered the NRAS RNA G4 showed this G4 structure is a very stable with 100 mM KCl (Kumari et al. 2007), which is different from the ECD findings observed above. The different oligonucleotide studied need to be considered. They used an RNA sequence containing only the G4-forming region plus three base pairs on the 5′-end (5′-UGUGGAGGGGCGGGUCUGGG-3′), while we have to adopt a longer NRAS WT\textsubscript{ext} DNA sequence containing not only the G4-forming region, but also 14 flanking base pairs on both 5′- and 3′-end of the G4 to enable clamp binding. Studies have shown that G4 structures formed by RNA are thermodynamically more stable than G4s formed by their DNA counterparts, which might contribute the differences observed above (Agarwala et al. 2015b; Arora and Maiti 2009; Joachimi et al. 2009). Also, the additional complementing base pairs added on the flanking region
of the G4 can also disrupt the G4 formation of the region. To address this problem, we added acetonitrile (ACN), an osmolyte, used to induce the physiologically relevant structure in a test tube (Miller et al. 2010). The effect of ACN and clamp 1 or clamp 2 on the G4 formation of WT$_{ext}$ were evaluated by EMSA and ECD (Figure 4-4).
Figure 4-3. Binding of linked clamps 1 and 2 to the NRAS G4. (A) The binding of clamps 1 and 2 to NRAS WT<sub>ext</sub> and a G4 knockout MT<sub>ext</sub> was examined by EMSA. Three upwardly shifted bands were observed to occur between WT<sub>ext</sub> and clamp 1, while only one upwardly shifted band was observed with clamp 2. With the G4 knockout NRAS sequence, clamp 1 produced only two upwardly shifted bands that matches the two faster migrating bands while clamp 2 produced only one upwardly shifted band, as with the WT sequence. (B) The control DNA alone (NRAS WT<sub>ext</sub>) demonstrated no G4 formation by ECD. Upon the addition of clamp 1 (dashed line) and clamp 2 (dotted line), a signature parallel G4 signal around 270 nm was induced. Also, a broadening of the spectrum around 275-280 nm, denoting the formation of dsDNA on the flanking region, was recorded. Compared to clamp 2, clamp 1 fostered more G4 with the NRAS WT<sub>ext</sub> sequence, which is consistent with the observations from EMSA.
Upon the addition of 40% ACN (according to a ACN gradient study with WT<sub>ext</sub>, data not shown), clamp 1 fostered notably more supra-shifted banding, as compared to without ACN, indicating G4 formation while clamp 2 mediated only the complementing dsDNA, as evidenced by the darker bandings, as compared to without ACN. ECD was used to probe the structural changes after the addition of ACN and clamps. Without osmolyte, either WT<sub>ext</sub> alone (blue) or with 100 mM KCl (red) showed no G4 formation which is consistent with spectra recorded previously (Figure 4-3). Upon the addition of 40% ACN (green), a pronounced signature G4 cotton effect was induced, as evidenced by the minimum and maximum at 240 nm and 265 nm respectively, indicating the formation of a parallel G4 structure. Upon the addition of clamp 2 (teal), the intensity of the G4 cotton effect was unchanged. Notably, upon the addition of clamp 1 (purple), a more intense 265 nm cotton effect was observed, as well as a broadening of the spectrum around 275-280 nm, indicating the induction of G4 and the formation of dsDNA on the flanking region. These results further demonstrated that clamp 1, but not clamp 2 is able to facilitate NRAS G4 formation.

The G4-induction ability of clamp 1 was evaluated by comparing the binding patterns of pre-annealing and post-annealing clamp 1 with WT<sub>ext</sub> (Figure 4-5). Without the addition of osmolyte (Figure 4-5B), the incubation of clamp 1 with NRAS WT<sub>ext</sub> only fosters supra-shifted bands in a dose-dependent manner when clamp 1 and NRAS WT<sub>ext</sub> were annealed together (pre-annealing). With post-annealing samples, no visible supra-shifted band was observed indicating no G4 was detected by clamp 1, which is consistent with the ECD data showing no stable G4 formation by NRAS WT<sub>ext</sub> without the presence of ACN. Upon the addition of osmolyte (Figure 4-5A), clamp 1 fostered more supra-shifted banding, as compared to the pre-annealed sample without ACN. Notably, with post-annealing samples, clamp 1 introduced a light supra-shifted
banding in the presence but not absence of ACN, indicating clamp 1 recognizes the NRAS G4 formation.

