THERMODYNAMIC CHARACTERIZATION OF TETRAPLEX DNA STRUCTURES

by

BYUL KIM

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy

Department of Pharmaceutical Sciences
Leslie Dan Faculty of Pharmacy
University of Toronto

© Copyright by Byul Kim 2018
ABSTRACT

Thermodynamic Characterization of Tetraplex DNA Structures

Byul Kim

Doctor of Philosophy

Department of Pharmaceutical Sciences
Leslie Dan Faculty of Pharmacy

University of Toronto

2018

The role of counterion condensation as a dominant force governing the stability of DNA duplexes and triplexes is well established. In contrast, the effect of counterion condensation on the stability of G-quadruplex conformations is poorly understood. Unlike other ordered nucleic acid structures, G-quadruplexes exhibit a specific binding of counterions (typically, Na\(^+\) or K\(^+\)) which are buried inside the central cavity and coordinated to the O6 carbonyls of the guanines forming the G-quartets. While it has been known that the G-quadruplex-to-coil transition temperature, \(T_M\), increases with an increase in the concentration of the stabilizing ion, the contributions of the specific (coordination in the central cavity) and nonspecific (condensation) ion binding have not been resolved. In the first part of the work performed in this dissertation, we used G-quadruplexes formed by four different sequences derived from the human telomeric region, c-MYC, and VEGF genes to separate the two ionic contributions. We studied the change in \(T_M\) of preformed G-quadruplexes following the addition of nonstabilizing ions Li\(^+\), Cs\(^+\), and TMA\(^+\) (tetramethylammonium). Our data suggest that the stabilizing action of cations on the G-quadruplex conformation is, primarily, due to the central ions which act as specifically bound
ligands. Nonspecifically bound (condensed) counterions may slightly stabilize the G-quadruplex conformation or exert no influence on its stability (human telomeric G-quadruplexes) or may strongly destabilize it (c-MYC and VEGF). After investigating the salt effect in G-quadruplexes, the latter part of the thesis project was devoted to understanding the folding/unfolding transitions of i-motif, a distinct tetraplex structure formed by c-rich complementary sequences of G-quadruplexes. We described the pH-induced folding/unfolding transitions of i-motifs formed by c-MYC sequence by a linkage thermodynamics-based formalism in terms of three pK$_a$’s of cytosines, namely, an apparent pK$_a$ in the unfolded conformation, pK$_{au}$, and two apparent pK$_a$’s in the folded state, pK$_{af1}$ and pK$_{af2}$. 
I will love thee, O LORD, my strength.

Psalm 18:1
To my parents and family.
Acknowledgments

I would like to thank my supervisor Dr. Tigran V. Chalikian. I truly respect and appreciate the effort he put in to raise me as a doctor. I am the outcome of your patience and generosity, Dr. Chalikian!

I also would like to express appreciation to my committee members Dr. Robert B. Macgregor, Dr. Jeffrey Henderson, and Dr. Giuseppe Melacini. Thank you all for your encouragement and advice. I could not have done it without you.

Thank you, all my friends and colleagues. To name a few, I would like to express my sincere gratitude to Dr. Yuen Shek, Dr. Ikbae Son, Dr. Yang Li, Dr. Helen Fu, Pei Wang, Tameshwar Ganesh, Ujala Feroze, Lutan Liu, Alah Amsdr, and Jay Lee. I will not forget the quality times I spent with you. Thank you for encouraging me when I was down at times.

Lastly, but not least, I am very much indebted in love to my parents, sisters, and my one and only brother-in-law. Thanks Mom and Dad for always being there for me. Your prayer and words of encouragement have been and will always mean so much to me. I am blessed to be your daughter. My sisters, Jeeyeon and Sohyoung! I respect you big sissys for always holding up with my whining and moaning. Thank you for being my life-time best friends. My big brother Sungjin! I will never forget your patience in everything, especially in listening to my endless stories. Thank you so very much. Love y’all!
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................ VI

TABLE OF CONTENTS ........................................................................................................ VII

PUBLISHED WORKS .......................................................................................................... XI

LIST OF TABLES ................................................................................................................ XII

LIST OF FIGURES .............................................................................................................. XIII

**CHAPTER 1: General Introduction** ................................................................................ 1

1.1: History of Nucleic Acid Discovery .......................................................................... 1

1.2: Emergence of G-Quadruplexes .............................................................................. 1

1.3: G-quadruplexes *in vivo* ......................................................................................... 2

1.3.1 Telomeric DNA ...................................................................................................... 3

1.3.2 Oncogene Promoters ............................................................................................ 4

1.3.3 Drugs that Stabilize G-quadruplexes .................................................................. 5

1.4: Structural Properties of G-quadruplexes ................................................................. 6

1.4.1 Building Blocks ..................................................................................................... 6

1.4.2 Hydrogen Bonding ............................................................................................... 8

1.4.3 Stacking Interactions ............................................................................................ 9

1.5: Topological Diversity of G-quadruplexes ................................................................. 10

1.5.1 Tetramolecular G-quadruplexes ........................................................................ 11

1.5.2 Bimolecular G-quadruplexes ............................................................................. 11

1.5.3 Intramolecular G-quadruplexes ......................................................................... 13

1.6: Cations and Nucleic Acid Structure ...................................................................... 15

1.6.1 DNA as a polyelectrolyte .................................................................................... 15
CHAPTER 3: Conclusion

3.6: Conclusion .................................................................................................................. 93

3.7: Acknowledgements ..................................................................................................... 94

3.8: Reference .................................................................................................................... 95

CHAPTER 4: Ionic effects on VEGF G-quadruplex Stability .............................................. 113

4.1: Abstract ...................................................................................................................... 114

4.2: Introduction ............................................................................................................... 115

4.3: Materials and Methods ............................................................................................. 117

4.4: Results ...................................................................................................................... 119

4.5: Discussion ................................................................................................................. 130

4.6: Concluding Remarks ................................................................................................. 134

4.7: Supporting Information ............................................................................................. 135

4.8: Acknowledgements ................................................................................................... 136

4.9: References ............................................................................................................... 137

CHAPTER 5: Thermodynamic linkage analysis of pH-induced folding and unfolding
transitions of i-motifs ....................................................................................................... 144

5.1: Abstract ...................................................................................................................... 145

5.2: Introduction ............................................................................................................... 146

5.3: Experimental Section ............................................................................................... 155

5.4: Acknowledgements ................................................................................................... 156

5.5: Reference .................................................................................................................. 157

CHAPTER 6: General Conclusion and Future Perspectives ............................................. 162

6.1: Conclusions .............................................................................................................. 162

6.2: Future Prospect ........................................................................................................ 164
6.3: Reference ........................................................................................................... 166

APPENDIX A: Effect of Urea on G-quadruplex Stability ........................................ 167
A.1: Abstract .............................................................................................................. 168
A.2: Introduction ....................................................................................................... 169
A.3: Materials and Methods .................................................................................... 170
A.4: Results .............................................................................................................. 172
A.5: Discussion ........................................................................................................ 181
A.6: Conclusions ..................................................................................................... 189
A.7: Acknowledgements .......................................................................................... 190
A.8: Reference .......................................................................................................... 191
PUBLISHED WORKS


LIST OF TABLES

Chapter 1

Table 1-1  Ionic radii of G-quadruplex stabilizing cations.

Chapter 3

Table 3-1  Molecular volumes, $V_M$, and solvent accessible surface areas, $S_A$, of the c-MYC and human telomeric Tel22 and Tel26 G-quadruplexes. Values normalized per nucleotide are shown in parentheses.
LIST OF FIGURES

Chapter 1

Figure 1-1 Components of a Nucleotide.

Figure 1-2 Base Geometry (Syn and Anti glycosylic bonds).

Figure 1-3 Guanine Nucleotides.

Figure 1-4 G-quartet.

Figure 1-5 Stacking of G-tetrads.

Figure 1-6 Schematic diagrams of various topologies of bimolecular quadruplexes. (a) lateral loops connecting antiparallel strands to form head-to-head hairpin quadruplex; (b) lateral loops connecting antiparallel strands on different sides; (c) external and lateral loops; (d) external (propeller) chain reversal loops linking parallel strands; (e) diagonal linkers connecting opposite anti-parallel strands.

Figure 1-7 Basket-type Anti-parallel G-quadruplex.

Figure 1-8 Guanine Tetrads and Associated Waters.

Figure 1-9 Cytosine-Cytosine+ Base Pairs.
Chapter 2

Figure 2-1  CD spectra of Tel22 in the presence of 20 mM NaCl and 0 (●), 180 (○), 380 (■), and 780 (□) mM LiCl (panel A), CsCl (panel B), and TMACl (panel C).

Figure 2-2  CD spectra of Tel26 in the presence of 20 mM KCl and 0 (●), 180 (○), 380 (■), and 780 (□) mM LiCl (panel A), CsCl (panel B), and TMACl (panel C).

Figure 2-3  UV melting profiles at 295 nm of Tel22 at 100 mM NaCl (●) and Tel26 at 100 mM KCl (○).

Figure 2-4  (a) Dependences of the melting temperature, $T_M$, of Tel22 on the concentration of Na$^+$ (●), Li$^+$ (○), Cs$^+$ (■), and TMA$^+$ (□) cations; (b) Dependences of the melting temperature, $T_M$, of Tel26 on the concentration of K$^+$ (●), Li$^+$ (○), Cs$^+$ (■), and TMA$^+$ (□) cations.

Chapter 3

Figure 3-1  (a) CD spectra of MYC22-G14T/G23T in a 10 mM TBA-phosphate buffer at various KCl concentrations at 25 °C. (b) Dependence of the molar ellipticity of MYC22-G14T/G23T at 264 nm on the KCl concentration. Experimental data were approximated with Eq. (1) for n = 3 (solid line).

Figure 3-2  (a) Representative UV melting profile at 295 nm of the MYC22-G14T/G23T G-quadruplex at 1 mM KCl. (b) Dependence of the melting temperature, $T_M$, of the
MYC22-G14T/G23T G-quadruplex on the concentration of K\(^+\) ions.

Experimental data were approximated by an exponential function solely to guide the eye (solid line).

**Figure 3-3**  (a) CD spectra of the MYC22-G14T/G23T G-quadruplex at 300 μM KCl at various CsCl concentrations. (b) CD spectra the MYC22-G14T/G23T G-quadruplex preformed in 300 μM KCl with subsequently added 800 mM CsCl at temperatures ascending from 25 to 80 °C and descending to 25 °C.

**Figure 3-4**  (a) CD spectra of the MYC22-G14T/G23T G-quadruplex at 300 μM KCl at various TMACl concentrations. (b) CD spectra the MYC22-G14T/G23T G-quadruplex preformed in 300 μM KCl with subsequently added 800 mM TMACl at temperatures ascending from 25 to 80 °C and descending to 25 °C.

**Figure 3-5**  Dependences of the melting temperature, \(T_M\), of the MYC22-G14T/G23T G-quadruplex on the concentration of CsCl at 300 (a), 600 (b), and 1000 (c) μM KCl. Experimental data were approximated by an exponential function solely to guide the eye (solid lines).

**Figure 3-6**  Dependences of the melting temperature, \(T_M\), of the MYC22-G14T/G23T G-quadruplex on the concentration of TMACl at 300 (a), 600 (b), and 1000 (c) μM KCl. Experimental data were approximated by an exponential function solely to guide the eye (solid lines).
Figure 3-7  CsCl-dependence of the molar ellipticity of MYC22-G14T/G23T at 264 nm in 600 μM KCl at 51 °C. Experimental data were approximated with Eq. (3) (solid line).

Chapter 4

Figure 4-1  (a) CD spectra of Pu22-T12T13 in a 10 mM TBA-phosphate buffer at various KCl concentrations at 25 °C. (b) Dependence of the molar ellipticity of Pu22-T12T13 at 264 nm on the KCl concentration. Experimental data were approximated with Eq. (1) for n = 2 (solid line).

Figure 4-2  (a) Dependence of the melting temperature, T_M, of the Pu22-T12T13 G-quadruplex on the concentration of K⁺ ions. Experimental data were approximated by a linear function (solid line). (b) The G-quadruplex-to-coil transition van't Hoff enthalpies of Pu22-T12T13 plotted as a function of the transition temperature, T_M, at 2 (■), 5 (□), and 10 (●) mM KCl and TMACl concentrations ranging from 0 to 800 mM and 2 (○), 5 (♦), an 10 (◊) mM KCl and CsCl concentrations ranging from 0 to 100 mM.

Figure 4-3  CD spectra of the Pu22-T12T13 G-quadruplex at 2 (a), 5 (b), and 10 (c) mM KCl at 25 °C and various concentrations of TMACl.

Figure 4-4  CD spectra of the Pu22-T12T13 G-quadruplex at 2 (a), 5 (b), and 10 (c) mM KCl at 25 °C and various concentrations of CsCl.
Figure 4-5  Dependences of the melting temperature, $T_M$, of the Pu22-T12T13 G-quadruplex on the concentration of TMACl at 2 (a), 5 (b), and 10 (c) mM KCl. Experimental data were approximated by an exponential function solely to guide the eye (solid lines).

Figure 4-6  Dependences of the melting temperature, $T_M$, of the Pu22-T12T13 G-quadruplex on the concentration of CsCl at 2 (a), 5 (b), and 10 (c) mM KCl. Experimental data were approximated by an exponential function solely to guide the eye (solid lines).

Chapter 5

Figure 5-1  The CD spectra of the ODN between pH 9.5 and 4.5 (panel A) and pH 4.8 and 2.3 (panel B).

Figure 5-2  The pH-dependence of the molar ellipticity at 284 nm. Experimental data were approximated with Eq. (2) (solid line).

Figure 5-3  The pH-dependence of the differential number of protons bound to the folded and unfolded states of the ODN calculated with Eq. (3).

Figure 5-4  Transition profiles simulated with Eq. (2) for $pK_{au}$, $pK_{af1}$, and $pK_{af2}$ of 4.8, 6.0, and 3.6, respectively, and $K_0$ equal to 0.0001 (dark yellow), 0.001 (green), 0.005 (blue), 0.02 (black), and 0.1 (red).
Transition profiles simulated with Eq. (2) for i-motifs containing 6 (black) and 8 (red) C·C* base pairs with \( pK_{au} \), \( pK_{a1} \), and \( pK_{a2} \) of 4.8, 6.0, and 3.6, respectively, and \( K_0 \) of 0.0001.

**Appendix A**

**Figure A-1** The CD spectra of Tel22 at 25 °C in the absence of urea at various NaCl concentrations. The NaCl concentration (in mM) for each spectrum is given in the inset.

**Figure A-2** The CD spectra of Tel22 at 25 °C at various urea concentrations at 20 (panel a), 50 (panel b), 100 (panel c); 200 (panel d); 400 (panel e); 600 (panel f); 800 (panel g); and 1000 (panel h) mM NaCl. The urea concentrations (in M) for the spectra are given in the insets.

**Figure A-3** The CD spectra of the urea-induced unfolded state at 20 mM NaCl and 8 M urea (○) and temperature-induced unfolded states at 20 mM NaCl and 95 °C in the absence of urea (■) of Tel22. The inset presents the temperature dependence of the molar ellipticity at 257 nm of the urea-induced unfolded state at 20 mM NaCl and 8 M urea.

**Figure A-4** Representative UV melting profile at 295 nm Tel22 at 200 mM NaCl and 1 M urea. Experimental data were approximated with Eq. (1) (solid line).
**Figure A-5**  The G-quadruplex-to-coil transition temperature, $T_M$, transition temperatures of Tel22 as a function of urea at various NaCl concentrations. The NaCl concentration (in mM) for each set of $T_M$ is given in the inset.

**Figure A-6**  Compilation of transition enthalpies, $\Delta H_M$, of Tel22 measured at various combinations of NaCl and urea concentrations plotted against corresponding transition temperatures, $T_M$.

**Figure A-7**  The dependence of the molar ellipticity of the Tel22 at 295 nm on the urea concentration at 20 mM NaCl. Experimental data were approximated with Eq. (2) (solid line).

**Figure A-8**  The three-dimensional dependence of the transition temperature, $T_M$, on the concentrations of NaCl and urea.
CHAPTER 1: General Introduction

1.1: History of Nucleic Acid Discovery

In 1953, Franklin and Wilkins crystallized a DNA molecule and recorded its X-ray diffraction patterns. Subsequently, Watson and Crick used the diffraction patterns to construct three-dimensional model of the DNA. The proposed model envisages a double helical structure in which two separate 5’-3’ strands run in an anti-parallel manner. The two strands are interconnected via hydrogen bonds between complementary bases resembling steps of a ladder. In agreement with the chemical evidence of uniform composition offered by Chargaff and his colleagues, the Watson - Crick model of a DNA duplex offers a molecular rationalization of the complementarity of A and T and G and C bases.

The Watson and Crick model has since been corroborated through numerous fiber diffraction experiments and single-crystal X-ray measurements. Structural studies of nucleic acids not only produced a deeper understanding of the canonical nucleic acid structure but also unveiled the environment and sequence-dependent polymorphism of DNA. The conformational flexibility of nucleic acids under various environmental conditions and in the presence of DNA-binding proteins hints at the diversity of DNA structures adopted at different genomic loci and during various phases of the cell cycle.

1.2: Emergence of G-Quadruplexes

More than 40 years prior to the development of the double helical DNA model, the Norwegian chemist Ivar Christian Bang found that concentrated samples of guanylic acid from ox pancreas could form viscous gels. Guanylic acid at a concentration of 25.0 mg/mL forms a thick aqueous solution which suggests that self-association may take place. This simple yet novel proposal has caused many controversies as some researchers argued that guanylic acid, a
simple nucleotide similar to inosinic acid, could not form higher order structures\textsuperscript{10}. This argument persisted until 1962, when Gellert and collaborators employed X-ray crystallography to illustrate that guanine bases, indeed, arrange themselves into a large helical tetraplex structure now referred to as a G-quadruplex\textsuperscript{11}.

1.3: G-quadruplexes in vivo

In 2005, Huppert and Balasubramanian scanned through the human genome using the algorithm \textit{Quadparser} to assess the prevalence of the G-quadruplex-forming sequences\textsuperscript{12}. These authors have devised an algorithm that can be used to identify the sequences d(G\textsubscript{3+}N\textsubscript{1-7} G\textsubscript{3+}N\textsubscript{1-7} G\textsubscript{3+}N\textsubscript{1-7} G\textsubscript{3+}) that may fold into putative G-quadruplex structures\textsuperscript{12}. Theoretically, there are 375,000 candidate sequences capable of folding into G-quadruplexes\textsuperscript{12}. Although the sequences identified through the bioinformatics analysis may or may not adopt a G-quadruplex structure, the specific dispersion pattern suggests that the tetraplex formation may play a vital role at certain loci of the human genome\textsuperscript{12}.

The Balasubramanian group (2013) has also succeeded in visualizing G-quadruplexes in human cells using structure-specific antibodies\textsuperscript{13}. They have observed discrete G-quadruplexes distributed throughout the genome\textsuperscript{13}. Telomeric ends at chromosomes favour formation of G-quadruplexes given the abundance of guanine-rich tandem repeats and the presence of single-stranded 3’-overhangs\textsuperscript{14}. It is worth noting, however, that the absence of the complementary strand is not a requirement for formation of G-quadruplexes. The canonical double helical DNA can undergo transient unwinding in the middle of the transcription or replication events, thereby providing an ideal condition for tetraplex DNA formation at a number of functionally important genomic loci such as promoter regions of proto-oncogenes\textsuperscript{14}. Proto-oncogenes are essential for normal functioning of cells but once deregulated, they become oncogenes whose expression can lead to cancer initiation. The evidence of G-quadruplexes existing at sites crucial for transcriptional control and senescence marks these structures as potential targets for therapeutics.
1.3.1 Telomeric DNA

In the early 1990s, Blackburn and her colleagues identified G-rich tandem repeats positioned at the end of chromosomes\textsuperscript{15}. This region, which measures about 5-15 kilobases in length, is now referred to as a telomere\textsuperscript{16}. The G-rich tandem repeat sequences are highly conserved across different species. Telomeres are composed of a double-stranded section that is a few thousand nucleotides long followed by a short dangling terminal end consisting of 100-200 bases\textsuperscript{16}. Understandably, the duplex region has a complementary C-rich strand running in the opposite direction, whereas the single-stranded overhang is unprotected and exposed to the environment. Some common examples of the repeat units are TTAGGG in mammals\textsuperscript{17}, TTGGGG in the ciliate Tetrahymena\textsuperscript{18,19}, and TTTTGGGG in the protozoa Oxytricha nova\textsuperscript{20}. The biomedical relevance of telomeres stems from the fact that their length is an accurate timer of cellular aging\textsuperscript{21-25}. Telomeres are destined to shorten by about 40 to 50 bases with each DNA replication. When the chromosomal terminus reaches a critical length, the cell is signaled to undergo apoptosis\textsuperscript{26}. There is an enzyme that elongates telomeres in order to systematically decelerate this aging clock\textsuperscript{27}. The enzyme, telomerase, is a reverse transcriptase that has an RNA template complementary to the guanine rich telomeric sequence\textsuperscript{28}. Nonetheless, this molecule is rarely abundant in human tissues unless the cell requires a high degree of development\textsuperscript{156}. Hence, telomerase is active in germ cells or stem cells where cell proliferation is key\textsuperscript{29,30}. It is essential to control the telomerase activity. There is much evidence indicating that the maintenance of telomeric ends is central to making cells overproliferative\textsuperscript{31,32,156}. Telomerase activity is very high in about 90\% of immortalized cells, including tumour cells\textsuperscript{156}. G-quadruplexes may play an essential role in this respect, since the formation of a tetraplex structure in the telomeric region has been shown to inhibit the telomerase activity\textsuperscript{33,34}. Finding ways to control cellular immortalization and incessant replication is the ultimate pathway to targeting tumourigenesis. Thus, G-quadruplexes have been evolved as a potential target in cancer therapy.
1.3.2 Oncogene Promoters

Transcriptional regulatory regions within the genome also have a high propensity to fold into G-quadruplexes. Putative G-quadruplex-forming sequences rich in guanine bases are found in 50% of all identified human genes, in particular, in promoters and nuclease hypersensitive elements within 1000 bases upstream of the transcription initiation site. The nuclease hypersensitive elements and promoter regions are 230 times more likely to adopt G-quadruplex structures relative to other genomic loci. Nuclease hypersensitive site is an area proximal to the promoter that is easily detected by cleavage enzymes such as DNase. This region is characterized by the absence of histones (which make chromosomes compact) thereby enhancing the flexibility of the DNA and its propensity to fold into a G-quadruplex structure. Although there is a large number of oncogene promoters, the regulatory pathways of which still need to be investigated, there are quite a few of them that have been well scrutinized. One example is c-MYC that encodes a transcription factor that is essential for regulating cell proliferation. Deregulation of c-MYC expression triggers uncontrolled multiplication of cells which is associated with tumorigenesis in a variety of tissues. c-MYC can either act as a primary oncogene that becomes amplified or translocated or it can activate downstream pathways of other oncogenes in a chain reaction. G-quadruplex formation at the promoter of c-MYC is thought to provide steric hindrance and impair the binding of initiation proteins. This notion suggests that small ligands that stabilize tetraplex DNA can act as antitumor drugs.

Another well investigated proto-oncogene is vascular endothelial growth factor (VEGF). VEGF plays an essential role in angiogenesis by producing growth factors that are vital for generating embryonic circulatory system or for growing blood vessels from pre-existing vascular structure. Solid tumours that are about 1mm long in diameter usually depend on other host blood vessels to acquire oxygen and nutrients. However, when cancer cells cannot sustain their activities with the substances they obtain from existing blood cells, they become angiogenic themselves by gaining the capability to generate new blood vessels. The angiogenic ability is activated in tumour cells with elevated expression of VEGF. The over-expression of human VEGF observed in many tumor tissues can be regulated at the stage of transcription. One way of achieving such a regulatory effect is to induce a G-quadruplex at the promoter region of the
oncogene to prevent the binding and suppress the activity of transcriptional proteins. The oligomeric sequences derived from the c-MYC and VEGF promoter regions are the main subject of the experiments described in this thesis.