Next, the G4 stabilization ability of the NA-clamps on NRAS WT_{ext} sequence was examined by DMS footprinting study. All samples used in this study were pre-annealed under the conditions used in Figure 4-6. NRAS WT_{ext} alone demonstrated no observable protection of the four guanine runs capable of participating G4 formation. Compared to WT alone, ACN facilitated G4 formation as evidenced by the lighter bandings from the four-participating guanine runs (red). Upon the addition of clamp 1, a more noticeable guanine protection pattern from the four guanine runs (red) was observed compared to either the control sample or sample annealed with ACN, demonstrating NRAS G4 stabilization by clamp 1. As compare to control, no observable guanine protection was detected when NRAS WT_{ext} was annealed with clamp 2, suggesting clamp 2 is not effective at stabilizing the NRAS G4.
Figure 4-4. Binding of clamps to NRAS G4 under physiologically relevant condition. (A) ECD confirmed that 40% ACN (teal) facilitated the formation of NRAS G4, as compared to either NRAS WT alone (blue) or with 100 mM KCl (red). Upon the addition of clamp 1, a more notable signature parallel G4 cotton effect plus a broadening of the spectra around 275-280 nm was observed. While with clamp 2, no change was observed. (B) Clamps were incubated with NRAS WT<sub>ext</sub> with ACN pre-annealingly. Similar binding patterns were observed with both clamp 1 and clamp 2 as previously. With clamp 1, ACN promoted the formation of the supra-shifted band. With clamp 2, only the slower migrating band was observed.
Figure 4-5. NRAS G4 induction by annealing with clamp 1. (A) Without the presence of ACN, incubation of clamp 1 with WT<sub>ext</sub> induced G4 formation only when clamp 1 was added before pre-annealing but not post-annealing. NRAS G4 was induced by as low as 0.5 equivalence of clamp 1 through pre-annealing incubation, but no G4 induction was observed even with two equivalences of clamp 1 by post-annealing incubation. (B) Upon the addition of 40% ACN that mimics the physiological condition, one equivalence of clamp 1 induced a lighter supra-shifted band with post-annealing incubation, indicating clamp 1 binds to the G4 formed by WT<sub>ext</sub> sequence. With pre-annealing incubation, supra-shifted bands were observed at both ACN present and absent groups, which further demonstrated that clamp 1 facilitates NRAS G4 formation at both conditions.
Figure 4-6. **Clamp 1-mediated stabilization of NRAS G4.** DMS-footprinting was utilized to confirm the NRAS G4 stabilization by clamp 1. NRAS G4-forming region was marked in red. Compared to without ACN, slight protection of all four guanine runs was observed indicating some G4 formation. Upon the addition of clamp 1, lighter bands from all four guanine runs were observed, either compared to the control or ACN group, demonstrating G4 stabilization by clamp 1. With clamp 2, no protection pattern of guanine was detected, indicating no G4 stabilization by clamp 2.

<table>
<thead>
<tr>
<th></th>
<th>NRAS WT &lt;sub&gt;ext&lt;/sub&gt;</th>
<th>KCl (100 mM)</th>
<th>40% ACN</th>
<th>Clamp 1</th>
<th>Clamp 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protection</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

![NRAS G4-forming region marked in red.](image)

GTGGAGGGGGCGGGTGCCCTGC
4.3. *Intracellular activity of NRAS clamp 1*

The experiments conducted above demonstrate that clamp 1 is able to recognize and stabilize NRAS G4. In this section, the cytotoxicity of clamp 1 will be evaluated on cancerous melanoma WM 3682 and non-cancerous HEK-293 cell lines. Also, the effect of clamp 1 and clamp 2 on NRAS expression will be assessed on melanoma cells.

NRAS-mutated cancers are often addicted to their constitutive RAS activity, and down-regulating protein expression is sufficient to decrease cell survival in melanoma cells only driven by NRAS mutation (Eskandarpour et al. 2005). A previous study successfully induced apoptosis of melanoma cells harboring only an NRAS mutation with targeted RNA interference, confirming that down-regulating NRAS mRNA expression is able to elicit apoptosis in addicted cells (Eskandarpour et al. 2005). In this study, a human metastatic melanoma cell line WM 3682 (different from the study above), which contains only NRAS mutation (Q61L) but no BRAF, c-KIT, PTEN, or any other common melanoma related gene mutations, was used for cellular viability and protein expression assays.

To test the effect of clamp 1 on cell survival, we used the non-cancerous cell line HEK-293 and the malignant melanoma cell line WM 3682 (Figure 4-7). Increasing doses of clamp 1 were transfected into HEK-293 and WM 3682 cells in micelles formed with FuGene HD one day post cell seeding, standard MTS assay was performed to measure cellular viability 48 and 72 hr post-transfection. Dose-dependent decrease of cellular viability in both cell lines at both time points were observed. Forty-eight hr post-transfection, melanoma cells showed higher sensitivity to clamp 1 treatment, as compared to non-cancerous HEK-293 cells, their ED$_{50}$S were 670 and 1040 ng, respectively. The differential response of the two cell lines to clamp 1 treatment was gone by 72 hr transfection.
Based on the cytotoxicity data, a range of doses (500, 1000, 2500, 5000 ng) of clamp 1 and clamp 2 were transfected into WM 3682 melanoma cells to assess if the dose-dependent decrease of cellular viability of the melanoma cells by clamp 1 coincided with NRAS down-regulation. Western blot analysis was performed to determine NRAS protein levels on samples collected 48 hr post-transfection (Figure 4-8A). Based on the densitometry analysis (Figure 4-8B), no dose-dependent decrease of NRAS protein level was observed by clamp 1 and a dose dependent increase was observed with clamp 2; further investigations of translational regulation at more time points remains necessary. Also, the intracellular stability of the clamps need to be evaluated.
Figure 4-7. Cytotoxicity of clamp 1 on cancerous melanoma WM 3682 and non-cancerous HEK-293 cell lines. Dose-dependent decreases of cellular viability were observed with both cell lines 48 and 72 hr post-transfection of clamp 1. At 48 hr, WM 3682 melanoma cells showed a notable lower ED_{50} compared to non-cancerous HEK-293 cells. However, the difference of the IC_{50} was diminished 72 hr post-transfection.
Figure 4.8. Effect of clamps on NRAS translation. (A) Clamp 1 was transfected into WM 3682 melanoma cell by FuGene HD. Cells were lysed and a standard western blot was conducted to determine NRAS protein levels 48 hr post-transfection. Beta-actin was used as an internal control for protein expression. (B) Densitometry analysis of the immunoblots from (A) was performed using LiCor Odyssey software to quantify NRAS protein levels. No dose-dependent NRAS expression changes were observed with clamp 1 transfection. With clamp 2, a dose-dependent increase of NRAS expression was observed.
4.4. Discussion

RNA G4s play important roles in translational regulation. Pioneering research using a eukaryotic cell free translation system demonstrated a naturally occurring G4 formed in the 5’-UTR of NRAS mRNA is important in modulating the translational efficiency of the NRAS oncogene (Kumari et al. 2007). After this milestone work, other G4s formed in the 5’-UTR or 3’-UTR of mRNAs were also shown to inhibit the translation of their corresponding genes, such as ZIC-117, BCL-2, MP3- MMP18, ERS1, TRF2, and PIM1 (Arora et al. 2008; Arora and Suess 2011; Balkwill et al. 2009; Gomez et al. 2010; Morris and Basu 2009; Shahid et al. 2010). Notably, the research on ZIC-117 5’-UTR mRNA G4 was the first study that showed an RNA G4 structure that can repress gene translation in living human HeLa cells (Arora et al. 2008). Many more potential G4s formed in the regulatory regions of mRNA, based on bioinformatics analysis, remain to be studied (Huppert et al. 2008). For the current work, we focused on the first discovered NRAS 5’-UTR G4 to evaluate the efficacy of the NA-clamp approach to stabilize naturally occurring RNA G4s, thus modulating gene translation.