1.3.3 Drugs that Stabilize G-quadruplexes

The ultimate aim of scrutinizing the distribution and biological relevance of G-quadruplex-forming sequence is to develop an ability to induce G-quadruplexes at selected genomic loci via intervening low molecular weight compounds (drugs). In 1991, Zahler and his colleagues reported the inability of Oxytricha telomerase to elongate the telomeric oligonucleotide coiled into a G-quadruplex form in vitro\(^48\). Given this observation, much work has been invested to find natural molecules and their derivatives that can trigger G-quadruplex folding and disrupt telomere elongation in vivo\(^49\). Numerous compounds representing various chemical classes (such as porphyrin\(^50\), acridine\(^51\), perylene\(^52\), and bisquinolinium\(^51,53\)) have been shown to impede telomerase activity.

Telomestatin is one of the better understood high-affinity G-quadruplex binding molecules. This natural compound isolated from Streptomyces anulatus is a powerful telomerase inhibitor\(^54\). It stabilizes the G-quadruplex conformation, particularly, the anti-parallel topology, thereby blocking the telomerase activity\(^54\). Telomestatin blocks telomere elongation within time periods that are too short to measure if used at a dose higher than 5\(\mu\)M\(^55,56\). At concentrations lower than 2\(\mu\)M, telomestatin has been shown to promote a gradual telomere degradation or programmed cell death in numerous tumor cell lines in vivo\(^57\). Telomestatin also induces a quick dissociation of the telomere capping proteins TRF2 and Pot1\(^58\). Other low molecular weight compounds that bind G-quadruplexes and are capable of uncapping the 3’ ends of chromosomes are BRACO19 and the acridine RHPS4\(^59,60\). The selective uncapping of chromosome by these drugs in cancer relative to normal somatic cells was a matter of concern at the early stages of research\(^59,60\). It is presently known that the affinity of these drugs for tumour cells is significantly higher compared to that for noncancerous cells for not very well understood reasons\(^56,61\).
A number of agents that bind to G-rich promoter sequences and modulate the transcription of proto-oncogenes have been identified and investigated. A porphyrin derivative TMPyP4 down-regulates the expression of c-MYC proto-oncogene in HeLa cells\(^\text{62}\). The Hurley lab has shown that TMPyP4 binds specifically to the c-MYC nuclease hypersensitive element (NHE)\(^\text{63}\). When the NHE\(\text{III}_1\) region of c-MYC in Burkitt’s lymphoma cell lines was removed, the porphyrin had no effect on c-MYC expression studied\(^\text{63}\).

There are a few G-quadruplex-binding ligands that have gone through clinical investigations\(^\text{64}\). Quarfloxin, a fluoroquinolone-based antitumor agent, is the first G-quadruplex interactive small molecule to have progressed through Phase II clinical trials\(^\text{64}\). This molecule was discovered by Hurley and his colleagues\(^\text{65, 66}\). It is tolerated well in neuroendocrine cancer patients and its effect is mediated through disruption of the interaction between nucleolin and G-quadruplex in rDNA, an important interaction amplified in tumour cells\(^\text{67}\). Other prospective G-quadruplex stabilizing molecules currently undergoing clinical trials include CX5461 and AS1411\(^\text{67}\). The quest for G-quadruplex-binding agents for specific therapeutic applications within the genome has evolved into an important component of anticancer strategies.

1.4: Structural Properties of G-quadruplexes

1.4.1 Building Blocks

The building block of a G-quadruplex is a G-quartet, a planar assembly of four guanines interlinked through Hoogsteen hydrogen bonds\(^\text{68}\). Similar to other DNA structures, G-quadruplexes are composed of nucleotides that involve pentose sugars attached to phosphate groups (Figure 1-1)\(^\text{7}\). Though there are four bases (adenine, cytosine, thymine, and guanine) that are indiscriminately used in other DNA structures, G-quadruplexes need guanine bases to form their primary building blocks\(^\text{35}\). The torsion angle created by the bond between the base and the 1’ carbon of the sugar is referred to as angle \(\chi\). Depending on the angle, the nucleotide can be classified as \textit{anti} (-120<\(\chi\)<180°) or \textit{syn} (0<\(\chi\)<90°) (Figure 1-2). The \textit{anti} configuration is sterically more favourable relative to the \textit{syn} configuration; the latter can be stabilized through
various external factors\textsuperscript{35}. The degree of this angle is a significant variant among different G-quadruplex conformations\textsuperscript{35}.

![Components of a Nucleotide](image.png)

Figure 1-1. Components of a Nucleotide\textsuperscript{35}. 

![Syn and Anti conformations](image.png)
Figure 1-2. Base Geometry (Syn and Anti glycosyl bonds).

The pentose sugar configuration is also a distinctive feature of different nucleic acid conformations. The C2’-endo sugar pucker observed in B-DNA is the stable form in G-tetrads\textsuperscript{35}.

1.4.2 Hydrogen Bonding

Nucleic acid bases contain hydrogen donors (amino or imino hydrogen) and acceptors (carbonyl oxygen or aromatic nitrogen)\textsuperscript{7}. Hydrogen bonding between bases dictates the specific conformation DNA can fold into at given environmental conditions. In a double-helical DNA, pyrimidines form Watson-Crick hydrogen bonds with complementary purine bases. In G-quadruplexes, guanine bases engage in two hydrogen bonds with alternation of Hoogsteen and Watson-Crick faces as shown in Figure 3\textsuperscript{35}. In a G-quartet, the nitrogen and the oxygen attached to the ring carbon, N7 and O6, serve as hydrogen bond acceptors with N2-H and N1-H serving donors (Figure 4)\textsuperscript{35}.

Figure 1-3. Guanine Nucleotides\textsuperscript{35}.
1.4.3 Stacking Interactions

DNA bases are non-polar and their interactions with water, a polar solvent, are energetically unfavourable\(^7,\,72\). In order to minimize the area exposed to the aqueous environment, nucleic acid bases with planar ringed structures form stacking interactions. In the case of G-quadruplexes, planar G-tetrads, each composed of four guanines, aggregate on top of each other to create a structure similar to a jenga tower\(^71,\,72\). The stacking that contributes to DNA stability originates from a combination of hydrophobic, electrostatic, and van der Waals forces\(^35,\,70-73\). When the planar bases gather in close proximity to each other, \(\pi\) orbitals cause repulsion and create an unfavourable energetic state. However, stacked aromatic bases enable formation of \(\pi-\pi\) bonding central to G-quadruplex stabilization\(^7,\,35\). Concurrently, the hydrogen atoms around the ring become electron-poor. Base stacking in G-quadruplexes interfaces is
important for their stability\textsuperscript{35}. Due to steric constraints, G-quartets stack with a twist of $30^\circ$ and a rise of 3.3 Å with each elevation (Figure 5)\textsuperscript{35}. The eight carbonyl oxygen atoms of stacked bases eventually arrange into a bipyramidal antiprismatic construct\textsuperscript{35,74}.

![Figure 1-5. Stacking of G-tetrads\textsuperscript{35}.](image)

### 1.5: Topological Diversity of G-quadruplexes

G-quadruplexes can be widely classified based on the number of strands involved in the structure. Intramolecular conformations comprise a single strand, whereas intermolecular quadruplex structures can be composed or two or more (2 and 4, typically) sequences\textsuperscript{35}. The orientation of the strand involved in G-quadruplex formation offers another level of classification. If the adjacent strands have opposite directionalities, the conformation is designated as anti-parallel, while G-quadruplex is parallel if all strands have the same orientation\textsuperscript{49}. G-quadruplexes can also consist at a mixture of parallel and anti-parallel strands. The conformation adopted by a
G-quadruplex depends on the nucleotide sequence, number of stacked G-tetrads, and numerous environmental factors such as the cation type and concentration and the extent of hydration\textsuperscript{35}.

### 1.5.1 Tetramolecular G-quadruplexes

A tetramolecular G-quadruplex is composed of four strands rich in guanine bases. The guanines of each strand interact with the guanines from adjacent strands. Theoretically, there are a number of possibilities in which the strands can be oriented, but, experimentally, only the all-parallel structure has been observed\textsuperscript{35}. High resolution X-ray and NMR studies have reported that tetramolecular tetraplexes are extremely stiff\textsuperscript{104}. They fold from four $X_nG_mX_n$ or $G_mX_nG_m$ sequences\textsuperscript{120}. However, tetramolecular G-quadruplexes did not receive as much attention because it is not found in human genome.

### 1.5.2 Bi-molecular G-quadruplexes

When less than four strands come together to form a G-quadruplex structure, loops come into play. The linkers (loops) can be classified as lateral, diagonal, or external\textsuperscript{14, 35}. As shown in Figure 6, the lateral loops of a bimolecular G-quadruplex can be placed either on the same or the opposite side. When the linkers are on the same face, a head-to-head hairpin G-quadruplex forms\textsuperscript{120}. However, when the loops are on the opposite side, all strands become anti-parallel to each other, creating a head-to-tail lateral loop dimer\textsuperscript{120}. When at least one strand is anti-parallel to the other strands, the G-quadruplex conformation is classified as anti-parallel\textsuperscript{120}. The anti-parallel configuration has guanines with alternating \textit{syn} and \textit{anti} glycosidic angles\textsuperscript{35}.
Figure 1-6. Schematic diagrams of various topologies of bimolecular quadruplexes; (a) lateral loops connecting antiparallel strands to form head-to-head hairpin quadruplex; (b) lateral loops connecting antiparallel strands on different sides; (c) external and lateral loops; (d) external (propeller) chain reversal loops linking parallel strands; (e) diagonal linkers connecting opposite anti-parallel strands.\textsuperscript{35}
1.5.3 Intramolecular G-quadruplexes

The G-quadruplex conformation can be adopted by a monomolecular sequence containing G-tracts, separated by regions that may consist of any combination of residues. The consensus sequence of an intramolecular G-quadruplex forming motif is \((G_mL_nG_mL_pG_mL_qG_m)_r\) where \(m\) is the number of guanine repeats, and \(L_n, L_p, \) and \(L_q\) denote any combination of nucleotides forming loops.

The length and the sequence of the loops determine the specific G-quadruplex conformation of the DNA molecule. Monomolecular structures may contain propeller, lateral, and diagonal loops. Hazel and his colleagues have postulated a general loop-length dependent folding pattern of G-quadruplex structures as described below. Using a combination of circular dichroism measurements and molecular dynamics simulations on ten intramolecular G-quadruplexes, it has been reported that three single nucleotide linkers facilitate formation of a parallel structure. When single nucleotide loops are present along with longer loops, DNA may fold into both parallel and anti-parallel G-quadruplexes. Guedin et al. (2010), based on the analysis of more than 80 DNA sequences, have concluded that single nucleotide linkers displayed the highest stability. The two studies are in agreement with each other as it has been reported that parallel structures are thermodynamically more stable than anti-parallel G-quadruplexes.

Conformational preferences of intramolecular G-quadruplexes also depend on the identity of the stabilizing cation. The same sequence may frequently adopt different conformations depending on the cation present in the solution. For instance, the human telomeric sequence d(CTAGGG)₄ may fold into different conformations in the presence of Na⁺ or K⁺ cations. In a Na⁺ solution, the 22-base telomeric sequence 5'-AGGG(TTAGGG)₃-3' forms a basket-type anti-parallel G-quadruplex as shown in Figure 7. However, in a buffer with K⁺ ions, the same sequence folds into multiple G-quadruplex conformations with comparable free energies. It has been reported that alternating between Na⁺ and K⁺ can trigger interconversion between anti-parallel and parallel G-quadruplexes. This agrees with one study reporting a minimal energy barrier between parallel and anti-parallel G-quadruplexes adopted by intramolecular sequences with TTA loops. Thus, the conformational diversity of G-
quadruplexes results from and is governed by the balance between the energetic contributions of a variety of structural and environmental factors and interactions.\textsuperscript{35}

Figure 1-7. Basket-type Anti-parallel G-quadruplex.

The presence of a mixture of conformations complicates thermodynamic and kinetic studies and interpretation of macroscopic results. For longer and more complex sequences, the polymorphic nature of G-quadruplexes becomes even more apparent. The 27 nucleotide long c-MYC parent sequence contains five tracts of three or more guanine bases which, depending on the combination of the guanines involved in G-quartet formation, may adopt structurally diverse G-quadruplex conformations.\textsuperscript{125} Mutations involving shortening the five runs of guanine bases to four or adding bases to elongate the loops may reduce the number of conformations available for sampling.\textsuperscript{125} Conformational conversions may also occur depending on molecules crowding conditions induced by PEG\textsuperscript{126} or the identity and concentration of divalent ions.\textsuperscript{127} Parallel conformations are favoured under molecular crowding conditions\textsuperscript{126} and when divalent cations are present at high concentrations.\textsuperscript{127}
1.6: Cations and Nucleic Acid Structure

DNA and RNA are highly charged molecules given the backbone rich in negatively charged phosphate groups. Association of two or more nucleic acid strands accompanying the formation of higher order structures causes a significant increase in charge density. High charge density attracts oppositely charged ions, making the discussion of ionic behavior important when examining the physical properties and the biological role of nucleic acids. Cations are especially important for G-quadruplexes because the structure cannot form in the absence of counterions specifically bound inside the central cavity. The ab initio calculations showed that specific cation binding in the central cavity contributes more to the stabilization of G-quadruplex structures than the hydrogen bonding or stacking interactions. Electrostatic repulsion caused by the gathering of O6 atoms of four guanine bases in a G-quartet and clustering of phosphate groups must be compensated by a neutralizing effect of counterions. It is therefore intuitively clear that folding, stability, and conformational preferences of G-quadruplexes are strongly modulated by the type and concentration of the cations present in solution.

1.6.1 DNA as a polyelectrolyte

Polyelectrolyte is a linear polymer with ionizable units along the backbone. Nucleic acids in solution are natural polyelectrolytes given the abundance of negatively charged phosphates with the resulting high charge density. The electrostatic interactions between counterions and nucleic acids modulate the ion distribution within the solution, thereby controlling the salt-dependent equilibria of conformations available for sampling. Stabilization of duplex DNA structures that are crucial for storing and expressing the hereditary information requires counterion condensation neutralizing the repulsive forces originating from the clustering of charged phosphate groups. The charge density of a duplex is higher compared to that of its isolated constituent single strands. Hence, the degree of counterion condensation around the double helix should be greater relative to the single-stranded conformation, causing counterion release upon the duplex-single strand transition. Since maintaining a cloud of counterions is entropically unfavourable with the entropic penalty increasing with a decrease in
the concentration of counterions in the bulk, the thermodynamic stability of duplex DNA increases with salt despite the fact that the local concentration of counterions in the vicinity of a DNA is essentially independent of that in the bulk\textsuperscript{84, 86, 87}. On a similar note, an increase in salt, generally, results in an increase in the thermal stability of G-quadruplexes\textsuperscript{87}. However, G-quadruplexes, in addition to the polyelectrolyte effect and the related cloud of counterions, bind cations in a specific manner inside their central cavity where they become coordinated to O6 oxygens of the guanines\textsuperscript{88, 89}.

One of the oldest theoretical tools that has been used to quantify the distribution of mobile ions around polyelectrolytes in solution is the Poisson-Boltzmann (PB) theory\textsuperscript{90, 91}. The PB theory links the electrostatic potential of a system and the charges that contribute to this potential\textsuperscript{80}. The basic Poisson-Boltzmann equation is the following:

$$\nabla^2 \psi(r) = \frac{-(4\pi/\varepsilon) \rho(r)}{\varepsilon}$$

(where $\psi(r)$ is the electrostatic potential at a distance $r$ in a medium with a dielectric constant $\varepsilon$ and a charge density $\rho$)\textsuperscript{92}. The charge density is given by the sum $q_i \Sigma z_i n_i(r)$ with the local concentration (ions/mL) of ions of type $i$, $n_i(r)$, and the ionic charge of $q_i z_i$. The Boltzmann equation relates the local concentration of a species to its bulk concentration $n_{bi}$ and energy $E_i(r)$ at $r$, which is determined by its charge and potential at $r$\textsuperscript{92}.

$$n_i(r) = n_{bi} e^{-E_i(r)/k_B T} = n_{bi} e^{-z_i q_i \psi(r)/k_B T}$$

Substituting equation (2) into equation (1) yields the final Poisson-Boltzmann equation\textsuperscript{92}. 
The PB equation is a mean-field theory with the mobile ions in the region far from the polyion surface affecting each other only through their average contribution to the mean-field potential. In other words, the PB theory does not take into consideration the ion-ion correlation. The PB equation is mathematically complicated because it is non-linear with the exponential term containing the electrostatic potential. In addition, the PB theory incorporates the Gouy-Chapman model where the polyelectrolyte is viewed as an infinitely long double-layered cylinder with condensed ions placed within the space between the inner and the outer layers. The mean charge density is measured by the number of charges present.

The counterion condensation theory described by Manning simplifies the mathematical complexity of the PB theory, although the two theories are based on the same basic principle. The counterions cluster around the polyelectrolyte, while the ions of the opposite charge are excluded from the vicinity. The number of accumulated counterions depends only on the valency of the counterion and the axial charge density of the polyelectrolyte, being independent of the concentration in the bulk solution. In the case of DNA, the theory predicts that cations are accumulated and anions are excluded and that the bulk concentration of ions does not affect the local concentration of the counterions in the close vicinity of the DNA. The counterions are free to change positions along the iso-potential energy paths created by the phosphates.

If a macromolecule can be modeled by a linear array of \( N \) univalent (negative) charges with an average of spacing \( b \), the extent of association of monovalent counterions \( M^+ \) with the polyanion can be described by a single dimensionless parameter \( \xi \).

\[
\xi = \frac{q^2}{\varepsilon kTb}
\]
Here, \( q \) signifies the protonic charge, \( \varepsilon \) is bulk dielectric constant of solvent, \( k \) is Boltzmann’s constant, \( T \) is the thermodynamic temperature, and \( b \) is average axial intercharge spacing along the helical axis for DNA or along the contour axis for more flexible polymers.

If only univalent ions are present, the \( \xi \) value greater than unity signifies that the molecule is highly charged. On the other hand, if the net charge density is less than one, counterions do not condense\(^7,84\). The counterion condensation theory, as proposed by Manning and Record, is useful in studying the ionic distribution in polyelectrolyte solutions. However, the basis of this model that presents the polyelectrolyte as a linear charge with no thickness and counterions as point charges has limitations in explaining the competitive binding of distinct counterion molecules\(^92\).

### 1.6.2 Salt Effects in G-quadruplexes

Among the major classes of nucleic acid structures, G-quadruplexes are unique in that they interact with cations in both specific and non-specific manners\(^97\). The specific binding occurs within the central cavity where mono- or divalent cations of an appropriate size form inner sphere coordination bonds with O6 oxygens of guanine bases\(^98,99\).

Arnott and colleagues (1974) have pioneered in suggesting that the O6 carboxyl oxygens of guanines form coordination bonds with cations\(^100\). Sequestered cations serve to weaken unfavourable interactions between carbonyl oxygens and consequently promote base-base stacking interactions\(^102,103\). In agreement with this notion, molecular dynamics simulations reveal that G-quadruplexes are destabilized in the absence of centrally-bound cations\(^104,105\).

The effect of specific internal cations was clearly demonstrated through studies observing the folding of G-quadruplexes in the atmosphere of Na\(^+\) and K\(^+\) ions\(^106\). The folding of the d(TG\(_4\)T) G-quadruplex increases by a factor of 200 following a 10-fold increase in the concentration of Na\(^+\) ions\(^35\). The physiological abundance and functionality of K\(^+\) and Na\(^+\) ions particularly make them important targets of characterization in G-quadruplex studies\(^35\).
Depending on the size, centrally bound cations may be located within the G-quartet plane or be sandwiched between the planes. Ions such as K⁺ and NH₄⁺ (with radii 1.33 and 1.48 Å, respectively) are too large to be coordinated in the plane of a G-quartet, whereas Na⁺ (with a radius 0.95 Å) fits within the plane of a G-quartet. In general, n-1 large cations are accommodated in n G-quartets. In this respect, Hud and collaborators used the [d(G₄T₄G₄)]₂ sequence to demonstrate that three NH₄⁺ ions are located within the four quartets. In contrast, the smaller Na⁺ ion allows for in-plane coordination. Multiple Na⁺ ions are not as restricted to the localized space as larger ions; they can move further away from each other to reduce electrostatic repulsions.

An early study aimed to characterize the role of cations in the stability of G-quadruplexes have produced the following order of the stabilizing action of cations K⁺ > Ca²⁺ > Na⁺ > Mg²⁺ > Li⁺ and K⁺ > Rb⁺ > Cs⁺. A subsequent study has revealed the following ranking of the stabilizing role of cations Sr²⁺ > Ba²⁺ > Ca²⁺ > Mg²⁺ and K⁺ > Rb⁺ > Na⁺ > Li⁺ = Cs⁺. In agreement with this result, Sr²⁺ and Ba²⁺ induce a thermodynamically more stable thrombin aptamer G-quadruplex than that induced by Mg²⁺ and Ca²⁺.

Cations do not always make G-quadruplexes stable. Certain counterions can also destabilize the structure. Even at low concentrations, divalent transition metal cations such as Mn²⁺, Co²⁺ or Ni²⁺ have been observed to disrupt K⁺-induced G-quadruplex structures. Some studies have also shown that the destabilizing effect of divalent cations depends on the concentration of monovalent ions. Hardin and his colleagues reported that divalent cations of concentration 10-40 mM stabilize the G-quadruplex conformation when the concentration of monovalent counterions is higher than 50 mM. However, when the monovalent cation concentration is lower than 50 mM, divalent ions destabilize G-quadruplex structures. Although Li⁺ is commonly suggested to destabilize G-quadruplexes, there are some indications that it neither stabilizes nor destabilizes G-quadruplex structures.
1.7: Energetics/Thermodynamics of Quadruplex Formation

G-quadruplexes have two or more G-tetrads stacked on top of another. Some researchers argue that G-quadruplex formation is accompanied by a decrease in free energy equal to or even exceeding that associated with Watson-Crick base pair formation\textsuperscript{128-131}. Other researchers, however, suggest that G-quadruplex stability depends stronger on the identity of cations, the length and sequence of loops, and the presence co-solvents\textsuperscript{132}. Kogut et al. (2016) have measured the energetics of dissociation of a single G-quartet from the telomeric G-quadruplex\textsuperscript{133}. The change in free energy results from an interplay between the favourable $\Delta H$ and unfavourable $-T\Delta S$ terms\textsuperscript{133}. A change in enthalpy, $\Delta H$, is around -35 kcal/mol for the tetrad formed in the telomeric sequence, whereas the value of $-T\Delta S$ is 29 kcal/mol\textsuperscript{133}. Thus, G-quartet formation and the subsequent stacking is an enthalpically driven event. This notion is in agreement with the results of Boncina et al. (2012) and Sengar et al. (2014) which reveal that G-quadruplex folding is enthalpy driven with an enthalpy change per G-quartet of -14 to -19 kcal/mol\textsuperscript{129,134}.

The van’t Hoff analysis of melting curves enables one to calculate changes in enthalpy of association or dissociation. The applicability of this analysis in intermolecular G-quadruplexes is based on the two-state assumption\textsuperscript{35}. Nonetheless, it has been reported that the enthalpies of intermolecular G-quadruplexes are comparable to those of intramolecular structures\textsuperscript{35}. Calorimetric studies of the (TG\textsubscript{4}T\textsubscript{4}) tetramolecular G-quaruplex reveal a change in enthalpy of -20 kcal/mol per quartet\textsuperscript{106,135}.