To date, small molecules have been the main approach taken to target RNA G4 structures. Particularly for the NRAS RNA G4, two small molecules - RR110 and RGB1, have been developed to stabilize this G4 thus inhibiting the translation of NRAS. Notably, RGB1 not only showed inhibition efficacy of NRAS translation in vitro, and downregulated endogenous NRAS protein level in breast cancer cells (Bugaut et al. 2010; Katsuda et al. 2016). RGB1 was shown to be RNA versus DNA G4 selective, but selectivity for the NRAS RNA G4 specifically remains to be demonstrated. RR110 was shown to stabilize NRAS G4, but not selective over DNA G4 or dsDNA based on their screening strategy. Compared to small molecule strategy, we employed an NA-clamp-based approach with high specificity for targeting the NRAS RNA G4 structure.
In this work, we designed two NA-clamps (clamp 1 and 2) to recognize and stabilize NRAS RNA G4. However, only clamp 1 demonstrated its NRAS G4 recognition and stabilization capability. These two clamps are both composed of two pieces of oligonucleotides complementing the 5’- and 3’- flanking regions of NRAS RNA G4 and a polyethylene glycol linker connecting these together. Clamp 1 contains 28, as compared to 20, bps of complementary nucleotides, which would be expected to grant clamp 1 higher binding affinity to the NRAS G4. Our previous work targeting the MYC DNA G4 with clamp A, using the same polyethylene linker, showed that 18 bps on the 5’-side plus 5 bps on the 3’-side of the clamp, 23 bps altogether, was sufficient to grant high binding affinity and specificity to its target G4 sequence (Hao et al. 2016). Compared to clamp 2, clamp 1 possesses only four more base pairs on each side of the complementing oligonucleotides; However, these eight extra base pairs made a marked difference in terms of NRAS G4 recognition and stabilization, indicating the sufficient base pairs complementing the flanking region of the NRAS G4 is required to achieve stabilization.

DNA oligonucleotides rather than RNA nucleotide sequences was used in this study in order to be more cost efficient. Previous research established that RNA G4s are more stable than DNA G4s thermodynamically (Agarwala et al. 2015b; Arora and Maiti 2009; Joachimi et al. 2009). The initial work, which described NRAS RNA G4, showed this parallel G4 structure is very stable thermodynamically with a $T_M$ of $> 95 \ ^\circ C$ (Kumari et al. 2007). In this work, we showed that the DNA version of the same NRAS WT sequence, containing only the G4-forming region, had a much lower $T_M$ (79 $^\circ C$) compared to its RNA counterparts. Moreover, the longer flaking region further disrupted the formation of this G4.

Therapeutically, RAS proteins, including NRAS, have been well validated as molecular targets for a variety of cancer types. Particularly for NRAS, around 15% of melanomas are driven
by NRAS mutations, and are considered untreatable. Targeted therapeutic development for melanomas harboring NRAS mutation is urgent and necessary. Previous RAS targeting approaches have focused on inhibiting the membrane localization of the proteins or targeting downstream signaling pathways (Brock et al. 2016; Fiordalisi et al. 2003; Z. Liu et al. 2017; Prasad et al. 2016; Smakman et al. 2005; Vojtek et al. 1993; F. Xu et al. 2014). However, none of the above approaches has made it to the clinic. Other than small molecules, siRNA has also been used to target NRAS expression with promising efficacy on downregulating NRAS expression. However, the delivery of siRNA systemically is an obstacle for patients. In addition, NRAS mutations commonly occur in codon 61 with three major different mutation types, including Q61K, Q61R, Q61L (Hou et al. 2012), thus requiring siRNAs to be multiple designed against individual mutations to achieve high selectivity. In contrast, modulating NRAS translation by targeting the 5′-UTR G4 of its mRNA is not restricted by the mutation types of the gene. Therapeutics developed by targeting NRAS G4 therefore will have a wider application window compared to siRNAs designed for individual mutations.

Despite demonstrating stabilization of the NRAS 5′-UTR G4, endogenous protein downregulation was not observed in melanoma cells 48 hr post-transfection of clamp 1. This result could be caused by an instability of clamp 1 in cytoplasm; stability studies need to be completed. Previous study showed that an RNA G4 stabilizing small molecule, RGB1, was able to downregulate endogenous NRAS protein expression in cells, indicating a role for the structure itself (Katsuda et al. 2016). On the other hand, a recent study showed that RNA G4 structures are globally unfolded in eukaryotic cells (Guo and Bartel 2016). Researchers from the same study above proposed that the NRAS G4 sequence can function as a protein binding site, and it is the proteins bound to the sequence that sequester translational machinery, thus inhibiting translation.
rather than the G4 structure formed by the sequence (Guo and Bartel 2016). Notably, if there was no initial G4 formation, and the sequence was pre-occupied by proteins, the access of clamp 1 to the G4-forming sequence would be blocked. Further research should be carried out to identify the potential binding proteins of the NRAS G4 to unravel the mystery of G4 targeting with NA-clamp approach.
CHAPTER V. CONCLUSION