<table>
<thead>
<tr>
<th>Element</th>
<th>$K^+$</th>
<th>$Na^+$</th>
<th>$NH_4^+$</th>
<th>$Rb^+$</th>
<th>$Cs^+$</th>
<th>$Li^+$</th>
<th>$Ca^{2+}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ionic radius</td>
<td>1.52</td>
<td>1.16 (0.97)</td>
<td>1.43\textsuperscript{71}</td>
<td>1.66</td>
<td>1.81</td>
<td>0.9</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 1-1. Ionic radii of G-quadruplex stabilizing cations\textsuperscript{35}.
In contrast to consistent values of enthalpy reported in literature, the reported changes in entropy depend strongly on the molecularity of G-quadruplexes and the loop length and identity\textsuperscript{35,132}.

1.8: Solvent Effects on the Stability of G-quadruplexes

In aqueous solutions, macromolecules are hydrated\textsuperscript{157}. Hydration refers to the binding of water molecules to solutes\textsuperscript{159}. It has been reported that hydration of some biopolymers is not limited to a single layer of water molecules but can involve two or more layers of water molecules\textsuperscript{158}. The presence of cosolvents modulates the solute-solvent interactions of biomolecules and changes the hydrogen-bond dynamics of water molecules\textsuperscript{136}. Nucleic acids reside and function in cellular environment where biomolecules occupy about $1/3$ of the whole volume\textsuperscript{137} with the macromolecular concentration ranging from 50 to 400 g/L\textsuperscript{138}. DNA molecules including G-quadruplexes are surrounded by water molecules forming a hydrogen bonding network as illustrated in Figure 8.
The stability of DNA and RNA duplexes and triplexes involving Watson-Crick hydrogen bonds decreases with an increase in the concentration of crowding agents, but the effect of these agents on the stability of G-quadruplexes is opposite\textsuperscript{139}. Sugimoto and coworkers have reported that cosolvents can induce transitions between G-quadruplex conformations\textsuperscript{141}. In particular, polyethylene glycol (PEG) triggers a transition of d(G\textsubscript{4}T\textsubscript{4}G\textsubscript{4}) from the antiparallel to a more stable parallel conformation\textsuperscript{141}. A similar result has been reported by Li et al. (2005)\textsuperscript{142}. These findings suggest that the energetics of G-quadruplex polymorphism is modulated not only by nucleotide sequence but also by environmental factors including solvent composition.

The antiparallel G-quadruplex formed by a 15-base thrombin aptamer is stabilized under the molecular crowding conditions at various concentrations of PEG 200\textsuperscript{143}. The stability of the G-quadruplex decreases linearly as the activity of water increases\textsuperscript{143}. The number of water molecules released per nucleotide upon formation of an anti-parallel G-quadruplex has been estimated to be 4.5\textsuperscript{143}. The release of water molecules upon G-quadruplex formation suggests
that the folded conformation is less hydrated than the unfolded form. In contrast, according to osmotic stress experimental results, the folded state of a DNA duplex is more hydrated than the unfolded state\textsuperscript{143}. The number of water molecules taken up following the formation of a duplex is 3.4 per nucleotide\textsuperscript{143}. Thus, under molecular crowding conditions, non-canonical DNA structures like G-quadruplexes may be favoured over duplexes.

### 1.9: i-motif Nucleic Acid Structures

In the genomic DNA, G-quadruplex-forming guanine-rich strands are complementary to strands rich in cytosine, with an exemption of the telomeric single-stranded 3’-overhang. In 1993, Gehring and his colleagues published the first report of the unique structure adopted by these cytosine-rich sequences (Figure 9)\textsuperscript{145}. These tetraplex structures consist of two parallel duplexes are intercalated in an antiparallel manner. They are, generally, formed at slightly acidic conditions facilitating hemi-protonation of cytosines and formation of cytosine-cytosine\textsuperscript{+} base pairs (Figure 9).

![Figure 1-9. Cytosine-Cytosine\textsuperscript{+} Base Pair\textsuperscript{145}](image)
There are two topological forms of i-motifs. In one form (3'E), the cytosine-cytosine base pair is located at the 3’ end, while in the other form (5'E), the base-pairing occurs at the 5’ end. They are named 3'-E (R-form) and 5'-E (S-form). Some sequences show a tendency to interconvert between the two forms consistent with the picture in which the energy barrier between the two configurations is not large. The third less common topology is the T-form where the base pair at both ends are not intercalated. Analogous to G-quadruplexes, i-motifs may be inter- or intramolecular.

Although i-motifs generally exist under slightly acidic conditions, they also can form under neutral or even basic conditions. In particular, experiments conducted under conditions mimicking the crowded environment of the cell reveal that i-motifs can form at neutral pH. Negative superhelicity has also been shown to facilitate the i-motif formation. In one study by Sun and Hurley (2009), the c-MYC oncogene promoter sequence has been placed into a supercoiled plasmid. Importantly, the G-rich and C-rich strands of the sequence have been shown to fold into the G-quadruplex and i-motif conformations, respectively. These authors have used enzymatic and chemical footprinting to show that negative superhelicity supports the formation of tetraplex structures even at neutral pH. Paradoxically, a more recent paper has refuted this finding and argued that supercoiling itself is not a sufficient condition to induce G-quadruplex formation in other sequences. Thus, further studies are required to clarify the influence of supercoiling on tetraplex structure formation in the genomic DNA.

1.10: Methods to Study G-quadruplex Structures

A basic structural method of studying G-quadruplex structures is X-ray crystallography. This technique involves crystallization of purified DNA sample. By analyzing the diffraction patterns of the X-ray shone on the crystal structure, one can determine the three-dimensional image of the DNA of interest. One disadvantage of X-ray crystallography is that only the solid state of a DNA structure can be detected. The technique is of limited applicability when the structure is polymorphic, which is a common case for G-quadruplexes. X-ray crystallography will detect and characterize only the conformation most favoured by crystallization.
Another technique that provides detailed structural data about G-quadruplex structures is nuclear magnetic resonance (NMR) spectroscopy\textsuperscript{161}. Though high purification is needed, this method demands less work for sample preparation compared to X-ray crystallography\textsuperscript{161}. NMR is useful for G-quadruplexes because guanine imino protons show a unique shift when they form hydrogen bonding\textsuperscript{161}. In addition, NMR provides information about the polymorphism because it can be used for dynamics and kinetics studies\textsuperscript{161}.

Native gel electrophoresis probes the anomalous migratory behaviour of folded G-quadruplexes relative to the unfolded DNA\textsuperscript{150}. Owing to their compact structure, intramolecular G-quadruplexes migrate faster through a cation-containing gel than their linear counterparts\textsuperscript{151}. To characterize other possible structures present in the bands, one must pay a careful attention to the speed (voltage) of the run\textsuperscript{151}. DNA can be visualized through labeling or EtBr staining\textsuperscript{35}.

Circular dichroism (CD) spectroscopy has been traditionally used as a tool that discriminates between various G-quadruplex conformations\textsuperscript{35}. This method involves a sample exposed to circularly polarized light. If there is a chiral center in the molecule of interest, it will generally interact asymmetrically with the light, with the asymmetry varying with the wavelength. Although it may be difficult to unequivocally link a CD spectrum to a specific DNA structure, characteristic spectra corresponding to different G-quadruplex conformations have been determined empirically\textsuperscript{35}. If used over a range of temperatures, CD spectroscopy can be used to observe a melting transition of a G-quadruplex and hence determine thermodynamic parameters such as the Tm, ΔH, and ΔG used to quantify structural stabilities\textsuperscript{35}. UV spectroscopy is a technique that can monitor the dissociation/association and folding/unfolding reactions of DNA structures\textsuperscript{35}. For a duplex DNA, a hyperchromic shift is indicative of the unstacking of bases and unfolding of the structure\textsuperscript{35}. This method has been applied to G-quadruplexes with UV absorption changes monitored at 295nm\textsuperscript{35,121}. The Tm, the temperature at which 50% of the structures is unfolded, can be determined from the melting profile and additional thermodynamic data can be obtained based on the two-state approximation of the melting curve\textsuperscript{35,154}. The temperature dependence of the equilibrium constant of the reaction yields the van’t Hoff enthalpy of the reaction.
Differential scanning calorimetry (DSC) is a fundamental method for examining the thermal stability of G-quadruplexes. DSC provides a thermogram in which the apparent heat capacity is measured as a function of temperature. This technique allows a direct measurement of the enthalpy change resulting from heat-induced unfolding of DNA. The model-independent enthalpy obtained from DSC can be compared to the model-dependent van’t Hoff enthalpy to determine if the transition contains intermediates. If the two enthalpies coincide, the transition proceeds as a two-state event. However, if the values are different, the transition involves intermediates or aggregation takes place.

A powerful method for detecting and characterizing the G-quadruplex configuration and dynamics is single-molecule Förster resonance energy transfer (FRET). In FRET, oligonucleotides are labeled with a donor (e.g., 6-carboxy fluorescein FAM at 5’ end) and an acceptor (e.g., tetramethyl-6-carboxyrhodamine TMA at the 3’ end) fluorophores. Upon DNA folding, the donor fluorophore transfers its energy to the acceptor, with an efficiency that depends on their distance and relative orientation. When the donor and the acceptor are close, the fluorescence is quenched, but when apart, the molecule becomes detectable with fluorescence. G-quadruplex melting proceeds with the separation of the two fluorophores with an increase in temperature. This technique is most appropriate for elongated sequences that form intramolecular G-quadruplexes with the folded conformation bringing the two ends into close proximity.

1.11: Gap in Literature and the Experimental Layout

The importance of G-quadruplexes as therapeutic targets continually increases with their in vivo existence and important functionalities being revealed on a virtually, daily basis. Hence, the controlled modulation of G-quadruplex structures and the search of drugs triggering G-quadruplex formation being of high interest to researchers. Enhancing or inhibiting the formation of G-quadruplexes at specific sites within the genome is a promising approach to control expression of proto-oncogenes and control and prevent the immortality of tumour cells. Though there are many factors that contribute to stabilization of G-quadruplexes, one universal component that exists only in G-quadruplexes is the specific ion binding within the central cavity.
The thermal stability of G-quadruplexes have been shown to strongly depend on the concentration of salt\textsuperscript{35}. However, prior to our work, the effects of specific and non-specific cations have not been separated. Ionic effects are more complex in G-quadruplexes, since there are two types of ion binding. This thesis aims at resolving the differential effect of specific and non-specific ionic binding on the stability of G-quadruplexes. Our laboratory has devised an experimental approach to separate the stabilizing effect of centrally bound cations from that of the cloud of counterions. The method involves the initial saturation of the G-quadruplex by the stabilizing ions (Na\textsuperscript{+} or K\textsuperscript{+}) followed by the addition of either a stabilizing or a non-stabilizing cation (Li\textsuperscript{+}, Cs\textsuperscript{+}, TMA\textsuperscript{+}). This methodology, which started with telomeric sequences, has been subsequently applied to sequences derived from the promoter regions of the c-Myc and VEGF oncogenes.


1.12 Reference


CHAPTER 2: Polyelectrolyte Effects in G-quadruplexes

"Reprinted from Biophysical Chemistry, 184, Byul G. Kim, Yuen Lai Shek, and Tigran V. Chalikian. Polyelectrolyte Effects in G-quadruplexes, 95-100. Copyright (2013), with permission from Elsevier B.V. ".

Author’s Contribution: Byul G. Kim performed all experiments, analyzed the data, and participated in the preparation of the manuscript.
2.1 Abstract

The role of counterion condensation as a dominant force governing the stability of DNA duplexes and triplexes is well established. In contrast, the effect of counterion condensation on the stability of G-quadruplex conformations is poorly understood. Unlike other ordered nucleic acid structures, G-quadruplexes exhibit a specific binding of counterions (typically, Na\(^+\) or K\(^+\)) which are buried inside the central cavity and coordinated to the O6 carbonyls of the guanines forming the G-quartets. While it has been known that the G-quadruplex-to-coil transition temperature, \(T_M\), increases with an increase in the concentration of the stabilizing ion, the contributions of the specific (coordination in the central cavity) and nonspecific (condensation) ion binding have not been resolved. In this work, we separate the two contributions by studying the change in \(T_M\) of preformed G-quadruplexes following the addition of nonstabilizing ions Li\(^+\), Cs\(^+\), and TMA\(^+\) (tetramethylammonium). In our studies, we used two G-quadruplexes formed by the human telomeric sequences which are distinct with respect to the folding topology and the identity and the number of sequestered stabilizing ions. Our data suggest that the predominant ionic contribution to G-quadruplex stability comes from the specifically bound Na\(^+\) or K\(^+\) ions and not from counterion condensation. We offer molecular rationalizations to the observed insensitivity of G-quadruplex stability to counterion condensation and emphasize the need to expand such studies to assess the generality of our findings.
2.2 Introduction

Although the genomic DNA overwhelmingly exists in the B-double helical conformation, many genomic sequences may give rise to stable non B-DNA conformations. For example, guanine-rich single-stranded stretches are susceptible to spontaneous dissociation from their complementary strands and folding into various G-quadruplex structures in which guanine bases associate with each other in G-quartets, stable hydrogen-bonded arrangements\textsuperscript{1-3}. The biomedical relevance of G-quadruplexes stems from the fact that the consensus G-quadruplex sequence motifs are found in critical loci of the genome, including telomeres, centromeres, immunoglobulin switch regions, mutation-prone hot spots, and promoter regions of many oncogenes\textsuperscript{2,4-8}. Telomeric sequences have become an important target for anti-cancer drug development, since drug-induced G-quadruplex structures are incompatible with telomerase function and induce rapid senescence in cancer cells\textsuperscript{9}. Search for G-quadruplex-stabilizing molecules have quickly emerged as part of novel anti-cancer strategies\textsuperscript{6, 8}. The ultimate success of such strategies depends, in part, on our ability to understand thermodynamic forces stabilizing G-quadruplex conformations relative to their respective double- and single-stranded conformations and modulation of these forces by environmental conditions and the presence of G-quadruplex-stabilizing agents (drugs).

Owing to their polyelectrolyte nature, DNA and RNA are surrounded by a cloud of mobile counterions which exert a powerful influence on the stability of nucleic acid structures\textsuperscript{10-15}. Counterion condensation in the vicinity of DNA and the concomitant entropic penalty introduce a thermodynamic leverage that can produce shift in an equilibrium of virtually any helix-to-coil and ligand binding reaction simply by a change in salt\textsuperscript{15}. Helical forms of DNA, due to their higher axial charge density, $\xi = e^2/(e k_B T b)$ (where $e$ is electron charge, $\varepsilon$ is the dielectric constant of the solvent, $k_B$ is the Boltzmann constant, $T$ is the thermodynamic temperature, and $b$ is the average axial distance between the polyelectrolyte charges), condense more counterions per phosphate relative to the single-stranded conformation\textsuperscript{10,14-17}. Consequently, helix-to-coil transitions of nucleic acids are accompanied by a release of counterions to the bulk, while an increase in salt results in an increase in the thermal and thermodynamic stability of the helical forms due to a diminution of the mixing entropy penalty\textsuperscript{10,13-16}. At low ionic strength, the
number of associated counterions per phosphate is given by $\theta = 1 - \xi^{-1}$. For example, the values of $\theta$ for a double-stranded B-DNA and a single-stranded DNA are 0.76 and 0.44, respectively$^{10}$. Thus, the helix-to-coil transition of a B-DNA is accompanied by a release of 0.32 monovalent cations to the bulk.

The conventional way of experimental determination of the number of cations, $\Delta n = \theta_h - \theta_c$, released to the bulk upon a helix-to-coil transition of a DNA relies on measuring the salt dependence of the transition temperature, $T_M$:

$$\Delta n_{M^+} = (\Delta H_M/RT_M^2)(\delta T_M/\Delta \ln[M^+]) \quad (1)$$

where $[M^+]$ is the molar concentration of the counterion$^{11,16,18,19}$.

In contrast to all other ordered nucleic acid structures, G-quadruplexes, in addition to non-specific binding (condensation), exhibit a specific binding of counterions (typically, Na$^+$ or K$^+$) which are buried inside the central cavity and coordinated to the O6 carboxyls of the guanines forming the G-quartets$^{3,5,6}$. Hence, application of Eq. (1) to G-quadruplexes melting transitions, yields the number of cations released to the bulk from both the central cavity (bound specifically) and the cloud of counterions (bound nonspecifically). The problem of discrimination between the two cation populations remains unsolved which prevents one from elucidating the role of the polyelectrolyte effect in stabilization of G-quadruplex structures$^3$. In one study, it has been concluded that the specific cation binding plays a more prominent role in stabilizing a G-quadruplex than counterion condensation$^{20}$. This conclusion is in agreement with our results presented below. However, it has been reached based on comparative G-quadruplex stability studies in binary solvents containing water and a variety of cosolvents. In particular, it has been assumed that the influence of a particular cosolvent on G-quadruplex stability is predominantly determined by the dielectric constant of the medium. However, such an oversimplification may not be warranted given the complexity of the effect of cosolvents on the stability of biopolymers$^{21-23}$. 
In this work, we evaluate the influence of the polyelectrolyte effect on G-quadruplex stability by separating the effect of the stabilizing cations in the central cavity from the effect of the cations condensed in the vicinity of the DNA. More specifically, after saturating a G-quadruplex by the stabilizing cation (Na\(^+\) or K\(^+\)), a further increase in the solution ionic strength is accomplished by the addition of either the stabilizing or a non-stabilizing cation [Li\(^+\), Cs\(^+\), TMA\(^+\) (tetramethylammonium)]. The latter cannot penetrate the central cavity being either too large or too small and, therefore, influence the G-quadruplex stability only via the effect of counterion condensation. We study the differential effect of salt on the G-quadruplex-to-single strand transition temperature, \(T_M\), when the ionic strength is modulated by the stabilizing or non-stabilizing cations. To increase the generality of our work and the ensuing conclusions, we have chosen two G-quadruplexes varying in topology and the identity and number of the sequestered stabilizing cation (Na\(^+\) and K\(^+\)). Our results suggest that, in contrast to the duplex and T-rich triplex conformations, the polyelectrolyte effect plays a small to negligible role in maintaining G-quadruplex stability. The observed salt-dependent increase in G-quadruplex stability predominantly reflects the specific binding of the Na\(^+\) or K\(^+\) cations.

### 2.3 Materials and Methods

#### Materials

The oligodeoxyribonucleotides, d[A(G\(_3\)T\(_2\)A)\(_3\)G\(_3\)] (Tel22) and d[A\(_3\)(G\(_3\)T\(_2\)A)\(_3\)G\(_3\)A\(_2\)] (Tel26), containing four repeats of the human telomeric DNA sequence were synthesized and cartridge purified by ACGT, Inc. (Toronto, ON, Canada). Potassium chloride, sodium chloride, cesium chloride, lithium chloride, tetramethylammonium chloride, and phosphoric acid were purchased from Sigma-Aldrich Canada (Oakville, ON, Canada). EDTA (free acid) was purchased from Fisher Biotech (Fair Lawn, NJ, USA). These reagents were of the highest grade commercially available and used without further purification. All solutions were prepared using doubly distilled water.
Prior to all experiments, the Tel22 oligonucleotide was dissolved in and exhaustively dialyzed against a pH 7.0 buffer consisting of 10 mM sodium phosphate, 0.1 mM EDTA, and 20 mM NaCl. The Tel26 oligonucleotide was dissolved in and dialyzed against a pH 7.0 buffer consisting of 10 mM potassium phosphate, 0.1 mM EDTA, and 20 mM KCl. In these conditions, both Tel22 and Tel26 are fully folded in their respective G-quadruplex conformations [24,25]. Dialysis was carried out in 1000 Da molecular weight cut-off Tube-O-Dialyzers from G Biosciences (St. Louis, MO, USA). For UV melting and CD experiments, the ionic strengths of the DNA solutions were adjusted to the desired level by addition of aliquots of NaCl, KCl, LiCl, CsCl, or TMACl on top of the existing 20 mM of NaCl (Tel22) or KCl (Tel26).

The CD and UV melting experiments were performed at the following ionic conditions. For the experiments involving Tel22, the ionic conditions were 20, 50, 100, 200, 400, and 800 mM NaCl or 20 mM NaCl plus 30, 80, 180, 380, and 780 mM LiCl, CsCl, or TMACl. For the experiments involving Tel26, the measurements were performed at 20, 50, 100, 200, 400, and 800 mM KCl or 20 mM KCl plus 30, 80, 180, 380, and 780 mM LiCl, CsCl, or TMACl.

The concentrations of the oligonucleotides were determined from the UV light absorbance at 260 nm measured at 25 °C with a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, Ontario, Canada) using molar extinction coefficients of 228,500 and 278,200 M$^{-1}$ cm$^{-1}$ for the unfolded conformation of Tel22 and tel26, respectively. These extinction coefficients were calculated using an additive nearest neighbor procedure as described by Dr. Richard Owczarzy (http://www.owczarzy.net/extinct.htm). For our CD measurements and temperature-dependent UV light absorption measurements, the DNA concentrations were ~30 and ~3 μM, respectively.

**Circular Dichroism Spectroscopy**

The CD spectra of the Tel22 and Tel26 were recorded in a 1 mm path-length cuvette at 25 °C using an Aviv model 62 DS spectropolarimeter (Aviv Associates, Lakewood, NJ). CD spectroscopy was used to probe the conformations adopted by the oligonucleotides at the various ionic conditions employed in the present study.
UV Melting Experiments

UV light absorption at 295 nm was measured as a function of temperature in DNA samples contained in a 1 cm path-length cuvette. These measurements were performed by a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, Ontario, Canada). The temperature was changed at a rate of 1 °C per minute. The G-quadruplex-to-coil transition temperatures, $T_M$, and van't Hoff enthalpies, $\Delta H_{vh}$, were evaluated from our experimental UV melting profiles using standard procedures\textsuperscript{26,27}.

It should be noted that temperature scanning experiments performed on Tel22 in NaCl solutions at a rate of 0.2 °C per minute did not reveal any change in the shape of the melting profiles. By extension, we assume that the scanning rate of 1 °C per minute is optimal for all DNA melting experiments reported here.

2.4 Results

CD spectra. In the presence of Na$^+$ ions, the Tel22 human telomeric sequence, d[A(G3T2A)3G3], forms an antiparallel structure with three Na$^+$ ions coordinated within the central cavity\textsuperscript{28}. To monitor the G-quadruplex formation in the presence of NaCl and its maintenance throughout our experimental protocols (in the presence of LiCl, CsCl, and TMACl), we measured the CD spectrum of the DNA at each salt. Figure 2-1 presents the CD spectra of Tel22 in the presence of 20 mM NaCl and various concentrations of LiCl (panel A), CsCl (panel B), and TMACl (panel C). Previous investigations have shown that the CD spectrum of antiparallel telomeric G-quadruplexes display a negative peak at 260 nm and a positive peak at 295 nm\textsuperscript{3}. The CD spectra of Tel22 in Figure 1a, b, and c with the characteristic bands at 260 and 295 nm are consistent with the formation of an antiparallel G-quadruplex structure.
In the presence of K\(^+\) ions, the Tel26 sequence d[A\(_3\)G\(_3\)(T\(_2\)AG\(_3\))\(_3\)A\(_2\)] overwhelmingly exists as a hybrid-type (hybrid-1) intramolecular G-quadruplex consisting of three G-tetrads linked with mixed parallel/antiparallel G-strands\(^{29-31}\). The structure has a sequential double-chain-reversal side loop and two lateral loops each consisting of three nucleotides (-TTA-) thereby making the third G-strand antiparallel. In contrast to Tel22 with three sequestered Na\(^+\) ions, the Tel26 G-quadruplex coordinates only two K\(^+\) ions within its central cavity. The CD signature of the hybrid-1 conformation of Tel26 is a double-headed positive domain with the maxima at 260 and 295 nm and a negative domain with the minimum at 240 nm\(^{29}\). The CD spectra of Tel26 in the presence of 20 mM KCl and various concentrations of LiCl (Figure 2-2a), CsCl (Figure 2-2b), and TMACl (Figure 2-2c) display the characteristic bands at 240, 260, and 295 nm. Thus, the CD spectra of Tel26 are consistent with the CD spectrum reported for the hybrid-1 conformation of Tel26\(^{25,29}\). The CD spectra are representative of more than three repeated trials of equivalent experiments.
Chapter 2: Polyelectrolyte Effects in G-quadruplexes

Figure 2-2. CD spectra of Tel26 in the presence of 20 mM KCl and 0 (●), 180 (○), 380 (■), and 780 (□) mM LiCl (panel A), CsCl (panel B), and TMACl (panel C).

UV Melting Profiles. Figure 2-3 shows representative UV melting profiles measured at 295 nm for Tel22 in 100 mM NaCl and Tel26 in 100 mM KCl. These and similar profiles determined at other salt conditions all exhibit a sigmoidal shape. The UV melting profiles were approximated by the analytical relationship for the two-state transition:

\[ \alpha = \left[ 1 - \exp \left( \frac{\Delta H_vH}{T - T_M - 1/R} \right) \right]^{-1} \]  

(2)

where \( \Delta H_vH \) is the van’t Hoff enthalpy of the transition, \( T_M \) is the transition temperature, \( \alpha = [A(T) - A_N(T)] / [A_D(T) - A_N(T)] \) is the fraction of the DNA that is unfolded; \( A(T) \) is the UV light absorption at a temperature \( T \); and \( A_N(T) \) and \( A_D(T) \) signify the native and denatured baselines, respectively.
Our measured UV melting profiles were used in conjunction with Eq. (2) to determine the transition temperatures, \( T_M \), and enthalpies, \( \Delta H_v \), for Tel22 and Tel26 at each experimental salt condition. Figure 4 plots the melting temperatures, \( T_M \), against the natural logarithm of cation concentration for Tel22 (panel A) and Tel26 (panel B).

It needs to be pointed out that G-quadruplexes, in general, may not melt in the two-state manner with no single temperature to describe the melting transition\(^{32,33}\). Therefore, the \( T_M \), as determined from Eq. (2), should be qualified as an effective or apparent melting temperature. Nevertheless, it represents a reliable measure of the thermal stability of G-quadruplexes in comparative studies reported here.
2.5 Discussion

Counterions can influence the stability of a G-quadruplex in two ways by binding specifically or binding nonspecifically by condensing around the DNA polyanion. While specific binding may cause conformational transitions of G-quadruplex-forming sequences, no conformational changes are expected to accompany nonspecific ion binding. Any release of condensed counterions upon a helix-to-coil transition reflects a difference in charge density between the folded and unfolded forms. In the present study, we are interested in this type of binding (counterion condensation).

The Na\(^+\) and K\(^+\) ions bind specifically by penetrating the central cavity and stabilizing the G-quadruplex conformation\(^3\). In contrast, the Li\(^+\), Cs\(^+\), and TMA\(^+\) ions do not fit inside the cavity being either too small or too large. To additionally verify that these ions do not support G-quadruplex formation, we have recorded the CD spectra of Tel22 and Tel26 in the absence of the Na\(^+\) and K\(^+\) ions but in the presence of Li\(^+\), Cs\(^+\), and TMA\(^+\) up to 800 mM (data not shown). No G-quadruplex formation was detected.

Inspection of Figures 1 and 2 reveals that, in the presence of Li\(^+\), Cs\(^+\), and TMA\(^+\), the CD spectra of the preformed Tel22 and Tel26 G-quadruplexes undergo only quantitative, although, sizeable changes. The observed changes do not involve appearance of new CD bands or disappearance of the existing ones. Instead, the Li\(^+\), Cs\(^+\), and TMA\(^+\) ions only bring about a diminution of the intensity of the existing CD bands. This observation is consistent with the picture in which the two G-quadruplexes retain their respective global conformations with no new species (including unfolded species) appearing in the presence of the three nonstabilizing cations. Thus, we conclude that the Li\(^+\), Cs\(^+\), and TMA\(^+\) ions affect the thermal stability of the Tel22 and Tel26 G-quadruplexes predominantly via the effect of counterion condensation.

A note of caution is in order here. TMA\(^+\) ions interact in a differential manner with the AT and GC base pairs in DNA duplexes\(^34-38\). Hence, the effect of TMA\(^+\) on G-quadruplex stability may not be limited to counterion condensation. While not ignoring this possibility, we, nevertheless, do not separate the effect TMA\(^+\) ions from that of the Li\(^+\) and Cs\(^+\) ions because of the qualitative similarity of the influence of the three ions on the spectral and thermal properties.
of the G-quadruplexes. In fact, the CD spectral changes of Tel22 and Tel26 caused by TMA$^+$ are less pronounced compared to those caused by Li$^+$ and Cs$^+$ (see Figures 1 and 2).

Inspection of Figures 4a and b reveals that steep increases in $T_M$ for both the Tel22 and Tel26 oligonucleotides accompany increases in the concentration of NaCl and KCl, respectively. In other words, the thermal stability of the two G-quadruplexes increases with the rise in the concentration of their respective stabilizing ions (Na$^+$ for Tel22 and K$^+$ for Tel26). We use the slope ($\Delta T_M/\Delta \ln[M^+]$) in conjunction with Eq. (1) to calculate the number of sodium or potassium ions, $\Delta n_{M^+}$, released to the bulk upon G-quadruplex unfolding$^{11,18}$. The numbers of counterions released to the bulk are 1.7±0.5 Na$^+$ for Tel22 and 1.5±0.2 K$^+$ for Tel26. These numbers are qualitatively similar to 3 and 2, the structurally detected numbers of Na$^+$ and K$^+$ cations coordinated within the central cavities of the Tel22 and Tel26 G-quadruplexes, respectively$^{28,29,39}$. 

![Graph](image)
Further inspection of Figures 4a and b reveals that an increase in the concentration of the nonstabilizing ions Li$^+$, Cs$^+$, and TMA$^+$ brings about only modest (Tel26) or no (Tel22) changes in the melting temperature, $T_M$. Specifically, for Tel22, an increase in the concentration of these ions from 0 to 800 mM does not result in any increase in $T_M$. For Tel26, an increase in the concentration of the Li$^+$, Cs$^+$, and TMA$^+$ ions causes modest, although non-zero, changes in $T_M$. The strongest change is observed for Li$^+$ ions for which an increase in concentration from 0 to 800 mM causes $T_M$ to increase from 46 to 58 °C. This is only a fraction of the increase in $T_M$ from 46 to 77 °C that was detected to accompany an increase in K$^+$ concentration from 20 to 800 mM.

In the absence of fortuitous compensations, the observed insensitivity of the thermal stability of Tel22 with respect to the presence of the nonstabilizing ions Li$^+$, Cs$^+$, and TMA$^+$ is consistent with the picture in which the polyelectrolyte effects exert a negligible influence on the stability.
of the Tel22 G-quadruplex. By extension, this inference implies that the Tel22 oligomeric sequence condenses the same amount of counterions in its folded and unfolded states. It also suggests that the observed increase in the stability of the Tel22 G-quadruplexes upon the addition of NaCl results overwhelmingly from the specific binding of Na\(^+\) ions in the central cavity with the former acting as a ligand. From this perspective, the observed increase in G-quadruplex stability due to Na\(^+\) ions is similar to an enhancement of protein or DNA stability facilitated by ligand binding.

In contrast to Tel22, the thermal stability of the Tel26 G-quadruplex moderately increases with an increase in the concentration of the nonstabilizing ions Li\(^+\), Cs\(^+\), and TMA\(^+\). As is seen from Figure 4b, the increase in \(T_M\) is only 15 to 30\% of that caused by the stabilizing ion K\(^+\). Thus, although, the effect of counterion condensation in Tel26 is not zero, the predominant ionic contribution to its stability still comes from the specific K\(^+\) binding and not from counterion condensation.

The greatly reduced effect of counterion condensation in G-quadruplex stability is related, at least, in part to the presence of coordinated cations within the central cavity with the concomitant decrease in the charge density of the DNA. This notion is supported by the fact that Tel22 with its three sequestered Na\(^+\) ions exhibits no change in \(T_M\) with an increase in the concentration of nonstabilizing cations. On the other hand, Tel26 with only two sequestered K\(^+\) ions and, hence, a higher charge density, exhibits a non-zero slope of the dependence of \(T_M\) on the concentration of nonstabilizing cations. This interpretation is also in line with the lack of counterion condensation contribution to the stability of C-rich triplexes which, due to the presence of protonated cytosines in the third strand, exhibit a reduced charge density\(^{40, 41}\). Another contributing factor may be the globular shape of a G-quadruplex that increases the distance between the G-strands thereby diminishing the net charge density to the level characteristic of the single-stranded conformation. This conjecture is supported by the results of recent theoretical studies by Manning that revealed reduced counterion condensation around a charged spherical construct relative to a cylinder\(^{12, 13}\).

The results presented in this paper suggest that, for the two G-quadruplexes differing in topology and the type of the stabilizing cations, the polyelectrolyte effects play a small to
negligible role in maintaining their stability. This is in stark contrast to the thermodynamics of duplex and T-rich triplex DNA where counterion condensation provides a major contribution to thermal and thermodynamic stability\textsuperscript{10, 11, 14, 15, 18, 41}. Our data suggest that the duplex-G-quadruplex equilibrium in a system with a G-rich strand and its complementary C-rich strand will shift towards the duplex state upon an increase in the concentration of nonstabilizing counterions. This notion may be of practical importance, for example, for understanding the balance of forces governing the conformational preferences of G-rich strands in the promoter regions of oncogenes where G-quadruplex formation is preceded by duplex dissociation.

The generality of our finding is supported by the ionic, structural, and topological dissimilarities between the two G-quadruplex structures studied in this work. However, further studies involving a larger and more diversified group of G-quadruplexes are needed to ascertain the generality of the relative insignificance of counterion condensation in maintaining G-quadruplex stability. In particular, the influence of the G-quadruplex topology, molecularity, sequence, loop length, and composition must be studied.

2.6 Conclusion

G-quadruplexes are non-canonical nucleic acid constructs that exist in the genome and have been implicated in cancer research. G-quadruplexes are unusual in that, in contrast to other DNA structures, they specifically bind cations in their central cavity. There has been a long-standing question about the ionic stabilization of G-quadruplexes. It was not clear if the stabilization comes from the specific binding, nonspecific binding (counterion condensation), or both. In this work, we resolve the specific and nonspecific contributions to the thermal stability of G-quadruplexes by separating the effect of the centrally bound ions from the effect of the cations condensed around the DNA. Our data suggest that the predominant ionic contribution to G-quadruplex stability comes from the specifically bound Na\textsuperscript{+} or K\textsuperscript{+} ions and not from counterion condensation around the DNA. Our results shed light on the balance of forces governing the conformational preferences of guanine-rich sequences while also being of practical importance for controlled induction of G-quadruplexes in the genome.
2.7 Acknowledgements

This work is supported by a grant from NSERC to TVC.
2.8 Reference


11. M.T. Record, C.F. Anderson and T.M. Lohman, Thermodynamic analysis of ion effects on binding and conformational equilibria of proteins and nucleic acids. Roles of ion association or release, screening, and ion effects on water wctivity, Q.Rev.Biophys. 11 (1978) 103-178.


CHAPTER 3: Effects of Salt on the Stability of a G-quadruplex from the Human c-MYC Promoter


Author’s Contribution: Byul G. Kim performed all experiments, analyzed the data, and participated in the preparation of the manuscript.
3.1 Abstract

In the presence of potassium ions, a modified c-MYC NHE III $I_1$ sequence with two G-to-T mutations (MYC22-G14T/G23T) forms a highly thermostable parallel-stranded G-quadruplex. The G-quadruplex exhibits a steady increase in its melting temperature, $T_M$, upon an increase in the concentration of the stabilizing cation $K^+$. On the other hand, an increase in the concentration of nonstabilizing $Cs^+$ or $TMA^+$ cations at a constant concentration of $K^+$ causes a sharp decline in $T_M$ followed by leveling off at ~200 mM $Cs^+$ or $TMA^+$. At 51 °C and 600 mM $K^+$, an increase in $Cs^+$ concentration from 0 to 800 mM leads to a complete unfolding of the G-quadruplex. These observations are consistent with the picture in which the unfolded state of MYC22-G14T/G23T exhibits a stronger polyelectrolyte effect than the G-quadruplex state. We estimate that the unfolded state condenses roughly one extra counterion relative to the G-quadruplex state. Taken together with results of our earlier study of human telomeric G-quadruplexes, our data suggest that the stabilizing action of cations on the G-quadruplex conformation is, primarily, due to the central ions which act as specifically bound ligands. Nonspecifically bound (condensed) counterions may slightly stabilize the G-quadruplex conformation or exert no influence on its stability (human telomeric G-quadruplexes) or may strongly destabilize it (MYC22-G14T/G23T). We offer a structural rationalization for the extraordinary thermal stability and reduced polyelectrolyte effect of the MYC22-G14T/G23T G-quadruplex.
3.2 Introduction

Guanine-rich single-stranded DNA stretches exhibit a propensity to spontaneously dissociate from their complementary strands and fold into various G-quadruplex structures in which guanine bases associate into G-quartets - plain hydrogen-bonded arrangements \(^1\)-\(^3\). More generally, the \(G_mL_1G_mL_2G_mL_3G_m\) sequence motifs (\(m\) refers to the number of guanine repeats within the G-stretch, while \(L_1\), \(L_2\), and \(L_3\) signify a nucleotide combination forming the loops between the respective G-stretches) may, under favorable conditions, form G-quadruplexes even when the complementary DNA strand is present \(^3\)-\(^5\). There is growing evidence that DNA and RNA G-quadruplexes exist \textit{in vivo} with ensuing biological implications \(^6\)-\(^9\). Consensus G-quadruplex sequence motifs are found in important loci of the genome including telomeres, centromeres, immunoglobulin switch regions, mutation-prone hot spots, and promoter regions of oncogenes (such as \(c\)-\textit{myc}, \(c\)-\textit{myb}, \(c\)-\textit{fos}, and \(c\)-\textit{abl}) \(^2\),\(^5\),\(^10\)-\(^13\).

The \(c\)-\textit{MYC} oncogene encodes a transcription factor that upregulates expression of many genes thereby playing an important role in cell proliferation, growth, differentiation, angiogenesis, and apoptosis \(^14\)-\(^16\). The aberrant overexpression of \(c\)-\textit{MYC} is linked to the progression of a variety of human cancers including colon, cervical, breast, prostate, and small-cell lung cancers, as well as osteosarcomas, leukemias, and lymphomas \(^14\),\(^16\),\(^17\). Roughly 90% of \(c\)-\textit{MYC} transcriptional activation is controlled by the highly conserved nuclease-hypersensitive element \(\text{III}_1\) (\(\text{NHE III}_1\)) which is located 142-115 bp upstream from the \(P1\) promoter of the oncogene \(^17\)-\(^20\). \(\text{NHE III}_1\) consists of the complementary coding pyrimidine-rich and non-coding purine-rich strands which, \textit{in vitro}, can separately adopt the i-motif and G-quadruplex conformations, respectively \(^18\),\(^21\). Under conditions of transcriptionally induced negative superhelicity, the \(\text{NHE III}_1\) domain exists in a slow equilibrium between the B-DNA double stranded conformation, the single stranded conformation, and the combination of the tetra-stranded i-motif (coding strand) and G-quadruplex (non-coding strand) conformations \(^17\),\(^22\),\(^23\). The G-quadruplex that forms within \(\text{NHE III}_1\) acts as a silencer element facilitating downregulation of \(c\)-\textit{MYC} transcription \(^17\),\(^24\).
Chapter 3: Effects of Salt on the Stability of a G-quadruplexes from the Human c-MYC Promoter

The purine-rich strand of the c-MYC NHE III, 5'-TGGGGAGGGTGGGGAGGTTGGGAAGG-3' (Pu27), is a 27-mer containing five consecutive runs of guanines. Both the I, II, III, and IV and II, III, IV, and V consecutive guanine runs have the ability to fold into G-quadruplexes. Nevertheless, in the single-stranded state of Pu27, the first run of guanines (G2 to G5) is not involved in the formation of the physiologically relevant G-quadruplex. The II, III, IV, and V runs of guanines of NHE III predominantly form the basis of its folding into the G-quadruplex-based silencer element\textsuperscript{24,25}. However, when incorporated into a supercoiled plasmid and, consequently, in the context of duplex DNA under negative superhelicity, the more stable G-quadruplex appears to be the one which involves the I, II, III, and IV runs of guanines\textsuperscript{23}.

Studies of the c-MYC NHE III\textsubscript{1} sequence as well as its myc-1234 (5'-TGGGGAGGGTTTTTAGGGTGGGA-3') and myc-2345 (5'-TGAGGGTGGGGAGGGTGGGGAA-3') fragments and their modifications have established that, in the presence of K\textsuperscript{+} ions, all of them adopt the parallel G-quadruplex conformation with three double-chain-reversal loops bridging G-quartet layers\textsuperscript{25-28}. The presence of the two four-guanine stretches in the myc-1234 and myc-2345 sequences gives rise to ambiguities with respect to the actual topology of the formed G-quadruplex structures involving three G-quartets. Introduction of two G-to-T mutations at residue positions of 14 and 23 of the original c-MYC (Pu 27) sequence (5'-TGAGGGTGGGTAGGGTGGGTAA-3') restricts the modified sequence (MYC22-G14T/G23T) to adopting a single, parallel-stranded G-quadruplex conformation\textsuperscript{27}. This singularity renders the MYC22-G14T/G23T sequence an attractive candidate for thermodynamic investigations affording a more straightforward interpretation of resulting data while also being of biologically relevance.

Identification and thermodynamic characterization of forces that stabilize/destabilize G-quadruplex conformations of guanine-rich strands relative to their double- and single-stranded conformers is of fundamental and practical importance\textsuperscript{3,29}. In particular, they are important for elucidating mechanisms of biological control of G-quadruplex-based genetic switches and devising ways for operational intervention in the induction/prevention of G-quadruplex formation at desired genomic loci.
One ubiquitous force that modulates the conformational preferences of nucleic acids emerges from the differential counterion condensation (polyelectrolyte effect) of the interconverting (e.g., helix and coil) species \(^{30-34}\). Given the entropic forces originating from the polyelectrolyte effect, an increase in the solution ionic strength shifts the equilibrium between nucleic acid species towards the conformer that condenses more counterions in its vicinity \(^{30-34}\). This is the case with the helix-to-coil transitions of nucleic acid duplexes and T-rich triplexes when an increase in salt favors the helical states with a concomitant increase in the melting temperature, \(T_M\) \(^{35,36}\).

The thermal stability (\(T_M\)) of G-quadruplexes increases with an increase in the concentration of the stabilizing ion \(^3\). However, the situation with G-quadruplexes is more complicated, since, unlike other nucleic acid structures, G-quadruplexes display specific ion binding inside the central cavity in addition to counterion condensation (nonspecific binding). Consequently, the observed increase in \(T_M\) may originate from both the specific and nonspecific ion binding. In a previous publication, we have developed an experimental approach that enables one to separate the specific and nonspecific contributions to the salt-induced increase in G-quadruplex stability \(^{37}\).

In this approach, the thermal stability (melting temperature, \(T_M\)) of a G-quadruplex preformed at saturating concentrations of the stabilizing ion is measured as a function of the concentration of stabilizing or nonstabilizing [such as Cs\(^+\) or tetramethylammonium (TMA\(^+\)) ion] monovalent cations \(^{37}\). A change in \(T_M\) upon addition of the stabilizing cation (Na\(^+\) or K\(^+\)) is governed by a combination of specific (internal binding) and nonspecific (counterion condensation) salt effects. In contrast, a change in \(T_M\) upon the addition of a nonstabilizing cation is solely determined by counterion condensation. Hence, comparison of the dependences of \(T_M\) on the concentration of stabilizing and nonstabilizing cations enables one to separate the effect of specific and nonspecific binding on G-quadruplex stability.

With this approach, we have studied two G-quadruplexes formed by human telomeric sequences which are distinct with respect to their folding topology and the identity and the number of sequestered stabilizing ions (Na\(^+\)-stabilized antiparallel and K\(^+\)-stabilized hybrid structures) \(^{37}\). Our data were consistent with the picture in which the overwhelming electrostatic
contribution to G-quadruplex stability is due to the specifically bound Na\(^+\) or K\(^+\) ions and not from the effect of counterion condensation\(^3^7\). The latter effect has no or only a slightly stabilizing influence on the thermal stability of the two telomeric G-quadruplexes\(^3^7\).

In the present work, we extend our studies to characterizing the K\(^+\)-stabilized parallel G-quadruplex formed by the modified MYC22-G14T/G23T sequence. As mentioned above, MYC22-G14T/G23T is an attractive subject of thermodynamic investigations given its biological relevance and the singularity of its structure. Another peculiarity of the MYC22-G14T/G23T G-quadruplex is its unusually high thermal stability and the ability to form G-quadruplex at extremely low K\(^+\) concentrations\(^2^7,3^8\). Our results indicate a strongly destabilizing influence of the polyelectrolyte effect on the G-quadruplex formed by the MYC22-G14T/G23T sequence. We also find that the MYC22-G14T/G23T G-quadruplex can be unfolded by an increase in the concentration of nonstabilizing counterions. This is a remarkable result that, to the best of our knowledge, has never been reported for nucleic acid structures.

### 3.3 Materials and Methods

**Materials.** The MYC22-G14T/G23T oligonucleotide (ODN) sequence 5' - TGAGGTTGAGGTAGGGTGTTAAA-3' was purchased from two sources - ACGT, Inc. (Toronto, ON, Canada) and Integrated DNA Technologies (Coralville, IA, USA). We did not observe any statistically significant differences in our experimental results obtained on the ACGT or IDT DNA samples. Potassium chloride, cesium chloride, tetramethylammonium chloride, tetrabutylammonium hydroxide, and phosphoric acid were obtained from Sigma-Aldrich Canada (Oakville, ON, Canada). EDTA (free acid) was purchased from Fisher Biotech (Fair Lawn, NJ, USA). All reagents were of the highest commercially available grade and used without further purification. The solutions were prepared using doubly distilled water.

The ODN was initially dissolved in and dialyzed against a buffer consisting of 10 mM phosphoric acid titrated to pH 7.0 by tetrabutylammonium hydroxide, 0.1 mM EDTA, and the desired levels of KCl, CsCl, or TMACl. Dialysis was performed in 1000 Da molecular weight
cut-off Tube-O-Dialyzers from G Biosciences (St. Louis, MO, USA). In the absence of any added salt, the tetrabutylammonium phosphate buffer does not support the folded G-quadruplex conformation of the ODN. Given their large size, tetrabutylammonium (TBA\(^+\)) ions are fully excluded from the central cavity of G-quadruplex and, therefore, cannot stabilize a G-quadruplex.

The concentrations of the oligonucleotide were determined from the UV light absorbance at 260 nm measured at 25 °C with a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, Ontario, Canada) using molar extinction coefficient of 228,700 M\(^{-1}\) cm\(^{-1}\) for the unfolded conformation. The latter was calculated using an additive nearest neighbor procedure as described by Dr. Richard Owczarzy (http://www.owczarzy.net/extinct.htm). For our CD measurements and temperature-dependent UV light absorption measurements, the ODN concentrations were adjusted to ~30 and ~3 μM, respectively.

**Circular Dichroism Spectroscopy.** The CD spectra of the ODN were recorded in a 1 mm path-length cuvette at 25 °C using an Aviv model 62 DS spectropolarimeter (Aviv Associates, Lakewood, NJ). CD spectroscopic measurements were used to probe the conformations adopted by the ODN at various ionic conditions employed in the current study.

**UV Melting Experiments.** UV light absorption at 295 nm was measured as a function of temperature in ODN samples contained in a 1 cm path-length cuvette. The measurements were performed with a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, Ontario, Canada). The temperature was changed at a rate of 1 °C per minute. The G-quadruplex-to-coil transition temperatures, \(T_M\), and van't Hoff enthalpies, \(\Delta H_{vH}\), were evaluated from our experimental UV melting profiles using the standard procedures \(^{39,40}\).