G-quadruplexes (G4s) are non-canonical secondary nucleic acid structures formed by DNA or RNA sequences, usually in the biologically important regions such as promoters of DNA, 5’-UTRs of mRNA and telomeres (Verma et al. 2008). This biological importance has resulted in the emergence of targeting G4 structures as a therapeutic strategy for the regulation of oncogene expression on both DNA and RNA levels (Balasubramanian et al. 2011, Cimino-Reale et al. 2016). G4 mediated transcriptional and translational silencing mainly functions through sequestering the binding of transcription factors, polymerases and translational machineries to the promoters of DNA or the 5’-UTRs of mRNA. This has allowed for the development of multiple strategies to stabilize G4 structures in the promoter or 5’-UTR of oncogenes in order to downregulate oncogene expression.

Extensive efforts have been exerted to develop G4 stabilizing small molecules (Maji and Bhattacharya 2014). Notably, two G4-interacting compounds, quarfloxin and CX5461, have entered phase II clinical trials for treating carcinoid/neuroendocrine and triple negative breast cancer, respectively. While quarfloxin failed phase II clinical trial due to high albumin binding, it represents the first in class G4-interacting compound that provided insights for G4 targeting in cancer treatment. CX5461 is still under active clinical investigation to determine its clinical efficacy. For MYC specifically, NSC338258, a representative small molecule, displayed preferential binding to MYC promoter G4 with nanomolar affinity and downregulated MYC
expression in lymphoma cells. However, this compound also showed binding and stabilization to other promoter G4s, such as the ones in VEGF and HIF-1α genes — stressing the importance of target specificity (Brown et al. 2011).

Nucleic acid-based strategies have also been employed to target certain oncogene promoters, such as MYC and Bcl-2. In particular, MYC G4-forming sequence pu27 has been utilized in leukemia cells, which resulted in profound damage to both telomeric and non-telomeric regions of DNA, inducing cancer cell apoptosis (Islam et al. 2014; Sedoris et al. 2012). Guanine rich PNAs have also been used to target the G4-forming region of MYC by forming PNA/DNA hybrid quadruplexes thus downregulating MYC expression (Gupta et al. 2013; Roy et al. 2007). Other types of DNAi work have showed promising effect in treating cancer. Notably, one unmodified oligonucleotide PNT 2258 was utilized to target the distal region of the BCL-2 promoter to downregulate BCL-2 expression and inducing apoptosis of cancer cells. PNT 2235 is currently under phase II clinical trials for diffuse large B cell lymphoma (Ebrahim et al. 2016; Tolcher et al. 2014). Although PNT 2228 is a non-G4-mediated approach, it still highlights the potential of using nucleic acid based approaches as cancer therapeutics.

To target individual G4s with high specificity, we adopted a novel nucleic acid-based approach to stabilize G4s. Clamps comprise two segments of oligonucleotides complementing to the 5’ and 3’- flanking regions of the G4s with a linker connecting the two segments together. The complementing oligonucleotides on the flanking region of the G4 grant the clamps high binding specificity, while the linker is optimized to the proximal distance of the G4 — enabling the clamps to recognize and stabilize the G4 structure. In this work, we designed clamps that stabilize MYC promoter G4 and NRAS 5’-UTR G4 to demonstrate this clamp approach on both DNA and RNA level.
MYC promoter G4 is the most well characterized G4 structure to date. Previous research has identified multiple G4 structures formed in the guanine and cytosine rich NHE III region of the MYC promoter, including G4_{1-4}, G4_{2-5}, and G4_{1-5} (Dettler et al. 2011; Siddiqui-Jain et al. 2002; Sun and Hurley 2009). To date, the physiologically relevant MYC G4 has been identified as a parallel structure G4_{1-4}, with the loop length of 1:2:1 from the 5’ to 3’ direction, under supercoiled DNA conditions, which you would find in a cell under physiological conditions (Sun and Hurley 2009). MYC promoter G4 is a silencing element of MYC transcription, and stabilizing this physiologically-relevant MYC G4 is a practical strategy to downregulate MYC expression (Brooks and Hurley 2010). For example, small molecule NSC 338258 downregulates MYC expression in lymphoma cells by stabilizing the physiologically relevant MYC promoter G4 (Brown et al. 2011).

In this work, we utilized a different stabilizing strategy, NA-clamps, that possess higher specificity towards the physiologically relevant MYC G4_{1-4}.

Clamp A presented specific recognition to its designed G4 target G4_{1-4} but not other G4 isoforms in the equilibrium within the NHE III region. Also, clamp A stabilized G4_{1-4} in vitro, as demonstrated by DMS-footprinting. More importantly, clamp A was able to stabilize this physiologically relevant G4_{1-4} in supercoiled plasmids and downregulates MYC promoter activity in living cells, as evidenced by luciferase reporter assay. A cellular viability assay also showed that clamp A possesses differential cytotoxicity profile towards cancerous vs non-cancerous cell lines.

To reduce the cost and increase the modification flexibility of clamp A, we optimized clamp A by replacing the original polyethylene glycol linker with five thymines that share the same length with the linker on clamp A, and yielded a new clamp termed clamp A 5T. With a dramatic cost reduction and more flexible modification potential, we demonstrated that clamp A
5T was able to specifically recognize and stabilize the physiologically relevant MYC G4 just as clamp A. Using fluorescently labeled clamp A 5T, we identified that clamp A 5T was able to be transfected into living cells, and more importantly, localized in the nucleus. By using biotinylated clamp A 5T, we were able to immunoprecipitate the physiological binding target of clamp A 5T in human breast cancer cells and confirmed its binding to the endogenous MYC promoter DNA, as supported by ChiP-PCR experiment.

All nucleic acids based therapeutics, including the NA-clamp approach discussed herein, face one common challenging problem — the intracellular delivery. To deliver NA-clamps in cells, especially hard-to-transfect lymphoma cells, we collaborated with a chemical engineering group that synthesized nine block copolymers to facilitate delivery. Among these nine block copolymers, we identified that the ethyl series copolymers were able to effectively transfect both RAJI and CA46 Burkitt’s lymphoma cells with fluorescent labeled clamp A 5T. Based on fluorescent microscopy studies, we discovered that FAM-clamp A 5T was only stable up to 24 hour intracellularly. Lastly, we evaluated the effect of clamp A 5T on the transcription of MYC in MCF-7 and RAJI cells. Even though no significant downregulation of MYC was observed 24 hr post-transfection in both cell lines, there was a decrease of MYC in cells treated with clamp A 5T in both cells. The in-cell stability of clamp A 5T needs to be optimized before its effect on MYC transcription could be confirmed.

Lastly, we explored the potential application of clamp A 5T as a clinical diagnostic companion for the indication of G4 small molecule employment. Through G4 induction assay, we showed that fluorescently labeled clamp A 5T was able to detect induced MYC G4 formation in fixed and permeabilized breast cancer and lymphoma cells by increasing concentrations of MYC G4 selective small molecule NSC 338258. This exploration further demonstrated that clamp A 5T
is highly specific for MYC G4 recognition. Clamp A 5T could be potentially be developed as a diagnostic companion in the clinics to detect the hyperactivity of MYC transcription in cancer cells and inform the usage of MYC selective G4 stabilizing small molecules.

At the translational level, RNA G4s are repressors of gene translation into their protein products (Bugaut and Balasubramanian 2012). Particularly, NRAS 5’-UTR G4 was the first reported mRNA G4 that functioned as a translation silencing element (Kumari et al. 2007). Small molecules have been developed to target NRAS RNA G4 to downregulate NRAS expression. Most notably, RGB1, a RNA G4 selective G4 stabilizing compound, has demonstrated its efficacy in downregulating endogenous NRAS protein level in breast cancer cells (Katsuda et al. 2016). However, the specificity of RGB1 over NRAS G4 and its mechanism of action still requires further elucidation. In this work, we applied NA-clamp approach to the NRAS 5’-UTR G4 by designing two clamps (clamp 1 and clamp 2) for NRAS G4 specifically. Based on the binding assays, we determined that clamp 1 but not clamp 2 was able to recognize NRAS G4 formation, with clamp 1 exhibiting NRAS G4 stabilization as demonstrated by DMS-footprinting. Through cellular viability assay, we observed that clamp 1 was cytotoxic to both cancerous melanoma and non-cancerous HEK-293 cells. However, clamp 1 showed much lower IC50 on cancerous melanoma cell versus non-cancerous HEK-293 cells, indicating selective cytotoxicity before clamps were digested in cells. Unfortunately, no NRAS protein downregulation was observed in melanoma cells after treating with either clamp 1 or clamp 2. However, a recent finding revealed that RNA G4s are universally unraveled in eukaryotic cells, and the G4-forming sequences are bound by multiple proteins to maintain the single stranded form of RNA (Guo and Bartel 2016). If this holds true in melanoma cells, then it is understandable that clamp 1 was not able to get access to its physiological target in cells thus further pharmacological effects would not be observed.
Overall, we demonstrated that NA-clamp approach could be applied to both promoter G4 and 5’-UTR G4 for specific recognition and stabilization, providing the foundation for the future development of NA-clamps into effective therapeutics, or as diagnostic companions for MYC, NRAS, or other G4-related cancer types.