**Computation of Solvent Accessible Surface Areas and Molecular Volumes.** Our analysis was based on the NMR-determined potassium-induced G-quadruplex structure of MYC22-G14T/G23T (PDB entry 1XAV) obtained from the RCSB Protein Databank. We calculated the solvent-accessible surface area, \(S_A\), for the G-quadruplex as the sum of the accessible surface areas of all atoms in the structure. The intrinsic volume, \(V_M\), of the ODN was calculated as molecular volumes as described by Richards \(^{41,42}\). The program MSP (Molecular Surface Package) Version 3.9.3 was obtained from Dr. Michael Connolly at
www.biohedron.com and used to calculate the solvent-accessible surface area and molecular volume, using a 1.4 Å probe radius on a Linux platform.

### 3.4 Results

**K⁺-induced G-quadruplex Formation.** Figure 1a presents the CD spectra of MYC22-G14T/G23T in a 10 mM TBA-phosphate buffer at various KCl concentrations. Inspection of Figure 1a reveals that MYC22-G14T/G23T undergoes a K⁺-induced transition to a G-quadruplex conformation. The CD spectrum of the G-quadruplex exhibits a positive maximum at 264 nm and a negative minimum at 244 nm. These features are characteristic of the parallel structure with three double-chain-reversal loops bridging the G-quartets. This notion is consistent with the NMR-derived structure of the K⁺-stabilized MYC22-G14T/G23T G-quadruplex. Note that the isoelliptic point observed at 254 nm is suggestive of the two-state nature of the K⁺-induced folding transition of the ODN.

Figure 1b shows the K⁺-dependence of the molar ellipticity of MYC22-G14T/G23T at 264 nm. Inspection of Figure 3-1b reveals that the G-quadruplex formation is complete at ~80 μM K⁺ ions. Given the oligonucleotide concentration of ~35 μM, this observation suggests an extremely tight binding of potassium ions to MYC22-G14T/G23T with its subsequent folding into G-quadruplex. For comparison, human telomeric sequences form G-quadruplexes at Na⁺ or K⁺ concentrations of 10 to 20 mM.
Figure 3-1. (a) CD spectra of MYC22-G14T/G23T in a 10 mM TBA-phosphate buffer at various KCl concentrations at 25 °C. (b) Dependence of the molar ellipticity of MYC22-G14T/G23T at 264 nm on the KCl concentration. Experimental data were approximated with Eq. (1) for n = 3 (solid line).
We used UV absorption at 295 nm to monitor the heat-induced G-quadruplex unfolding transitions at 100, 300, 600, and 1000 μM KCl. Figure 3-2a presents a representative UV melting profile of MYC22-G14T/G23T at 1mM KCl. The melting profiles were approximated by the two-state transition model from which the melting temperatures, $T_M$, and van’t Hoff enthalpies, $\Delta H_{\text{vH}}$, were determined. It should be noted, however, that G-quadruplex melting is, generally, not a two-state event and may involve one or more intermediates between the folded and unfolded states.
Figure 3-2. (a) Representative UV melting profile at 295 nm of the MYC22-G14T/G23T G-quadruplex at 1 mM KCl. (b) Dependence of the melting temperature, $T_M$, of the MYC22-G14T/G23T G-quadruplex on the concentration of K$^+$ ions. Experimental data were approximated by an exponential function solely to guide the eye (solid line).

Strictly speaking, if intermediates are present, a single $T_M$ becomes physically meaningless. More rigorous definitions may be the temperatures at which 50% of the DNA is in the folded, $T_F$, or fully unfolded, $T_U$, state. Determination of $T_F$ and $T_U$ is experimentally demanding and, generally, cannot be accomplished based on a single UV melting profile. However, the $T_M$ operationally defined as the transition midpoint is between $T_F$ and $T_U$ being close to $(T_F + T_U)/2$. In this respect, $T_M$ is a useful and convenient qualitative characteristics of G-quadruplex thermostability. Moreover, if the width of the transition does not change, a change in $T_M$ accompanying an alteration in solution conditions (e.g., ionic strength or solvent composition), $\Delta T_M$, is roughly equal to $\Delta T_F \approx \Delta T_U$. Should the width of the transition change, the deviation of $\Delta T_M$ from $\Delta T_F$ and $\Delta T_U$ should be on the order of the change in the transition width. Hence, $\Delta T_M$ can be treated as a robust measure of the G-quadruplex stability response to a change in environmental conditions.
On the other hand, the model-dependent van’t Hoff enthalpy, $\Delta H_{\text{vH}}$, determined from UV melting profiles may significantly differ from the true transition enthalpy. Given the uncertainty related to the $\Delta H_{\text{vH}}$ values, we limit their use to estimating the differential number of counterions, $\Delta n_{\text{M}^+}$, condensed around the G-quadruplex and coil states of the ODN with Eqs. (2) and (3). We treat our determined values of $\Delta n_{\text{M}^+}$ as qualitative rather than quantitative estimates.

Figure 2b plots the transition temperatures, $T_M$, of the MYC22-G14T/G23T G-quadruplex against the KCl concentrations. In agreement with a previous report, the MYC22-G14T/G23T G-quadruplex exhibits an unusually high thermal stability. For example, at 1 mM KCl, the $T_M$ is 79.9±0.2 C°. For comparison, the melting temperature of the K$^+$-stabilized hybrid-1 G-quadruplex formed by the 26-meric human telomeric sequence d[A$_3$G$_3$(T$_2$AG)$_3$A$_2$] is 70.6±0.3 C° at 200 mM KCl.

**Conformational Preferences of the K$^+$-stabilized G-quadruplex at Large Excess of Cs$^+$ and TMA$^+$ Ions.** Due to their size, Cs$^+$ and TMA$^+$ ions do not fit inside the central cavity of the G-quadruplex. If added to a preformed G-quadruplex, Cs$^+$ and TMA$^+$ ions should predominantly exert their influence on the G-quadruplex stability through the polyelectrolyte effect. We conducted CD spectroscopic measurements to see if the K$^+$-stabilized G-quadruplex undergoes any Cs$^+$- or TMA$^+$-induced transitions.

Figure 3a shows the CD spectra of the MYC22-G14T/G23T G-quadruplex in 300 μM KCl to which CsCl was incrementally added at concentrations ranging from 0 to 800 mM. Inspection of Figure 3a reveals the absence of Cs$^+$-induced conformational changes of the MYC22-G14T/G23T G-quadruplex. Similar observations were made at 600 and 1000 μM KCl (data not shown).

To ensure lack of kinetically trapped conformations of MYC22-G14T/G23T, the CD spectra of the G-quadruplex in 300 μM KCl to which 800 mM CsCl was subsequently added were recorded as the temperature was stepwise increased from 25 to 80 °C (above unfolding) and brought back to 25 °C. Figure 3-3b presents the CD spectra at ascending temperatures of 25, 40, 60, and 80 °C and descending temperatures of 60, 40, and 25 °C. At 25 °C, the CD spectra of the original and annealed G-quadruplex coincide, an observation incompatible with the existence
of kinetically trapped states. However, if CsCl is added to the ODN prior to KCl, the ODN adopts, at least, partially, a kinetically trapped state(s). We arrive at this conclusion based on ~20% increase in the CD peak at 265 nm of the annealed relative to the original ODN at 25 °C (data not shown).
Figure 3-3. (a) CD spectra of the MYC22-G14T/G23T G-quadruplex at 300 μM KCl at various CsCl concentrations. (b) CD spectra the MYC22-G14T/G23T G-quadruplex preformed in 300 μM KCl with subsequently added 800 mM CsCl at temperatures ascending from 25 to 80 °C and descending to 25 °C.

Figure 3-4a shows the CD spectra of the G-quadruplex preformed in 300 μM KCl to which TMACl was incrementally added at concentrations of up to 800 mM. Analogous to CsCl, no TMA⁺-induced changes in the conformation of the K⁺-stabilized MYC22-G14T/G23T G-quadruplex could be observed. We also did not observe TMA⁺-induced CD spectral changes for the G-quadruplex preformed in 600 and 1000 μM KCl (data not shown).

Figure 3-4b presents the CD spectra of the G-quadruplex preformed in the presence of 300 μM KCl, to which 800 mM TMACl was subsequently added, at temperatures increased in a stepwise manner from 25 to 80 °C and then reduced to 25 °C. The CD spectra of the initial and annealed ODN at 25 °C are identical which suggests the absence of kinetically trapped states. However, if TMACl was added to the ODN prior to KCl, the ODN adopts a kinetically trapped state(s). This inference is drawn based on a ~20 % increase in the intensity of the CD band at 265 nm when the sample is heated to 80 °C and subsequently cooled to 25 °C (data not shown).
Figure 3-4. (a) CD spectra of the MYC22-G14T/G23T G-quadruplex at 300 µM KCl at various TMACl concentrations. (b) CD spectra the MYC22-G14T/G23T G-quadruplex preformed in 300 µM KCl with subsequently added 800 mM TMACl at temperatures ascending from 25 to 80 °C and descending to 25 °C.

**Thermal Stability of the K⁺-stabilized G-quadruplex in the Presence of Cs⁺ and TMA⁺ Ions.** Figures 3-5a, b, and c show the CsCl-dependences of the melting temperatures, $T_M$, for the G-quadruplex preformed at 300, 600, and 1000 µM KCl, respectively. Inspection of Figure 3-5 reveals that, for all KCl concentrations, the $T_M$ steeply decreases as the CsCl concentration increases from 0 to ~200 mM and levels off at higher concentrations.
Chapter 3: Effects of Salt on the Stability of a G-quadruplexes from the Human c-MYC Promoter

\( a) \)

\[ T_m \, ^\circ C \]

\[ [\text{CsCl}], \text{mM} \]

\( b) \)

\[ T_m \, ^\circ C \]

\[ [\text{CsCl}], \text{mM} \]
Figure 3-5. Dependences of the melting temperature, $T_M$, of the MYC22-G14T/G23T G-quadruplex on the concentration of CsCl at 300 (a), 600 (b), and 1000 (c) μM KCl. Experimental data were approximated by an exponential function solely to guide the eye (solid lines).

Figures 3-6a, b, and c present the TMACl-dependences of the melting temperatures, $T_M$, for the MYC22-G14T/G23T G-quadruplex preformed in 300, 600, and 1000 μM KCl, respectively. Analogous to CsCl, the addition of TMACl causes a steep decrease in $T_M$ between 0 and ~200 mM TMACl followed by a plateau at higher concentrations.
Chapter 3: Effects of Salt on the Stability of a G-quadruplexes from the Human c-MYC Promoter

![Graph a)

![Graph b)

\( T_w \) °C vs. [TMAC] mM

---

84
Figure 3-6. Dependences of the melting temperature, $T_M$, of the MYC22-G14T/G23T G-quadruplex on the concentration of TMACl at 300 (a), 600 (b), and 1000 (c) μM KCl. Experimental data were approximated by an exponential function solely to guide the eye (solid lines).

Given the diminution in the stability of the MYC22-G14T/G23T G-quadruplex under the influence of Cs$^+$ and TMA$^+$ ions, it can be anticipated that, at some temperatures, the G-quadruplex-to-coil unfolding transition can be triggered by adding nonstabilizing cations, such as Cs$^+$ or TMA$^+$. To test this hypothesis, we performed CD spectroscopic characterization of the MYC22-G14T/G23T G-quadruplex in 600 μM KCl at 51 °C and increasing concentrations of CsCl. At 51 °C and no CsCl, MYC22-G14T/G23T overwhelmingly exists in its G-quadruplex conformation. However, addition of CsCl brings about an unfolding of the MYC22-G14T/G23T G-quadruplex as is seen from Figure 3-7. The latter presents the CD-detected CsCl-induced unfolding profile of the MYC22-G14T/G23T G-quadruplex at 51 °C. Inspection of Figure 3-7 reveals that the MYC22-G14T/G23T G-quadruplex undergoes a sigmoidal transition with a midpoint of ~75 mM CsCl.
Figure 3-7. CsCl-dependence of the molar ellipticity of MYC22-G14T/G23T at 264 nm in 600 μM KCl at 51 °C. Experimental data were approximated with Eq. (3) (solid line).

3.5 Discussion

The Stability of the MYC22-G14T/G23T G-quadruplex. The affinity of the ODN for K+ ions, with the G-quadruplex formation to follow, is very high. As is seen from Figure 3-1b, G-quadruplex formation is complete at a stoichiometric ratio of ~2 K+ ions per ODN which corresponds to the structural number of K+ ions sequestered inside the central hole of the G-quadruplex 27. The K+ induced coil (C)-to-G-quadruplex (Q) transition can be formally presented as the following binding reaction:

\[ C + nK^+ \leftrightarrow Q \]  
(Reaction 1)
By assuming a two-state transition, the binding constant of Reaction 1 can be expressed as follows:

\[ K_b = \frac{[Q]}{[C][K^+]^n} = \alpha / [(1 - \alpha)([K^+]_T - \alpha n[DNA])^n] \]  

(1)

where \([K^+]_T\) is the total concentration of potassium ions in the solution; \([DNA] = [C] + [Q]\) is the total concentration of the ODN in solution; \([K^+]\) is the concentration of unbound potassium ions; \(\alpha = [Q]/[DNA]\) is the fraction of the ODN in the G-quadruplex conformation; and, consequently, \((1 - \alpha) = [C]/[DNA]\) is the fraction of the ODN in the coil conformation.

The fraction \(\alpha\) can be evaluated for each experimental point of the titration profile presented in Figure 1b as the ratio \(\alpha = \Delta X / \Delta X_{\text{max}}\), where \(\Delta X\) is the change in an experimental observable (here, molar ellipticity at 264 nm) relative to its initial value at \([\text{KCl}] = 0\); and \(\Delta X_{\text{max}}\) is the asymptotic maximum change in X corresponding to the fully folded state of the ODN. We fitted the titration profile presented in Figure 1b with Eq. (1) with \(n = 1, 2, \text{ or } 3\). The fitting was performed numerically based on a direct search optimization technique for fitting nonlinear systems. The best fit of the data was achieved at \(n = 3\) with a binding constant \(K_b\) of \(1.19 \times 10^{14} \text{ M}^{-3}\).

With \(n\) equal to 3, one can hypothesize that, in addition to the two NMR-detected potassium ions sequestered inside the central cavity, there is an additional \(K^+\) ion bound to the G-structure albeit in a non-localized and/or unimobilized (and, therefore, invisible to NMR) manner. A note of caution is warranted here. Although the isoelliptic point at 254 nm in Figure 3-1a is consistent with the two-state nature of the \(K^+\)-induced MYC22-G14T/G23T folding, the folding/unfolding transitions of G-quadruplexes are, generally, non-two-state. Consequently, we do not exclude the possibility that the best fit obtained at \(n = 3\) may be an artifact related to the two-state approximation used to derive Eq. (1). At present, we view the resulting value of \(n\) of 3 as a qualitative estimate which calls for further studies to determine the nature of the \(K^+\)-induced transition of MYC22-G14T/G23T and the number of associated potassium ions.
In principle, the $K^+$-dependence of the melting temperature, $T_M$, presented in Figure 3-2b can also be used to evaluate the differential number of counterions ($M^+$) bound to the G-quadruplex and coil states:

$$\Delta n_{M^+} = \left(\frac{\delta T_M}{\delta \ln[M^+]}\right) \Delta H_M/RT_M^2$$  \hspace{1cm} (2)

The analysis yields a $\Delta n_{M^+}$ of 1.8±0.2. This result is close to 2, the number of centrally bound potassium ions. The analysis is compromised by the use in Eq. (2) of the model-dependent van't Hoff enthalpy that may significantly differ from the true transition enthalpy. Notwithstanding, our aggregate results suggest that the folding of MYC22-G14T/G23T is accompanied by an uptake of 2 to 3 potassium ions in good agreement with the structural data.

**Cs$^+$ and TMA$^+$ Ions Do Not Cause Structural Alterations in the Preformed G-quadruplex.** Once formed in the presence of $K^+$ ions, the MYC22-G14T/G23T G-quadruplex retains its conformation even at a large excess of Cs$^+$ or TMA$^+$ ions (see Figures 3-3a and 3-4a). We infer that the G-quadruplex conformation corresponds to the global minimum of free energy and not to a kinetically trapped local minimum. This notion follows from the observation that the annealing of the preformed G-quadruplex to which CsCl or TMACl was subsequently added does not lead to a change in the CD spectrum (see Figures 3-3b and 3-4b).

On the other hand, if the ODN is initially in 800 mM CsCl or TMACl and KCl is added later, a kinetically trapped state(s) is (are) formed. The kinetically trapped state(s) converts into the thermodynamically stable one once the solution is heated to 80 °C and cooled back to 25 °C. In other words, annealing the ODN eliminates kinetically trapped species converting them into the thermodynamically stable conformation.

In the UV melting experiments discussed below, the G-quadruplexes were preformed in KCl with the desired amounts of CsCl or TMACl added afterward. In addition, the ODN
samples were annealed prior to the experiments. Therefore, when interpreting our data, we assume a single, thermodynamically stable G-quadruplex conformation.

**Modulation of the stability of the K\(^+\)-stabilized G-quadruplex by Cs\(^+\) and TMA\(^+\) ions.** At every KCl concentration studied, the thermal stability, \(T_M\), of the G-quadruplex sharply decreases as the concentration of Cs\(^+\) or TMA\(^+\) ions is increased and levels off above ~200 mM (see Figures 3-5 and 3-6). At 300, 600, and 1000 μM KCl, Cs\(^+\)-induced decreases in \(T_M\) are 25±2, 21±2, and 23±2 °C (with an average of 23±1 °C), respectively, while TMA\(^+\) ions cause a depression of \(T_M\) by 15±1, 16±1, and 16±1 °C (with an average of 16±1 °C), respectively.

We have previously observed that nonstabilizing ions do not influence or weakly increase the thermal stability of the human telomeric G-quadruplexes. In the case of the c-MYC promoter sequence, the situation is more dramatic; Cs\(^+\) or TMA\(^+\) ions strongly destabilize the G-quadruplex conformation. We rationalize these experimental observations by proposing that counterion condensation is reduced in the G-quadruplex state of MYC22-G14T/G23T relative to the coil state. In other words, there are more counterions (Cs\(^+\) or TMA\(^+\)) associated with the coil state compared to the G-quadruplex state. The differential number, \(\Delta n_{M^+}\), of counterions bound to the G-quadruplex and coil states can be estimated from the initial slopes \(\delta T_M/\delta \ln[Cs^+]\) or \(\delta T_M/\delta \ln[TMA^+]\) in Figures 3-5 and 3-6 using Eq. (2). The estimates yield the values of \(\Delta n_{M^+}\) of -1.1±0.1 and -0.9±0.1 for Cs\(^+\) and TMA\(^+\), respectively.

The effect of differential counterion condensation on G-quadruplex stability ceases when the bulk concentration of the counterions approaches the local concentration in the vicinity of the coil state. Thus, the local counterion concentration in the vicinity of the coil state is on the order of ~200 mM compared to ~1 M, the estimate for duplex polymeric ODN.

**Cs\(^+\)-induced Unfolding of the K\(^+\)-stabilized G-quadruplex.** To the best of our knowledge, the unfolding profile shown in Figure 3-7 is the first reported salt-induced unfolding
transition of a nucleic acid structure. We analyzed the transition profile in Figure 3-7 within the framework of a two-state model in which the fraction unfolded, $\alpha$, is given by

$$\alpha = \frac{K}{1 + K} \quad (3)$$

where $K = K_0(1 + k[M^+])^{\Delta n_{M^+}}$ is the salt-dependent equilibrium constant for the reaction of interconversion between the coil and G-quadruplex states; $K_0$ is the equilibrium constant in the absence of Cs$^+$ ions ($M^+$).

When fitting the data in Figure 3-7 with Eq. (3), the value of $K_0$ was set to 0.0032. The latter was estimated from our UV melting data, as $K_0 = \exp(-\Delta G^\circ/RT)$, where $\Delta G^\circ = \Delta H_M(1 - T/T_M)$. For the experimental conditions in Figure 3-7 (600 μM KCl and 51 °C), the values of $T_M$, $\Delta H_M$, and $\Delta G^\circ$ are 74.3±0.3 °C, 54.7±1.4 kcal mol$^{-1}$, and 3.7±1.4 kcal mol$^{-1}$, respectively. The fit produces $\Delta n_{M^+}$ of -1.4±0.1. This value is in good agreement with -1.1, the estimate made above based on the Cs$^+$-dependence of $T_M$. The agreement lends credence to our analysis and the two-state approximation used in Eq. (3).

Enhanced Stability and Reduced Polyelectrolyte Effect of the MYC22-G14T/G23T G-quadruplex. The exceptionally high thermal stability of the MYC22-G14T/G23T G-quadruplex is consistent with the picture in which the intramolecular phosphate-phosphate repulsion, a major destabilizing force, is greatly reduced. The reduction may be caused by structural features and/or neutralization of the negative phosphate charges by the internally bound potassium ions. These factors contribute to the high thermal stability of a G-quadruplex and, concomitantly, to weakening the polyelectrolyte effect.

In G-quadruplexes, the polyelectrolye effect may be weakened, partially, due to the presence of cations within the central cavity that decrease the charge density of the DNA. In addition, it may be further weakened due to the structural factors as the near spherical shape of a
G-quadruplex facilitates an increase in the distance between the G-strands thereby decreasing the charge density in the system. This possibility is underscored by recent theoretical studies which suggest that a charged spherical construct displays reduced counterion condensation relative to a cylinder. While these arguments can be put forward with respect to any G-quadruplex, their effect should be stronger in the MYC22-G14T/G23T G-quadruplex compared to, for example, human telomeric G-quadruplexes. The latter exhibit a moderate stability, a lower affinity for the stabilizing ion, and a stronger polyelectrolyte effect compared to c-MYC.

As a first approximation, the neutralizing influence of the internally bound potassium ions should correlate to their average separation from the phosphates within the G-quadruplex, while the interphosphate repulsion should depend on the average distance between the phosphate oxygens. One way to analyze the structural disparities between the c-MYC and telomeric G-quadruplexes is to compare their solvent accessible surface areas, $S_A$, and molecular volumes, $V_M$. The solvent-accessible surface area of a G-quadruplex normalized per phosphate offers a measure of the average interphosphate distance, while the molecular volume per nucleotide characterizes the compactness of a G-quadruplex and, hence, the average distance between the central ions and the phosphates. Table 3-1 lists the computed data on $S_A$ and $V_M$ for the MYC22-G14T/G23T G-quadruplex and the telomeric Tel22 and Tel26 G-quadruplexes from our previous works.

Table 3-1. Molecular volumes, $V_M$, and solvent accessible surface areas, $S_A$, of the c-MYC and human telomeric Tel22 and Tel26 G-quadruplexes. Values normalized per nucleotide are shown in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>$V_M$, Å³</th>
<th>$S_A$, Å²</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-MYC</td>
<td>6178±16</td>
<td>4442±33</td>
</tr>
</tbody>
</table>
(281±1) (202±2)

---

<table>
<thead>
<tr>
<th></th>
<th>(292±1)</th>
<th>(178±5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tel22</td>
<td>6426±26</td>
<td>3907±99</td>
</tr>
<tr>
<td>Tel26</td>
<td>7092±27</td>
<td>3963±69</td>
</tr>
</tbody>
</table>

---

\(^a\) from ref. 44.

\(^b\) from ref. 45.

Inspection of data presented in Table 3-1 reveals that the molecular volume of the c-MYC G-quadruplex normalized per nucleotide is 281±1 cm\(^3\) mol\(^{-1}\). This number is intermediate between those of the Na\(^+\)-stabilized antiparallel Tel22 (292±1 cm\(^3\) mol\(^{-1}\)) and the K\(^+\)-stabilized hybrid-1 Tel26 (273±1 cm\(^3\) mol\(^{-1}\)) human telomeric G-quadruplexes. Thus, the c-MYC G-quadruplex exhibits an average compactness being less compact than the Tel26 G-quadruplex but more compact than the Tel22 G-quadruplex. Thus, we conclude that the central ion-induced neutralization of the phosphates cannot be the main reason for the observed extraordinary stability of the c-MYC G-quadruplex.