Biffi, G., Tannahill, D., and Balasubramanian, S. (2012), 'An intramolecular G-quadruplex


Boxer, R. B., et al. (2004), 'Lack of sustained regression of c-MYC-induced mammary adenocarcinomas following brief or prolonged MYC inactivation', *Cancer Cell*, 6 (6), 577-86.


Dahm, R. (2008), 'Discovering DNA: Friedrich Miescher and the early years of nucleic acid research', *Hum Genet*, 122 (6), 565-81.

Dalla Favera, Riccardo, Wong-Staal, Flossie, and Gallo, Robert C (1982), 'Onc gene amplification in promyelocytic leukaemia cell line HL-60 and primary leukaemic cells of the same patient'.

Dalla-Favera, R., et al. (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 U S A*, 79 (24), 7824-7.


Dexheimer, T. S., et al. (2009), 'NM23-H2 may play an indirect role in transcriptional activation of c-myc gene expression but does not cleave the nuclease hypersensitive element III(1)', Mol Cancer Ther, 8 (5), 1363-77.


Drygin, Denis, et al. (2008), 'Quarflvoxin (CX-3543) disrupts the Nucleolin/rDNA quadruplex complexes, inhibits the elongation by RNA Polymerase I and exhibits potent antitumor activity in models of cancer', (AACR).


Egler, R. A., et al. (2005), 'Regulation of reactive oxygen species, DNA damage, and c-Myc function by peroxiredoxin 1', Oncogene, 24 (54), 8038-50.


Hao, T., Gaerig, V. C., and Brooks, T. A. (2016), 'Nucleic acid clamp-mediated recognition and stabilization of the physiologically relevant MYC promoter G-quadruplex', Nucleic Acids


Ivancich, M., et al. (2017), 'Treating Cancer by Targeting Telomeres and Telomerase', *Antioxidants (Basel)*, 6 (1).


Jin, R., et al. (1992), 'Thermodynamics and structure of a DNA tetraplex: a spectroscopic and calorimetric study of the tetramolecular complexes of d(TG3T) and d(TG3T2G3T)', *Proc Natl Acad Sci U S A*, 89 (18), 8832-6.


Kumar, N., et al. (2008), 'Silencing c-MYC expression by targeting quadruplex in P1 promoter using locked nucleic acid trap', Biochemistry, 47 (50), 13179-88.


Levens, D. (2008), 'How the c-myc promoter works and why it sometimes does not', J Natl
Cancer Inst Monogr, (39), 41-3.


Liu, J., et al. (2006), 'The FUSE/FBP/FIR/TFIIH system is a molecular machine programming a pulse of c-myc expression', EMBO J, 25 (10), 2119-30.


Liu, Z., Sun, Q., and Wang, X. (2017), 'PLK1, A Potential Target for Cancer Therapy', Transl Oncol, 10 (1), 22-32.


Morris, M. J. and Basu, S. (2009), 'An unusually stable G-quadruplex within the 5'-UTR of the


Piazza, A., et al. (2010), 'Genetic instability triggered by G-quadruplex interacting Phen-DC


Xu, H., et al. (2017), 'CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours', Nat Commun, 8, 14432.


Zhang, Z., et al. (2016), 'BET Bromodomain Inhibition as a Therapeutic Strategy in Ovarian Cancer by Downregulating FoxM1', Theranostics, 6 (2), 219-30.

VITA

Taisen Hao was born in Shaanxi, China on October 17, 1995. He graduated from Beijing University of Chinese Medicine with a degree in Pharmaceutical Engineering in 2013. The following fall, he enrolled in the Ph.D. program of Pharmaceutical Sciences (Pharmacology) at The University of Mississippi. Taisen earned his doctoral candidacy in March 2016, and received his Ph.D. in Pharmaceutical Sciences (Pharmacology) in July 2017.