The solvent-accessible surface area occupied by each phosphate on the surface of the MYC22-G14T/G23T G-quadruplex is 15 to 30 % larger than that in the Tel22 and Tel26 G-quadruplexes. If the area assigned to each phosphate is viewed as a square, a 15 to 30 % increase in area translates into a 5 to 10 % increase in the interphosphate distance in the MYC22-G14T/G23T relative to the Tel22 and Tel26 G-quadruplexes. The detected increase in
interphosphate separation, in turn, signifies a lower charge-charge repulsion and a smaller charge
density of the MYC22-G14T/G23T G-quadruplex. This notion is in agreement with the
observed greater stability and lower polyelectrolyte effect of the MYC22-G14T/G23T G-
quadruplex compared to human telomeric G-quadruplexes.

In the aggregate, our simplified analysis is consistent with the picture in which the
outstanding stability of the MYC22-G14T/G23T G-quadruplex reflects and, in fact, results from
structural features. However, given the limited set of analyzed topologies, it is currently unclear
if the large interphosphate separation is an individual feature of the MYC22-G14T/G23T G-
quadruplex or a general characteristics of the parallel topology. Further studies are required to
shed light on the generality of our findings.

3.6 Conclusion

We report a systematic investigation of the thermal stability of a G-quadruplex formed by
the modified c-MYC NHE III1 sequence with two G-to-T mutations (5'-
TGAGGGTGGGTAGGGTGGGTAA-3') in the presence of stabilizing and non-stabilizing
cations. Introduction of two G-to-T mutations at residue positions of 14 and 23 of the original
c-MYC sequence restricts the modified sequence (MYC22-G14T/G23T) to adopting a single,
parallel-stranded K+-stabilized G-quadruplex conformation. The MYC22-G14T/G23T G-
quadruplex exhibits a steady increase in the melting temperature, TM, upon an increase in the
concentration of the stabilizing cation K+. On the other hand, an increase in the concentration of
nonstabilizing Cs+ or TMA+ cations at a constant concentration of K+ causes a sharp decline in
TM followed by leveling off at ~200 mM Cs+ or TMA+. Moreover, at 51 °C and 600 μM K+, an
increase in CsCl concentration from 0 to 800 mM leads to unfolding of the G-quadruplex. These
observations are consistent with the picture in which the unfolded state of MYC22-G14T/G23T
exhibits a stronger polyelectrolyte effect than the G-quadruplex state. In fact, we estimate that
the unfolded state of MYC22-G14T/G23T condenses roughly one extra counterion relative to the
G-quadruplex state.
Taken together with our earlier studies of human telomeric G-quadruplexes, our data suggest that the stabilizing action of cations on the G-quadruplex conformation is due, primarily, to the internally bound metal ions which act as ligands. Nonspecifically bound (condensed) counterions in the vicinity of DNA may exert a very slightly stabilizing or no influence on a G-quadruplex (human telomeric G-quadruplexes) or may strongly destabilize it (MYC22-G14T/G23T). The destabilizing influence of the polyelectrolyte effect on the MYC22-G14T/G23T G-quadruplex may be related to its structural features. In the MYC22-G14T/G23T G-quadruplex, the average solvent accessible surface area per phosphate is 15 to 30% larger than that in human telomeric G-quadruplexes which translates into a 5 to 10% greater interphosphate distance in the former. An increase in interphosphate separation signifies a reduced charge-charge repulsion and a lower charge density of the MYC22-G14T/G23T G-quadruplex. This notion is in agreement with the observed greater stability and lower polyelectrolyte effect of the MYC22-G14T/G23T G-quadruplex.

3.7 Acknowledgements

This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to TVC.
3.8 Reference


Chapter 3: Effects of Salt on the Stability of a G-quadruplexes from the Human c-MYC Promoter


CHAPTER 4: Ionic Effects on VEGF G-quadruplex Stability


Author’s Contribution: Byul G. Kim performed the majority of the experiments, trained the undergraduate student Ji Long to run some remaining experiments, analyzed the data, and participated in the preparation of the manuscript.
4.1 Abstract

In a potassium solution, a modified 22-meric DNA sequence Pu22-T12T13 from a region proximal to the transcription initiation site of the human VEGF gene adopts a single parallel-stranded G-quadruplex conformation with a 1:4:1 loop-size arrangement. We measured the thermal stability, T_M, of the K\(^+\)-stabilized Pu22-T12T13 G-quadruplex as a function of stabilizing K\(^+\) ions and nonstabilizing Cs\(^+\) and TMA\(^+\) ions. The thermal stability, T_M, of the Pu22-T12T13 G-quadruplex increases with the concentration of the stabilizing potassium ions, while sharply decreasing upon the addition of the nonstabilizing cations. We interpret these results as underscoring the opposing effects of internal binding and counterion condensation on the stability of the Pu22-T12T13 G-quadruplex. While centrally bound ions stabilize the G-quadruplex conformation, counterion condensation destabilizes it favoring the coil conformation. From the initial slopes of the dependences of T_M on the concentration of Cs\(^+\) and TMA\(^+\) cations, we estimate that the deleterious effect of counterion condensation stems from roughly one extra counterion associated with the coil relative to the G-quadruplex state of Pu22-T12T13. The reduced accumulation of counterions around the G-quadruplex state of Pu22-T12T13 relative to its coil state is due to the low surface charge density of the G-quadruplex reflecting its structural characteristics. Based on the analysis of our data along with the results of a previous study, we propose that the differential effect of internally (stabilizing) and externally (destabilizing) bound cations may be a general feature of parallel intramolecular G-quadruplexes.
4.2 Introduction

G-quadruplexes have emerged as a major class of non-canonical DNA structures of biological importance and a potential target of anticancer and antitumor strategies \(^1\)\(^-\)\(^6\). Guanine-rich DNA sequences with a potential to fold into a G-quadruplex conformation \textit{in vivo} have been found in important loci of the genome including telomeres, centromeres, immunoglobulin switch regions, mutation-prone hot spots, and promoter regions of oncogenes \(^4\)\(^,\)\(^6\)\(^-\)\(^12\). A growing body of experimental evidence suggests that DNA and RNA G-quadruplexes, in fact, do exist in the cell and play a regulatory role in genomic processes \(^13\)\(^-\)\(^15\). Drug-mediated stabilization of local G-quadruplexes in telomeric ends of chromosomes to inhibit telomerase activity and/or in the promoter regions of target oncogenes to prevent their expression is under development as prospective cancer therapy \(^3\)\(^,\)\(^16\).

Controlled regulation of G-quadruplex formation at specific DNA loci requires an understanding of the underlying thermodynamic forces that contribute to making a G-quadruplex conformation favored over the competing double- or single-stranded conformations. The G-quadruplex conformation is maintained by a complex interplay of a host of inter- and intramolecular interactions that act to stabilize or destabilize it \(^12\). Elucidation of the nature of these interactions and their thermodynamic characterization is also needed for rational design of low molecular weight compounds that would shift the population equilibrium between DNA structures towards the G-quadruplex conformation.

The electrostatic component of the balance of forces stabilizing ordered nucleic acid structures originates from the polyanionic nature of the DNA and is modulated by the identity and concentration of counterions present in solution \(^12\)\(^,\)\(^17\)\(^-\)\(^20\). In contrast to other ordered DNA structures, G-quadruplexes are subject to both specific and nonspecific ion binding \(^12\)\(^,\)\(^21\)\(^,\)\(^22\). The specific binding is due to metal ions coordinated inside the central cavity to the O6 oxygens of guanine bases, while the nonspecific binding involves external condensation of counterions around the phosphate groups \(^12\). In two recent publications, we have developed an experimental approach that permits one to discriminate between the contributions of specific and nonspecific cation binding to G-quadruplex stability \(^21\)\(^,\)\(^22\). In this approach, the thermal stability, T\(_M\), of a
preformed G-quadruplex is measured as a function of the concentration of stabilizing and nonstabilizing monovalent cations. The stabilizing ions are the cations that support G-quadruplex formation by penetrating into the central cavity and forming coordination bonds with the N6 oxygens of the guanines. Nonstabilizing cations are organic or inorganic cations that, given their size (being either too large or too small), do not fit inside the central cavity of the G-quadruplex and, therefore, per se do not support G-quadruplex formation. While the effect of stabilizing cations on $T_M$ includes both the internal and external binding, the effect of nonstabilizing cations is limited to external binding as counterions. Comparative analysis of the dependences of $T_M$ on the concentration of stabilizing and nonstabilizing cations can be used to discriminate between the influence exerted by the internal (specific) and external (nonspecific) binding on G-quadruplex stability.

Our combined results suggest that specific cation binding stabilizes the G-quadruplex conformation, while nonspecific binding may be neutral or may slightly stabilize or strongly destabilize it $^{21,22}$. Counterion condensation (the polyelectrolyte effect) has been found to dramatically decrease the thermal stability, $T_M$, of the parallel G-quadruplex formed by a modified DNA sequence from the promoter region of the c-MYC oncogene $^{22}$. Moreover, at some conditions, an increase in the concentration of Cs$^+$ ions in the solution of the K$^+$-stabilized c-MYC G-quadruplex brings about its complete unfolding to the coil conformation $^{22}$. This observation is in stark contrast with the counterion-induced increase in thermal stability conventionally observed for other nucleic acid structures establishing the parallel c-MYC G-quadruplex as a new paradigm for DNA-salt interactions $^{19,23-27}$.

Despite the progress made by these studies, our understanding of the electrostatic effects on the stability of G-quadruplexes is far from complete. In the present work, we assess the generality of the observed counterion-induced reduction in the stability of parallel G-quadruplexes commonly found in human promoter sequences. Specifically, we study a G-quadruplex formed by a modified 22-meric DNA sequence (Pu22-T12T13) from a region proximal to the transcription initiation site of the human vascular endothelial growth factor (VEGF) gene $^{28-30}$. VEGF is a homodimeric signal protein involved in vasculogenesis and angiogenesis $^{31,32}$. Overexpression of VEGF is implicated in development and progression of
many types of human cancers, rheumatoid arthritis, diabetic retinopathy, and age-related macular degeneration. The proximal 39-base pair polyG/polyC tract (-88 to -50 relative to the transcription initiation site) is functionally significant for inductive VEGF expression. The wild type VEGF-Pu22 sequence (5’-CGGGGCGGGGCAGCAGCAGCGG-3’) is a 22-meric oligonucleotide at the 5’-end of the G/C-rich tract (-86 to -65 relative to the transition site). The Pu22-T12T13 sequence (5’-CGGGGCGGGCGCTTGGGCGGGGT-3’) differs from the wild-type sequence VEGF-Pu22 by the wild-type G12 and G13 substituted with T12 and T13. In potassium solution, Pu22-T12T13 has been shown to adopt a single parallel-stranded G-quadruplex conformation with a 1:4:1 loop-size arrangement. In the Pu22-T12T13 G-quadruplex, the four-nucleotide middle loop stretches over the terminal quartet forming a specific caping structure involving the two flanking segments. This structural feature of Pu22-T12T13 is in contrast to our previously studied c-MYC G-quadruplex in which the two-nucleotide middle loop remains in the groove.

4.3 Materials and Methods

Materials. The Pu22-T12T13 oligonucleotide sequence 5’-CGGGGCGGGCGCTTGGGCGGGGT-3’ was purchased from two sources - ACGT, Inc. (Toronto, ON), and Integrated DNA Technologies (Coralville, IA). No statistically significant inconsistencies were observed in the experimental results obtained on the ACGT or IDT DNA samples. Potassium chloride, cesium chloride, tetramethylammonium chloride, tetrabutylammonium hydroxide, and phosphoric acid were obtained from Sigma-Aldrich Canada (Oakville, ON). EDTA (free acid) was purchased from Fisher Biotech (Fair Lawn, NJ).

The DNA was initially dissolved in and dialyzed against a denaturing buffer consisting of 10 mM phosphoric acid titrated to pH 7.0 with tetrabutylammonium hydroxide and 0.1 mM EDTA. Given their large size, tetrabutylammonium (TBA+) ions do not stabilize a G-quadruplex. Dialysis was carried out with 1000 Da molecular mass cut off Tube-O-Dialyzers from G
Biosciences (St. Louis, MO). Once the DNA was fully unfolded through dialysis, desired levels of KCl and CsCl or TMACl were added to the solution.

The concentrations of the oligonucleotide were determined spectrophotometrically at 25 °C using a molar extinction coefficient, $\varepsilon_{260}$, of 200,400 M$^{-1}$cm$^{-1}$ for the unfolded conformation. The latter was calculated using a nearest neighbor procedure as described by Dr. Richard Owczarzy (http://www.owczarzy.net/extinctionDNA.htm). In the CD measurements and temperature-dependent UV light absorption measurements described here, the DNA concentrations were adjusted to ~30 and ~3 μM, respectively.

**Optical Spectroscopy.** The circular dichroism (CD) spectra of Pu22-T12T13 were recorded at 25 °C in a 1 mm path-length cuvette using an Aviv model 62 DS spectropolarimeter (Aviv Associates, Lakewood, NJ). CD spectroscopic measurements were used to probe the conformations adopted by the DNA under each experimental condition employed in this study.

UV light absorption at 295 nm was measured as a function of temperature in Pu22-T12T13 samples contained in a 10 mm path-length cuvette. The measurements were performed with a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, ON). The temperature was changed at a rate of 1 °C/min. The G-quadruplex-to-coil transition temperatures, $T_M$, and van’t Hoff enthalpies, $\Delta H_{vH}$, were quantified from our experimental UV melting profiles using the standard procedures $^{37-40}$.

**Computation of Solvent Accessible Surface Areas and Molecular Volumes.** Our analysis was based on the NMR structure of the potassium-induced Pu22-T12T13 G-quadruplex (Protein Data Bank entry 2M27) obtained from the RCSB Protein Data Bank. As 2M27 contains ten models, we treated each model as a distinct structure. We calculated the solvent accessible surface area, $S_A$, for each structure as the sum of the accessible surface areas of all atoms in the structure $^{41,42}$. The intrinsic volume, $V_M$, of the DNA was calculated as a molecular volume as described by Richards $^{41,43}$. MSP (Molecular Surface Package) version 3.9.3 was obtained from Dr. Michael Connolly at www.biohedron.com and used to calculate the solvent accessible surface area and molecular volume using a 1.4 Å probe radius on a Linux platform.
4.4 Results

K⁺-induced Formation and Stability of the Pu22-T12T13 G-quadruplex. Figure 4-1a presents the CD spectra of Pu22-T12T13 in a TBA-phosphate buffer in the absence and presence of KCl. Inspection of the CD spectra presented in Figure 1a reveals the K⁺-induced formation of a parallel G-quadruplex as can be judged by a minimum at 244 nm and a maximum at 264 nm\(^ {44,45}\). The CD-detected parallel structure of the G-quadruplex adopted by Pu22-T12T13 in the presence of potassium ions is consistent with the NMR solution structure \(^ {35}\).

Figure 4-1b shows the folding profile of Pu22-T12T13 as monitored by the K⁺-dependence of the molar ellipticity at 264 nm. It can be seen from Figure 4-1b that the Pu22-T12T13 G-quadruplex formation is complete above ~2 mM KCl. This value is intermediate between 80 μM and 15 mM KCl observed for the K⁺-induced G-quadruplex formation by the MYC22-G14T/G23T sequence and the human telomeric Tel26 sequence, respectively\(^ {22,46}\).
Figure 4-1. (a) CD spectra of Pu22-T12T13 in a 10 mM TBA-phosphate buffer at various KCl concentrations at 25 °C. (b) Dependence of the molar ellipticity of Pu22-T12T13 at 264 nm on the KCl concentration. Experimental data were approximated with Eq. (1) for n = 2 (solid line).

G-quadruplex melting experiments performed at various KCl concentrations were monitored by UV absorption at 295 nm 39. Figure S1 of Supporting Information presents a representative UV melting profile obtained at 2 mM KCl in the absence of nonstabilizing cations. The transition temperatures, T_M, and van't Hoff enthalpies, ΔH_vH, were determined as described previously 37-39. Figure 4-2a presents the K⁺-dependence of the transition temperatures, T_M. As is seen from Figure 4-2a, the thermal stability, T_M, of Pu22-T12T13 increases in a nearly linear manner as the concentration of stabilizing K⁺ ions increases. This observation is in agreement with the behavior of all studied G-quadruplexes that display an increase in T_M with an increase in the concentration of stabilizing cations 12,21,22,46-48.

Figure 4-2b plots the transition van't Hoff enthalpies, ΔH_vH, against the melting temperature, T_M. The use of van't Hoff enthalpies, ΔH_vH, in this study is limited to calculation of the differential numbers of associated cations in the G-quadruplex and coil states with Eq. (2) (see below). For the salt conditions studied in this work, the combined dependences of the van't Hoff enthalpies, ΔH_vH, for the G-quadruplex-to-coil transitions on the transition temperature, T_M, can
be described by $\Delta H_{vH}(\text{kcal mol}^{-1}) = (-14 \pm 9) + (0.9 \pm 0.1)T_M(\text{°C})$. The slope of the dependence of $\Delta H_{vH}$ on $T_M$ of 0.9±0.1 kcal mol$^{-1}$ suggests a positive change in heat capacity, $\Delta C_P$, associated with the G-quadruplex-to-coil transition of Pu22-T12T13. This notion is consistent with the results of a study from the Shafer laboratory which reports that melting of G-quadruplexes are accompanied by significant changes in heat capacity$^{49}$. However, in a previous work, we have not observed any substantial variation of the transition enthalpy with $T_M$ for the human telomeric Tel26 G-quadruplex$^{46}$. The differential heat capacity, $\Delta C_P$, of the G-quadruplex and coil states may depend on the specific DNA sequence. Clearly, further studies are required to clarify this important point.
CHAPTER 4: Ionic Effects on VEGF G-quadruplex Stability

Figure 4-2. (a) Dependence of the melting temperature, $T_M$, of the Pu22-T12T13 G-quadruplex on the concentration of K$^+$ ions. Experimental data were approximated by a linear function (solid line). (b) The G-quadruplex-to-coil transition van't Hoff enthalpies of Pu22-T12T13 plotted as a function of the transition temperature, $T_M$, at 2 (■), 5 (□), and 10 (●) mM KCl and TMACl concentrations ranging from 0 to 800 mM and 2 (○), 5 (♦), and 10 (◊) mM KCl and CsCl concentrations ranging from 0 to 100 mM.

$Cs^+$- and TMA$^+$-induced Conformational Changes of Pu22-T12T13. Figures 3a, b, and c present the CD spectra of the Pu22-T12T13 G-quadruplex preformed in the presence of 2, 5, and 10 mM KCl at different concentrations of TMACl. Inspection of Figures 3a, b, and c reveals that, at all KCl concentrations studied here, the Pu22-T12T13 G-quadruplex retains its conformation as the concentration of TMA$^+$ ions increases up to 800 mM.
CHAPTER 4: Ionic Effects on VEGF G-quadruplex Stability

(a) 

(b)
CHAPTER 4: Ionic Effects on VEGF G-quadruplex Stability

Figure 4-3. CD spectra of the Pu22-T12T13 G-quadruplex at 2 (a), 5 (b), and 10 (c) mM KCl at 25 °C and various concentrations of TMACl.

Figures 4-4a, b, and c show the CD spectra of the Pu22-T12T13 G-quadruplex formed in the presence of 2, 5, and 10 mM KCl at different concentrations of CsCl. In contrast to TMA⁺ ions, Cs⁺ ions cause visible changes in the CD spectra of the preformed Pu22-T12T13 G-quadruplex as the concentrations of CsCl increases above 100 mM. However, below 100 mM CsCl, the CD spectra of the Pu22-T12T13 G-quadruplex are similar to the original spectrum in the absence of Cs⁺ ions. As a plausible explanation, these observations may reflect conformational alterations resulting from replacement of some of the central K⁺ ions by Cs⁺ ions at high cesium-to-potassium ratios.
CHAPTER 4: Ionic Effects on VEGF G-quadruplex Stability

a) 

b)
It should be also noted that folding of heat-induced unfolded Pu22-T12T13 triggered by a decrease in temperature from 90 to 25 °C in the presence of a large excess of Cs⁺ over K⁺ ions takes approximately two hours to complete (see Figure S2 of Supporting Information). The slow folding kinetics in the environment of Cs⁺ ions is distinct from the fast folding of Pu22-T12T13 in the presence of K⁺ ions or in K⁺/TMA⁺ mixtures which is on the order of the dead time of the experiment (~1 minute). In contrast to Cs⁺ ions, apparently, due to their larger size, the organic TMA⁺ ions do not compete with K⁺ ions even at very high concentrations.

This work aims at understanding the external effect of nonstabilizing cations as counterions. Consequently, we limit the analysis of our Cs⁺-dependent results to CsCl concentrations of up to 100 mM when their influence can be considered to arise only from external binding.

Cs⁺- and TMA⁺-induced Modulation of Melting Temperature, $T_M$. Panels A, B, and C of Figure 4-5 present the TMA⁺-dependences of the transition temperatures, $T_M$, of Pu22-T12T13 at
2, 5, and 10 mM KCl, respectively. Inspection of Figure 4-5 reveals sharp decreases in $T_M$ with an increase in the concentration of TMA$^+$ ions following by leveling off above ~200 mM. The net changes in $T_M$ are -14±1, -12±1, and -10±1 °C for 2, 5, and 10 mM KCl, respectively.
CHAPTER 4: Ionic Effects on VEGF G-quadruplex Stability

Figure 4-5. Dependences of the melting temperature, $T_M$, of the Pu22-T12T13 G-quadruplex on the concentration of TMACl at 2 (a), 5 (b), and 10 (c) mM KCl. Experimental data were approximated by an exponential function solely to guide the eye (solid lines).

Figures 4-6a, b, and c plot the changes in $T_M$ of the Pu22-T12T13 G-quadruplex as the concentration of CsCl increases from 0 to 100 mM at 2, 5, and 10 mM KCl, respectively. Inspection of Figure 4-6 reveals a steady decline in $T_M$ with as the concentration of Cs$^+$ ions increases from 0 to 100 mM. The changes in $T_M$ are -16±1, -15±1, and -11±1 °C for 2, 5, and 10 mM KCl, respectively.
Figure 4-6. Dependences of the melting temperature, T_M, of the Pu22-T12T13 G-quadruplex on the concentration of CsCl at 2 (a), 5 (b), and 10 (c) mM KCl. Experimental data were approximated by an exponential function solely to guide the eye (solid lines).

4.5 Discussion

*Stability of the Pu22-T12T13 G-quadruplex.* The folding profile presented in Figure 1b can be formally analyzed as an isothermal G-quadruplex folding reaction induced by cation binding. Within the framework of an association reaction in which the binding of n potassium ions to the unfolded coil-like state (C) results in the folded G-quadruplex state (Q), one can write:

\[ C + nK^+ \leftrightarrow Q \]  
(Reaction 1)
Under the assumption of a two-state coil-to-G-quadruplex transition, the apparent binding constant, $K_b$, of Reaction 1 is given by the relationship:

$$K_b = [Q] / ([C][K^+]^n) = \alpha / ((1 - \alpha) ([K^+]_T - \alpha n[DNA])^n) \tag{1}$$

where $[DNA] = [Q] + [C]$ is the total concentration of DNA; $[K^+]_T$ is the total concentration of potassium ions; $\alpha = [Q]/[DNA]$ is the fraction of the DNA in the G-quadruplex conformation.

The best fit of experimental data in Figure 1b with Eq. (1) is produced by $n$ equal to 2 and yields an apparent binding constant, $K_b$, of $(2.89\pm0.01) \times 10^6 \text{ M}^{-2}$. The $n$ of 2 is consistent with the structural number of potassium ions sequestered inside the central cavity of the G-quadruplex.\(^{35}\)

The number of counterions that become released to the bulk upon the G-quadruplex unfolding can also be evaluated from the $K^+$-dependence of the melting temperature, $T_M$, presented in Figure 2a\(^{18,52-54}\):

$$\Delta n_{M^+} = (\delta T_M/\delta \ln[M^+]) \Delta H_M/RT_M^2 \tag{2}$$

where $\Delta n_{M^+}$ is the differential number of counterions ($M^+$) thermodynamically associated with the G-quadruplex and coil conformations.

Using Eq. (2) in conjunction with the data presented in Figure 2a, we calculate a $\Delta n_{M^+}$ of $1.7\pm0.2$, consistent with our estimate based on Eq.(1) and Figure 1b.
**TMA⁺ Ions Do not Influence the Conformation of the K⁺-stabilized Pu22-T12T13 G-quadruplex.** The CD spectroscopic spectra shown in Figures 4-3 and 4-4 suggest that, while TMA⁺ ions do not influence the structure of the preformed Pu22-T12T13 G-quadruplex, Cs⁺ ions cause its structural alterations at concentrations higher than ~100 mM. The presence of Cs⁺ cations leads to a slow folding kinetics of the thermally unfolded G-quadruplex which additionally complicates the temperature-dependent measurements conducted in this work. We, therefore, limit discussions of Cs⁺-dependent ionic effects to the concentrations of up to 100 mM.

The influence of TMA⁺ (between 0 and 800 mM) and Cs⁺ ions (between 0 and 100 mM) on the stability of the Pu22-T12T13 G-quadruplex preformed in the environment of potassium ions can be viewed as originating from and reflecting the polyelectrolyte nature of DNA (external nonspecific binding). At higher concentrations, Cs⁺ is likely to compete with K⁺ for the sites inside the central cavity of the G-quadruplex which would, consequently, bring about conformational alterations and slow the folding kinetics.

**TMA⁺- and Cs⁺-induced Modulation of the Stability of K⁺-stabilized Pu22-T12T13 G-quadruplex.** Results of our previous studies suggest that counterion condensation may slightly increase the stability of a G-quadruplex, may not have any sizeable impact on it, or may significantly destabilize it²¹,²². The K⁺-stabilized parallel G-quadruplex formed by the modified MYC22-G14T/G23T sequence from the promoter region of the human c-MYC oncogene is structurally and functionally similar to the Pu22-T12T13 G-quadruplex studied here. The MYC22-G14T/G23T G-quadruplex preformed in the atmosphere of K⁺ ions is strongly destabilized in the presence of Cs⁺ and TMA⁺ cations²².

Inspection of Figures 4-5a, b, and c reveals that the influence of TMA⁺ ions on the stability of the Pu22-T12T13 G-quadruplex is qualitatively similar to their effect on the stability of the c-MYC G-quadruplex. At each K⁺ concentration studied in this investigation, the Tₘ decreases sharply as the concentration of TMA⁺ ions increase from zero to ~200 mM and levels off at higher concentrations. On a similar note, an increase in Cs⁺ concentration from 0 to 100 mM also results in a sharp decline in Tₘ (see Figures 6a, b, and c).
The observed decreases in \( T_M \) of the preformed K\(^+\)-stabilized Pu22-T12T13 G-quadruplex with an increase in the concentration of nonstabilizing Cs\(^+\) and TMA\(^+\) cations unequivocally suggest that counterion condensation (the polyelectrolyte effect) destabilizes the G-quadruplex conformation. This notion implies a smaller number of counterions associated with the G-quadruplex relative to the coil state. The differential number of counterions condensed around the G-quadruplex and coil states can be determined from the initial slopes \((\delta T_M / \delta \ln[M^+])\) in Figures 4-5 and 4-6 using Eq. (2). For TMA\(^+\) ions, our calculated values of \( \Delta n_{M^+} \) are all equal to -0.6±0.2 for 2, 5, and 10 mM KCl. For Cs\(^+\) ions, the values of \( \Delta n_{M^+} \) are equal to -0.6±0.2, -0.8±0.2, and -0.9±0.2 for 2, 5, and 10 mM KCl (on average, -0.8±0.2). These numbers are qualitatively similar to -0.9±0.1 and -1.1±0.1, the differential number of TMA\(^+\) and Cs\(^+\) ions bound to the G-quadruplex and coil states of MYC22-G14T/G23T \(^{22}\). Thus, similar to MYC22-G14T/G23T, the destabilizing effect of counterion condensation originates from roughly one extra counterion condensed around the coil relative to the G-quadruplex state of Pu22-T12T13.

There is one more important quantitative feature that is common for the Pu22-T12T13 and MYC22-G14T/G23T G-quadruplexes. The \( T_M \)-versus-[TMA\(^+\)] dependences of the former and the \( T_M \)-versus-[TMA\(^+\)] and \( T_M \)-versus-[Cs\(^+\)] dependences of the latter all level off at ~200 mM \(^{22}\). We have interpreted this experimental result as suggesting that the local counterion concentration in the vicinity of the coil state of MYC22-G14T/G23T is on the order of ~200 mM \(^{22}\). We now extend this estimate to the Pu22-T12T13 sequence and propose that a local counterion concentration of 200 mM may be a common property of single stranded DNA conformations.

**Structural Considerations.** A smaller number of counterions associated with the G-quadruplex state is consistent with a lower charge density of the folded relative to the unfolded conformations of Pu22-T12T13. A similar notion supported by subsequent structural calculations has been put forward with respect to the folded and unfolded states of MYC22-G14T/G23T \(^{22}\). The Pu22-T12T13 and MYC22-G14T/G23T sequences are both 22-meric oligonucleotides that fold in the presence of potassium ions into a parallel G-quadruplex with two K\(^+\) ions sequestered in the central cavity. The volume, \( \rho_q \), and surface, \( \sigma_q \), charge densities of a G-quadruplex can be obtained by dividing the number of formal charges, N, by the molecular,
volume, $V_M$, and solvent accessible surface area, $S_A$, respectively. For the Pu22-T12T13 G-quadruplex, $N$ equals 20 (22 negative phosphate charged minus two positive charges of the sequestered potassium ions). Our calculated values of $V_M$ and $S_A$ of the Pu22-T12T13 G-quadruplex are $6180\pm38 \text{ Å}^3$ and $4476\pm33 \text{ Å}^2$, respectively. These values are close to those calculated for the parallel MYC22-G14T/G23T G-quadruplex ($V_M = 6178 \text{ Å}^3$ and $S_A = 4442 \text{ Å}^2$) but differ significantly from those calculated for the Na$^+$-stabilized Tel22 antiparallel G-quadruplex ($V_M = 6426 \text{ Å}^3$ and $S_A = 3907 \text{ Å}^2$). The latter is bulkier being, at the same time, less solvent-exposed compared to the parallel G-quadruplexes.

The volume, $\rho_q$, and surface, $\zeta_q$, charge densities of the Pu22-T12T13 G-quadruplex are equal to $(3.24\pm0.02) \times 10^{-3} \text{ e/Å}^3 (20 / 6180)$ and $(4.47\pm0.03) \times 10^{-3} \text{ e/Å}^2 (20 / 4476)$, respectively. These charge densities are similar to those of the MYC22-G14T/G23T G-quadruplex ($\rho_q = 3.24 \times 10^{-3} \text{ e/Å}^3$ and $\zeta_q = 4.50 \times 10^{-3} \text{ e/Å}^2$) but differ significantly from those of the Tel22 G-quadruplex ($\rho_q = 2.96 \times 10^{-3} \text{ e/Å}^3$ and $\zeta_q = 4.9 \times 10^{-3} \text{ e/Å}^2$). In fact, the parallel Pu22-T12T13 and MYC22-G14T/G23T G-quadruplexes exhibit the lowest charge density amongst the G-duplexes and polymeric duplexes analyzed in our previous study. This result is consistent with and accounts for the strongly destabilizing effect of counterion condensation observed for the parallel Pu22-T12T13 and MYC22-G14T/G23T G-quadruplexes.

### 4.6 Concluding Remarks

We characterize the structural and thermodynamic properties of a G-quadruplex formed by a modified 22-meric DNA sequence from a region proximal to the transcription initiation site of the human VEGF gene in the presence and absence of stabilizing and nonstabilizing cations. In a potassium solution, the Pu22-T12T13 sequence adopts a single parallel-stranded G-quadruplex conformation with a 1:4:1 loop-size arrangement. The thermal stability, $T_M$, of the Pu22-T12T13 G-quadruplex increases with the concentration of the stabilizing K$^+$ ions. In contrast, the G-quadruplex-to-coil transition temperatures, $T_M$, sharply decrease as nonstabilizing Cs$^+$ and TMA$^+$ ions are added to the solution of the preformed K$^+$-stabilized G-quadruplex. Stabilizing ions modulate the conformational preferences of guanine-rich DNA sequences by acting both as
internally bound ligands and as counterions forming the external ionic atmosphere of the folded (G-quadruplex) and unfolded (coil) states. In contrast, nonstabilizing ions may only act as counterions. With this notion, our results emphasize the opposing effects of internal binding and counterion condensation on the stability of the Pu22-T12T13 G-quadruplex. While centrally bound ions stabilize the G-quadruplex conformation, counterion condensation destabilizes it favoring the coil conformation. From the slope of the dependence of $T_M$ on the concentration of $\text{Cs}^+$ and $\text{TMA}^+$ cations, we estimate that the destabilizing effect of counterion condensation originates from roughly one extra counterion condensed around the coil relative to the G-quadruplex state of Pu22-T12T13. The reduced counterion binding by the G-quadruplex state of Pu22-T12T13 compared to the coil state is due to the low surface charge density of the G-quadruplex reflecting its structural characteristics. Taken together with the results of a previous study, our present results are consistent with the picture in which the differential effect of internally (stabilizing) and externally (destabilizing) bound cations is a common feature of parallel intramolecular G-quadruplexes. Additional studies are, however, needed to prove or refute the generality of this notion.

4.7 Supporting Information

Figure S1 presents a representative UV melting profile of Pu22-T12T13 at 295 nm obtained in the presence of 2 mM KCl. Figure S2 presents the time dependence of the molar ellipticity of Pu22-T12T13 at 264 nm following a rapid decrease in temperature from 95 to 25 °C in the presence of 5 mM KCl and 200 mM CsCl. This material is available free of charge via the internet at http://pubs.acs.org.
4.8 Acknowledgements

The authors are grateful to Prof. Robert B. Macgregor, Jr., for useful discussions and constructive comments. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to TVC.
4.9 Reference


CHAPTER 5: Thermodynamic Linkage Analysis of pH-induced Folding and Unfolding Transitions of i-motifs


Author’s Contribution: Byul G. Kim performed all experiments, analyzed the data, and participated in the preparation of the manuscript.
5.1 Abstract

We describe the pH-induced folding/unfolding transitions of i-motifs by a linkage thermodynamics-based formalism in terms of three pK_a’s of cytosines, namely, an apparent pK_a in the unfolded conformation, pK_{au}, and two apparent pK_a’s in the folded state, pK_{af1} and pK_{af2}. For the 5'-TTACCCACCCTACCCACCCTCA-3' sequence from the human c-MYC oncogene promoter region, the values of pK_{au}, pK_{af1}, and pK_{af2} are 4.8, 6.0, and 3.6, respectively. With these pK_a’s, we calculate the differential number of protons bound to the folded and unfolded states as a function of pH. Analysis along these lines offers an alternative interpretation to the experimentally observed shift in the pH-induced unfolded-to-i-motif transitions to neutral pH in the presence of cosolvents and crowders.
5.2 Introduction

At slightly acidic to near-neutral solution pH, cytosine-rich DNA sequences exhibit a propensity to fold into an i-motif conformation. The latter represents a noncanonical DNA conformation with a tetrameric arrangement of two parallel-stranded duplexes which intercalate into each other in an antiparallel orientation\(^1\)\(^-\)\(^4\). Significantly, i-motifs are believed to be of biological importance, since cytosine-rich sequences capable of folding into i-motifs have been found in important genomic loci including but not limited to centromeric and telomeric regions of human chromosomes and the promoter regions of numerous genes\(^2\).

The building blocks of an i-motif are made of hemiprotonated cytosine·cytosine\(^+\) (C·C\(^+\)) base pairs which are stabilized by an unusual arrangement of three hydrogen bonds and hold together the pairs of the parallel strands forming the i-motif structure\(^2\)\(^,\)\(^5\). As cytosine protonation is involved in i-motif formation, the stability of i-motif structures is strongly pH-dependent with i-motifs being favored by slightly acidic pH. The maximum of i-motif stability is generally observed between pH 4 and 5 which corresponds to the pK\(_a\) of the N3 nitrogens of cytosines in a free mononucleotide (and, presumably, in an unfolded DNA) of ~4.6\(^2\)\(^,\)\(^6\)-\(^9\). A further decrease in pH leads to protonation of the remaining cytosines with a subsequent disruption of the i-motif around ~pH 3\(^2\)\(^,\)\(^8\)\(^,\)\(^10\).

The pH-dependence of i-motif stability, while being well documented, has never been analyzed in a rigorous way afforded by linkage thermodynamics. In general, the dissociation constant, pK\(_a\), of a titrable group is modulated by its microenvironment and the nature and extent of interactions with other groups. Clearly, the pK\(_a\) of cytosines in an unfolded oligonucleotide and a folded i-motif should be significantly different. Moreover, there should be two distinct dissociation constants, pK\(_a\), for the two sets of protonated and deprotonated cytosines forming the hemiprotonated C·C\(^+\) base pairs in a folded i-motif. The favorable change in free energy, \(\Delta G^o\), accompanying i-motif formation will drive the protonation of the initial set of cytosines in the hemiprotonated C·C\(^+\) base pairs to pH values higher than ~4.6, the pK\(_a\) of unfolded cytosines. Following the same logic, the protonation of the second set of cytosines (initially deprotonated in the i-motif conformation) will be suppressed and take place below pH 4.6.
With these comments, we analyze below the pH-induced folding/unfolding transition profile of a representative i-motif formed by a 22-meric oligonucleotide (5′-TTACCCACCCTACCCACCCTCA-3′) from the human c-MYC oncogene promoter region (ODN). The sequence is complementary to the guanine-rich G-quadruplex-forming sequence from the c-MYC promoter site that has been characterized in our recent work\textsuperscript{11}. Cytosine-rich sequences from the promoter region of c-MYC have been shown to fold into i-motif conformations at slightly acidic pH\textsuperscript{1,7,10,12-15}.

Figure 5-1 presents the CD spectra of the ODN recorded at pH ranging from 9.5 to 4.5 (panel A) and from 4.8 to 2.3 (panel B). Inspection of Figures 5-1a and b reveals several important observations. Firstly, a decrease in pH to ~5 results in formation of an i-motif structure as can be judged by the CD spectrum with a characteristic maximum at 285 nm and a minimum at 258 nm (see Figure 1a)\textsuperscript{2,9,10,16}. Secondly, the isoeilliptic points observed at 246 and 276 nm in Figure 1a suggest that the pH-induced i-motif folding is likely a two-state transition and does not involve populated intermediates. Thirdly, a decrease in pH to below ~3 results in an unfolding of the i-motif (see Figure 5-1b). Finally, isoeilliptic points at 233 and 271 nm in Figure 5-1b are consistent with the two-state nature of the second pH-induced transition (i-motif-to-unfolded transition). Given the isoeilliptic points in Figures 5-1a and b, we analyze the pH-induced folding/unfolding transitions of the ODN within the framework of a two-state model with equilibrium constant, $K = [F]/[U]$ (where [F] and [U] are concentrations of the folded and unfolded DNA states, respectively). A note of caution is warranted here. The observed isoeilliptic points are consistent with but cannot be viewed as an absolute proof of the two-state nature of the folding and unfolding transitions. While, in this work, we treat our data in a two-state manner, further studies are required to clarify the nature of the pH-induced folding/unfolding transitions of i-motifs, in general, and the ODN, in particular.
Figure 5-1. The CD spectra of the ODN between pH 9.5 and 4.5 (panel A) and pH 4.8 and 2.3 (panel B).
According to the thermodynamic linkage theory, the pH-dependence of the equilibrium constant, $K$, is given by the relationship:\(^{17-19}\):

$$
K = K_0 \frac{\prod_{i=1}^{6}(1+10^{pK_{af} - pH})}{\prod_{i=1}^{12}(1+10^{pK_{au} - pH})}
$$

where $K_0$ is the equilibrium constant in the absence of protons (at alkaline pH); $pK_{af}$ is the $pK_a$ of the $i$-th cytosine in the folded (i-motif) conformation; and $pK_{au}$ is the $pK_a$ of the same cytosine in the unfolded conformation. The specific values of $pK_{af}$ and $pK_{au}$ may depend on the internal and external factors. The former reflect the sequence and structure of the i-motif-forming DNA and the position of the specific cytosine in it, while the latter reflect temperature, pressure, and solution composition, including the presence of cosolvents and salts.

Figure 5-2 presents the pH-induced folding/unfolding profile of the ODN as traced by the molar ellipticity at 284 nm. Inspection of Figure 5-2 reveals two near-sigmoidal transitions centered at pH 6.1 and 3.6 representing the unfolded-to-i-motif and i-motif-to-unfolded transitions, respectively. We simplify Eq. (1) to fit the experimental data plotted in Figure 5-2. To this end, we assume that the cytosines of the ODN are uniformly characterized by a set of three distinct $pK_a$’s (in reality, the individual $pK_a$’s of cytosines may be distinct depending on their position). In the unfolded conformation, all 12 cytosines are assumed to exhibit the same dissociation constant of $pK_{au}$. Further, we assume that, in the i-motif conformation, the six initially protonating cytosines exhibit a higher dissociation constant of $pK_{af1}$, while the remaining six cytosines are characterized by a lower $pK_a$ of $pK_{af2}$. With these approximations, Eq. (1) simplifies to the form:

$$
K = K_0 \frac{(1+10^{pK_{af1} - pH})^6(1+10^{pK_{af2} - pH})^6}{(1+10^{pK_{au} - pH})^{12}}
$$

(2)
We used Eq. (2) to fit the experimental data points shown in Figure 5-2. The fit produced pK$_{au}$, pK$_{af1}$, and pK$_{af2}$ of 4.8±0.1, 6.0±0.1, and 3.6±0.1, respectively, with $K_0$ equal to 0.07±0.05. The equilibrium constant of 0.07±0.05 is rather large and consistent with a picture in which an i-motif population of a few percent persists even at alkaline pH with virtually no charged (protonated) cytosines. This notion is suggestive of the existence of highly unstable i-motifs containing unprotonated C·C base pairs. Alternatively, there is a possibility that this quantitative result is an artifact reflecting the approximations involved in deriving Eq. (2). Clearly, further studies are required to clarify this intriguing and important point.

**Figure 5-2.** The pH-dependence of the molar ellipticity at 284 nm. Experimental data were approximated with Eq. (2) (solid line).
The values of $pK_{au}$, $pK_{af1}$, and $pK_{af2}$ we obtained can be used to calculate the differential number of protons bound to the folded (i-motif) state relative to the unfolded state:

$$\Delta n_{H^+}(\text{pH}) = \frac{N_{cyt}}{2} \left( \frac{1}{1 + 10^{\text{pH}-pK_{af1}}} + \frac{1}{1 + 10^{\text{pH}-pK_{af2}}} - \frac{2}{1 + 10^{\text{pH}-pK_{au}}} \right)$$  \hspace{1cm} (3)

where $N_{cyt}$ is the total number of cytosines participating in i-motif formation.

Figure 5-3 plots the pH-dependence of the differential number of bound protons, $\Delta n_{H^+}$, computed with Eq. (3). Inspection of Figure 3 reveals that, as the pH decreases from neutral to acidic, $\Delta n_{H^+}$ initially increases from 0 to 2.7 (at ~pH 5.6), then decreases to -2.7 (at ~pH 4.0) crossing 0 at ~pH 4.8, and, finally, increases to zero at ~pH 2. Results of this simulation are in qualitative agreement with the results of previous works in which $\Delta n_{H^+}$ are treated as constants and determined from the slope ($\delta T_M/\delta \text{pH}$) \(^7,9\).
Eq. (2) yields valuable insights into an increase in the midpoint of the pH-induced i-motif formation to ~pH 7 or even higher in the presence of cosolvents and crowders\textsuperscript{12, 20, 21}. As an explanation, it has been proposed that the pK\textsubscript{a} of cytosines increases in the presence of cosolvents\textsuperscript{20} which, indeed, may contribute to the observed increase in the transition midpoint. In fact, all three pK\textsubscript{a}’s in Eq. (2) may be affected. As an alternative or supplemental explanation, we propose that, given the excluded volume effect, cosolvents and crowders may stabilize the compact i-motif state relative to the extended unfolded state with an ensuing increase in the equilibrium constant $K_0$. To illustrate this point, we used Eq. (2) to simulate the transition profiles of an i-motif for different values of $K_0$. Specifically, we calculate the fraction folded, $\alpha = K/(1 + K)$, as a function of pH. Figure 5-4 presents these profiles calculated for $pK_{au}$, $pK_{af1}$, and $pK_{af2}$ of 4.8, 6.0, and 3.6, respectively, and $K_0$ equal to 0.0001, 0.001, 0.005, 0.02, and 0.1.

**Figure 5-3.** The pH-dependence of the differential number of protons bound to the folded and unfolded states of the ODN calculated with Eq. (3).
Inspection of Figure 4 reveals that a decrease in $K_0$ below 0.005 leads to incomplete i-motif folding. At $K_0$ equal to 0.005, the folding is nearly complete with $\alpha \sim 98\%$ at pH 4.8 and a transition midpoint at pH 5.7. A 20-fold increase in $K_0$ from 0.005 to 0.1 ($\Delta \Delta G^\circ = \sim 1.8 \text{ kcal mol}^{-1}$) leads to a shift in the transition midpoint from pH 5.7 to 6.3. This observation readily offers an alternative rationalization to the effect of cosolvents and crowders on i-motif folding transitions.

**Figure 5-4.** Transition profiles simulated with Eq. (2) for $pK_{au}$, $pK_{af1}$, and $pK_{af2}$ of 4.8, 6.0, and 3.6, respectively, and $K_0$ equal to 0.0001 (dark yellow), 0.001 (green), 0.005 (blue), 0.02 (black), and 0.1 (red).

It should be noted that an increase in $K_0$ may be caused not only by environmental factors but also by modifications in structure and sequence, e. g. the number of C–C+ base pairs or loop sequence and length$^{2, 22}$. Clearly, such modulations of i-motif stability will be reflected in the
corresponding shift in the midpoint of the pH-induced folding/unfolding transition. As a final note, an increase in the number of C·C\(^+\) base pairs alone will cause an increase the degree of folding and shift the transition midpoint to a higher pH even in the absence of any change in K\(_0\). This notion is demonstrated in Figure 5-5 that plots the transition profiles of two i-motifs with 6 (black) and 8 (red) C·C\(^+\) base pairs with the same pK\(_a\)'s of cytosines and the same equilibrium constant, K\(_0\), of 0.0001. Inspection of Figure 5-5 reveals that the presence of two extra C·C\(^+\) base pairs leads to an increases in the maximum of the fraction folded, \(\alpha\), from 44 to 94 % while also shifting the folding transition midpoint to a higher pH.

![Figure 5-5](image)

**Figure 5-5.** Transition profiles simulated with Eq. (2) for i-motifs containing 6 (black) and 8 (red) C·C\(^+\) base pairs with pK\(_{au}\), pK\(_{af1}\), and pK\(_{af2}\) of 4.8, 6.0, and 3.6, respectively, and K\(_0\) of 0.0001.

In summary, the pH-induced folding/unfolding transitions of i-motifs can be effectively described by a linkage thermodynamics-based formalism in terms of three cytosine pK\(_a\)'s - the
apparent $pK_a$ of cytosines in the unfolded conformation, $pK_{au}$, and two apparent $pK_a$’s of cytosines in the folded conformation, $pK_{af1}$ and $pK_{af2}$. For a specific case of the 5’-TTACCCACCCTACCCACCCTCA-3’ sequence from the human c-MYC oncogene promoter region, the values of $pK_{au}$, $pK_{af1}$, and $pK_{af2}$ are 4.8, 6.0, and 3.6, respectively. These $pK_a$’s enable one to calculate the differential number of protons associated with the folded and unfolded states as a function of pH. Analysis along these lines offers an alternative interpretation to the experimentally observed shift in the pH-induced unfolded-to-i-motif transitions to neutral pH in the presence of cosolvents and crowders. Our simulations reveal that a significant increase in the transition midpoint pH can be achieved by an increase in the equilibrium constant between the folded and unfolded DNA conformations due to the excluded volume effect.

5.3 Experimental Section

The oligomeric sequence 5’-TTACCCACCCTACCCACCCTCA-3’ was purchased from Integrated DNA Technologies (Coralville, IA). Sodium chloride was obtained from Sigma-Aldrich Canada (Oakville, ON). EDTA (free acid) was purchased from Fisher Biotech (Fair Lawn, NJ). The DNA was initially dissolved in and dialyzed against an unbuffered solution containing 10 mM NaCl and 0.1 mM EDTA rather than a buffer to avoid complications related to the ionization-neutralization equilibria of the buffer in pH-titration experiments. Dialysis was carried out with 1000 Da molecular mass cut off Tube-O-Dialyzers from G Biosciences (St. Louis, MO). The concentrations of the oligonucleotide were determined spectrophotometrically at 25 °C with a molar extinction coefficient, $\varepsilon_{260}$, of 191,500 M$^{-1}$cm$^{-1}$ for the unfolded conformation. The latter was evaluated using a nearest neighbor procedure as described by Dr. Richard Owczarzy (http://www.owczarzy.net/extinctionDNA.htm). The UV light absorption spectral measurements were carried out with a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, ON).

The circular dichroism (CD) spectra of the DNA were recorded at 25 °C in a 1 mm path-length cuvette using an Aviv model 62 DS spectropolarimeter (Aviv Associates, Lakewood, NJ). CD spectroscopic measurements were used for conformational characterization of the ODN
under each solution condition employed in this study. CD pH titration experiments were performed by adding to the optical cell containing 0.25 mL of the ODN solution aliquots of HCl solution. The initial volume was delivered by a 1 mL Hamilton Syringe, while titrant solutions (50, 100, and 500 mM HCl) were added with 10 μL Hamilton syringes (Hamilton Co. Reno, NV). The syringes were equipped with Chaney adapter that allowed a relative delivery accuracy of ± 0.1%. The pH values were measured separately with four-fold larger volumes of the ODN solutions and the added HCl aliquots using a VWR brand Benchtop model 8015 pH-meter equipped with an Accumet Ag/AgCl combination microelectrode. The absolute error in these pH measurements was ±0.01 pH unit.

5.4 Acknowledgements

The authors are grateful to Prof. Robert B. Macgregor, Jr., for useful discussions and constructive comments. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to TVC.
5.5 Reference


CHAPTER 6: General Conclusion and Future Perspectives

6.1 Conclusions

As non-canonical DNA structures arise as potential therapeutic targets, understanding the factors that govern their stability and folding/unfolding transitions becomes increasingly important. G-quadruplex often occurs in important regulatory regions of the genome and acts to inhibit uncontrolled telomere elongation and oncogene expression. A unique characteristic that distinguishes a G-quadruplex from other DNA structures is that it requires specifically bound cations to be formed. Counterions of specific size become coordinated to negatively charged O6 atoms of the guanines within the central cavity of a G-quadruplex. Another tetraplex structure that is often studied in conjunction with G-quadruplexes is i-motif. In the genome, G-quadruplex-forming guanine-rich strands are complementary to strands rich in cytosine. This C-rich strand may form i-motif, a distinct structure that consists of hemi-protonated cytosine·cytosine\(^{+}\) base pairs that are, generally, formed under slightly acidic conditions. My doctoral thesis was devoted to understanding the thermodynamics of the stability of these tetraplex structures. Three projects of my thesis examined the salt effect in various G-quadruplex structures and one paper investigated the effect of pH on i-motif formation using the thermodynamic linkage analysis.

Chapters 2, 3, and 4 explore the influence of the polyelectrolyte effect on the stabilization of G-quadruplexes formed in telomeric and oncogene promoter regions of the genome. I have conducted a series of spectroscopic experiments aimed at characterizing the stability of various G-quadruplexes in solutions with varying concentrations of stabilizing (Na\(^{+}\) or K\(^{+}\)) and non-stabilizing (Li\(^{+}\), TMA\(^{+}\), Cs\(^{+}\)) counterions. The specific and non-specific ionic contributions to the thermal stability of G-quadruplexes were evaluated by separating the effect of the centrally bound ions from that of cations condensed around the DNA. The thermal stability of G-quadruplexes is known to increase with the concentration of stabilizing cations; however, the effects of specific and non-specific cation binding have not been separated\(^1\). We devised a method to discriminate between the effects of the stabilizing cations in the central cavity are counterions condensed in the vicinity of the DNA. The method used involved an initial
saturation of the G-quadruplex by the stabilizing ions (Na\(^+\) or K\(^+\)) followed by addition of either stabilizing or non-stabilizing cations (Li\(^+\), Cs\(^+\), TMA\(^+\)). Our data suggest that the predominant ionic contribution to G-quadruplex stability comes from the specifically bound Na\(^+\) or K\(^+\) ions and not from counterion condensation around the DNA. This is in contrast to the thermodynamics of duplex and T-rich triplex DNA where counterion condensation provides a major contribution to the conformational stability\(^2\)\(^-\)\(^3\). Our results clarified the balance of forces governing the conformational preferences of guanine-rich sequences while also being of practical importance for controlled induction of G-quadruplexes in the genome.

Following the study of the salt effect in G-quadruplex structures, we focused on understanding the molecular basis of i-motif stability. pH plays an important role in modulating the i-motif stability. Hence we conducted pH-dependent titration studies with the c-MYC i-motif and performed a thermodynamic linkage analysis of the pH-induced folding and unfolding transitions.
6.2 Future Prospect

Though the work for my dissertation provided new insight about the environment-dependent stabilities of tetraplex structures, it had several limitations. First, our studies hinted at the generality of the differential effect of internally (stabilizing) and externally (destabilizing) bound cations in parallel G-quadruplexes. However, experiments must be conducted with more oligomers derived from other parts of the genome to test if the results are broadly applicable. The second limitation is that the projects examining G-quadruplexes and i-motif separately do not perfectly reflect the physiological environment. In the genome, G-quadruplex structures are formed in the presence of complementary strands (with the exception of telomeric 3'-overhangs of eukaryotes). G-quadruplexes in gene promoter regions are constrained by the duplex nature of genomic DNA. Thus, an important question is how secondary structures such as G-quadruplexes energetically compete with duplex DNA, in particular, since putative GC-rich sequences that form these structures are thermally and thermodynamically more stable than AT-rich regions. Importantly, limiting the environmental factors simply to changes in salt and pH does not provide a big picture because there are many other molecules in the cell that can modulate the DNA-water interactions.

Salt Effects in G-quadruplexes derived from other regions of DNA

We have investigated the polyelectrolyte effect in four different G-quadruplex oligomers. The results showed that non-stabilizing cations can have no effect (tel-22 sequence), slightly stabilizing (tel-26), or destabilizing (VEGF and c-MYC) effect on G-quadruplex structures. It is suspected that the differing effects stem from the fact that the conformations of the G-quadruplexes used are different. However, more studies on G-quadruplex with varying topology must be conducted to draw concrete conclusions about the generalizability of our experimental results obtained.

Duplex-Single strand-Tetraplex Equilibria

Sugimoto and collaborators have shown that G-quadruplexes compete with duplexes depending on the cations and pH values. We are presently investigating the duplex-quadruplex
equilibria transitions using the complementary strands for a subset of our original G-rich sequences. We are modulating the environmental conditions including temperature, pH, and salt in order to observe how the equilibria between various conformations change as a function of solution conditions.

Effect of osmolytes and cosolvents

Characterizing the interactions of osmolytes and other cosolvents with proteins and nucleic acids represents an area of significant interest in our laboratory. This study is essential both for understanding the modulation of biological function and developing methods to stabilize specific conformational states of biopolymers. We have recently begun a cycle of experimental and theoretical works in which we extend volumetric measurements to studying the solvation properties of solutes in binary solvents. Our lab has applied this approach to characterizing the interactions of protein groups with urea and glycine betaine. We are extending this knowledge to nucleic acid structures, specifically to G-quadruplexes in both telomeric and non-telomeric sequences.
6.3 Reference


APPENDIX A: Effect of Urea on G-quadruplex Stability


Author’s Contribution: Byul G. Kim organized and analyzed the data, participated in the preparation of the manuscript, and trained the undergraduate student Jordan Ko to run the experiments.
A.1 Abstract

G-quadruplexes represent a class of noncanonical nucleic acid structures implicated in transcriptional regulation, cellular function, and disease. An understanding of the forces involved in stabilization/destabilization of the G-quadruplex conformation relative to the duplex or single-stranded conformation is a key to elucidating the biological role of G-quadruplex-based genomic switches and the quest for therapeutic means for controlled induction or suppression of a G-quadruplex at selected genomic loci. Solute-solvent interactions provide a ubiquitous and, in many cases, the determining thermodynamic force in maintaining and modulating the stability of nucleic acids. These interactions involve water as well as water-soluble cosolvents that may be present in the solution or in the crowded environment in the cell. We present here the first quantitative investigation of the effect of urea, a destabilizing cosolvent, on the conformational preferences of a G-quadruplex formed by the telomeric d[A(G3T2A)3G3] sequence (Tel22). At 20 mM NaCl and room temperature, Tel22 undergoes a two-state urea-induced unfolding transition. An increase in salt mitigates the deleterious effect of urea on Tel22. The urea m-value of Tel22 normalized per change in solvent-accessible surface area, ΔSA, is similar to those for other DNA and RNA structures while being several-fold larger than that of proteins. Our results suggest that urea can be employed as an analytical tool in thermodynamic characterizations of G-quadruplexes in a manner similar to the use of urea in protein studies. We emphasize the need for further studies involving a larger selection of G-quadruplexes varying in sequence, topology (parallel, antiparallel, hybrid), and molecularity (monomolecular, bimolecular, tetramolecular) to outline the advantages and the limits of the use of urea in G-quadruplex studies. A deeper understanding of the effect of solvent and cosolvents on the differential stability of the G-quadruplex and duplex conformations is a step towards elucidation of the modulating influence of different types of cosolvents on duplex-G-quadruplex molecular switches triggering genomic events.
A.2 Introduction

In a suitable environment, guanine-rich DNA and RNA sequences fold into one of the various G-quadruplex conformations in which guanines are collected into stacks of G-quartets\(^1\)\(^-\)\(^7\). A G-quartet is a planar structure formed by guanine bases pairwise-interlinked by Hoogsteen hydrogen bonds and additionally stabilized by a mono- or divalent cation of a proper size coordinated to the O6 carbonyls of the guanines within the central cavity of the core of a G-quadruplex\(^1\),\(^3\)\(^-\)\(^8\). Guanine-rich sequences potentially capable of folding into G-quadruplex have been identified in hundreds of thousand loci in the human genome including telomeres, centromeres, and promoter regions of oncogenes with clear connections between G-quadruplexes and cellular function and disease\(^2\),\(^8\)\(^-\)\(^14\). Human telomeric DNA consists of tandem repeats of the 5'-TTAGGG-3' sequence with a 3'-end single-stranded overhang of 100-200 nucleotides\(^7\),\(^11\),\(^15\). The structural, kinetic, and thermodynamic properties of monomolecular G-quadruplexes containing four repeats of 5'-TTAGGG-3' have been extensively studied\(^5\),\(^16\)\(^-\)\(^27\). In the atmosphere of Na\(^+\) ions, the 22-meric telomeric sequence d[A(G\(_3\)T\(_2\)A)\(_3\)G\(_3\)] (Tel22) forms a basket structure with antiparallel strands and a diagonal and two lateral loops\(^16\). In an aqueous solution, the Tel22 G-quadruplex is fully formed at ~20 mM NaCl and becomes further stabilized with an increase in the concentration of Na\(^+\) ions\(^25\).

Among all interactions governing the stability and conformational preferences of G-quadruplex-forming DNA sequences, solute-solvent interactions (solvation) stand out as a ubiquitous thermodynamic force prominently present in all major structural transitions involving nucleic acid structures (e.g., duplex-to-single strand) while also participating in fine tuning of more subtle transitions between conformational substates (e.g., A-to-B duplex transition or transitions between G-quadruplex topoisomers)\(^28\)\(^-\)\(^32\). Solvation should be viewed to include not only solute-water interactions but also interactions between DNA and water-miscible cosolvents that may be present in solution. The conformational preferences of guanine-rich DNA sequences have been shown to be strongly modulated by water-soluble organic compounds\(^30\)\(^-\)\(^31\),\(^33\)\(^-\)\(^35\).

Urea is a quintessential, widely used cosolvent with a strongly denaturing action on both proteins and nucleic acids\(^36\)\(^-\)\(^37\). It has been employed for over a century in protein studies as a
means to disturb and direct folded/unfolded equilibria. Remarkable progress has been made in developing an understanding of the action of urea on protein stability at the molecular level. Studies of the effect of urea on the stability of nucleic acid structures have been relatively scarce compared to proteins. While the influence of urea on the stability of DNA and RNA structures have been studied to some extent, quantitative investigations of the effect of urea on conformational equilibria involving G-quadruplexes are lacking. The deficiency is unfortunate for a number of reasons. Firstly, G-quadruplexes are unusual nucleic acid structures sharing a number of common features with globular proteins; they are globular in shape, less charged compared to other nucleic acids, and compressible owing to the presence of the central cavity. Thus, urea may interact with G-quadruplexes in a manner that is distinct compared to canonical DNA and RNA structures. Secondly, given the polymorphism of G-quadruplexes, urea may prove to be an effective and subtle tool in modulating the equilibria between various G-quadruplex conformations. Thirdly, an understanding of urea-G-quadruplex interactions may form the basis for subsequent studies of molecular forces underlying the duplex-to-G-quadruplex and single strand-to-G-quadruplex conformational switches within the genome in the crowded environment of the cell. Finally, urea is a natural cosolvent present in large concentrations in various living organisms which makes its effect on G-quadruplexes of biological significance.

In this work, we explore the conformational states of Tel22 as a function of temperature and the concentrations of urea and sodium ions. Our results represent the first systematic investigation of the stability of a G-quadruplex within the temperature-urea-salt conformational space, while also producing the necessary experimental basis for further investigations of the preferential interactions of this important cosolvent with G-quadruplex-forming DNA sequences in their folded and unfolded states.

A.3 Materials and Methods

Materials. The 22-mer, d[A(G3T2A)3G3] (Tel22), containing four repeats of the human telomeric DNA sequence was synthesized and purified by Integrated DNA Technologies (Coralville, IA). Sodium chloride, phosphoric acid, and urea were purchased from Sigma-Aldrich Canada.
(Oakville, ON, Canada). EDTA (free acid) was purchased from Fisher Biotech (Fair Lawn, NJ, USA). These reagents were used without further purification. All solutions were prepared using doubly distilled water.

All measurements were performed in a pH 7.0 buffer containing 10 mM monosodium phosphate/disodium phosphate, 0.1 mM EDTA, NaCl between 20 and 1000 mM, and urea at concentrations between 0 and 8 M. Solutions of urea with concentrations of 1, 2, 4, 6, and 8 M were prepared by weighing urea and adding pre-estimated amounts of the buffer to achieve the desired molalities, m. The molar concentration, C, of urea was computed from its molal concentration, m, using $C = \frac{1}{m \rho_b} + \frac{\phi V}{1000}$, where $\rho_b$ is the density of urea-less phosphate buffer and $\phi V = 44.1$ cm$^3$mol$^{-1}$ is the apparent molar volume of urea$^{52,53}$. Urea solutions were subsequently used as solvents for the DNA solutions.

The DNA was dissolved in and exhaustively dialyzed against a pH 7.0 phosphate buffer with NaCl and urea at desired concentrations. The dialysis was carried out in 1000 Da molecular weight cut-off Tube-O-Dialyzers from G Biosciences (St. Louis, MO, USA). To form the equilibrium structure, the DNA-containing solution was placed in a boiling water bath and allowed to cool to room temperature.

The concentration of the DNA was determined from the absorbance at 260 nm measured at 25 °C using a molar extinction coefficient of 228,500 M$^{-1}$ cm$^{-1}$ for the unfolded conformation. The latter was calculated using an additive nearest neighbor procedure described by Owczarzy and colleagues$^{54}$. For circular dichroism (CD) measurements and temperature-dependent UV light absorption measurements, the DNA concentrations were ~30 and 3 μM, respectively.

**Circular Dichroism Spectroscopy.** CD spectroscopy was used to probe the DNA conformation at each of the experimental conditions of this study. The CD spectra of Tel22 were recorded in a 1 mm path-length cuvette at 25 °C using an Aviv model 62 DS spectropolarimeter (Aviv Associates, Lakewood, NJ) or an Olis DSM 20 CD spectrophotometer (Olis, Inc., Bogart, GA).
UV Melting Experiments. UV light absorption at 295 nm was measured as a function of temperature in a DNA sample contained in a 1 cm path-length cuvette. These measurements were performed using a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, Ontario, Canada) or a Lambda 800 spectrophotometer (Perkin Elmer, Inc., Waltham, MA). The temperature was changed at a rate of 1 °C per minute. The G-quadruplex-to-coil transition temperatures, \( T_M \), and the van't Hoff enthalpies, \( \Delta H_vH \), were evaluated from the experimental UV melting profiles using standard procedures \(^{55-59}\).

A.4 Results

Figure A-1 presents the CD spectra of Tel22 at 25 °C in the absence of urea at NaCl concentrations between 20 and 1000 mM NaCl. In the absence of urea, Tel22 exhibits a CD spectrum, characteristic of the antiparallel G-quadruplex conformation with a positive peak at 295 nm and a negative peak at 260 nm \(^5,60-61\). The CD spectrum does not change when the concentration of NaCl increases from 20 to 1000 mM.
Figure A-1. The CD spectra of Tel22 at 25 °C in the absence of urea at various NaCl concentrations. The NaCl concentration (in mM) for each spectrum is given in the inset.

Figures A-2a, b, c, d, e, f, g, and h show the CD spectra of Tel22 at urea concentrations between 0 and 8 M at 20, 50, 100, 200, 400, 600, 800, and 1000 mM NaCl, respectively. As is seen from Figure A-2a, at 20 mM NaCl, Tel22 undergoes a urea-induced conformational transition as suggested by changes in the positions and the amplitudes of the original CD spectral peaks. However, an increase in salt leads to complete (at urea concentrations of up to 4 M) or partial (at 6 and 8 M urea) refolding of unfolded Tel22 (see Figures 2b to h).
APPENDIX A: Effect of Urea on G-quadruplex Stability

b) 50 mM NaCl

\[ [\theta], \text{deg M}^{-1}\text{cm}^{-1} \]
\[ \lambda, \text{nm} \]

- [ ] 0 M
- [ ] 1 M
- [ ] 2 M
- [ ] 4 M
- [ ] 6 M
- [ ] 8 M

100 mM NaCl

\[ [\theta], \text{deg M}^{-1}\text{cm}^{-1} \]
\[ \lambda, \text{nm} \]
APPENDIX A: Effect of Urea on G-quadruplex Stability

![Graph d) 200 mM NaCl](image)

![Graph e) 400 mM NaCl](image)
Figure A-2. The CD spectra of Tel22 at 25 °C at various urea concentrations at 20 (panel a), 50 (panel b), 100 (panel c); 200 (panel d); 400 (panel e); 600 (panel f); 800 (panel g); and 1000 (panel h) mM NaCl. The urea concentrations (in M) for the spectra are given in the insets.

Figure A-3 compares the CD spectra of the urea-induced unfolded state (at 20 mM NaCl and 8 M urea) and the temperature-induced (at 20 mM NaCl and 95 °C) unfolded states of Tel22. The two states are clearly distinct with the urea-induced unfolded state being more structured than the heat-induced unfolded state.
Figure A-3. The CD spectra of the urea-induced unfolded state at 20 mM NaCl and 8 M urea (○) and temperature-induced unfolded states at 20 mM NaCl and 95 °C in the absence of urea (■) of Tel22. The inset presents the temperature dependence of the molar ellipticity at 257 nm of the urea-induced unfolded state at 20 mM NaCl and 8 M urea.

Figure A-4 shows a representative UV melting profile of Tel22 recorded at 295 nm in a solution containing 200 mM NaCl and 1 M urea. All UV melting profiles have been approximated by the two-state model of thermal denaturation:

\[
A = A_N(T) + [A_D(T) - A_N(T)]/[1 + \exp(\Delta H_M(T^{-1} - T_M^{-1})/R)]
\]

where \(T_M\) and \(\Delta H_M\) are the transition temperature and van’t Hoff enthalpy, respectively; and \(A_N(T)\) and \(A_D(T)\) are linear functions that approximate the pre- and post-denaturation baselines, respectively.
Figure A-4. Representative UV melting profile at 295 nm Tel22 at 200 mM NaCl and 1 M urea. Experimental data were approximated with Eq. (1) (solid line).

Figure A-5 plots $T_M$ as a function of urea at various NaCl concentrations. Inspection of Figures A-5 reveals a steady decrease in $T_M$ as the concentration of urea increases, while $T_M$ increases with an increase in salt concentration.
Figure A-5. The G-quadruplex-to-coil transition temperature, $T_M$, transition temperatures of Tel22 as a function of urea at various NaCl concentrations. The NaCl concentration (in mM) for each set of $T_M$ is given in the inset.

Figure 6 presents a compilation of van’t Hoff transition enthalpies, $\Delta H_M$, plotted against respective transition temperatures, $T_M$. Linear regression analysis of the data plotted in Figure 6 yields a slope, $\Delta \Delta H_M / \Delta T_M$, of 0.14±0.04 kcal mol$^{-1}$K$^{-1}$ which may represent a change in heat capacity, $\Delta C_P$, associated with the heat-induced denaturation of Tel22. In fact, it has been reported that thermal unfolding of some G-quadruplexes is accompanied by an increase in heat capacity$^{62}$. On the other hand, our calorimetric measurements have not produced any significant changes in heat capacity accompanying thermal unfolding of the K$^+$-stabilized Tel26 telomeric G-quadruplex in the hybrid-1 conformation$^{26}$. 
A.5 Discussion

*Urea-induced unfolding of Tel22 is a two-state process.* Inspection of Figures 2a to h reveals that, at each NaCl concentration employed in this study, an increase in urea leads to a conformational alteration of Tel22 as manifested in a decrease in the amplitude of the positive peak at 295 nm, a redshift and a significant decrease in the amplitude of the negative peak at 264 nm, and a redshift and a modest decrease in the amplitude of the positive peak at 246 nm. Comparative analysis of the CD spectra presented in Figures 2a to h reveals that the denaturing effect of urea weakens as the concentration of NaCl in solution increases. This observation clearly reflects an increase in G-quadruplex stability accompanying an increase in the concentration of stabilizing ions (in this case, Na⁺ ions). Thus, at increasing NaCl concentrations, increasingly higher concentrations of urea are required to unfold the Tel22 G-quadruplex.

Further inspection of Figures 2a to h reveals the presence of nearly isoelliptic points at 252 and 275 nm. This observation is suggestive of a two-state nature of the urea-induced unfolding transition of the Tel22 G-quadruplex. In contrast to urea-induced unfolding of the Tel22 G-quadruplex, its heat-induced unfolding is not a two-state event but involves intermediates. Given the two-state nature of the urea-induced transition, it may be treated within the framework of the linear extrapolation model (LEM) in a manner similar to that applied to urea-induced denaturation of proteins. Figure 7 presents the urea dependence of the molar ellipticity at 295 nm of Tel22 at 20 mM NaCl. The experimental points were approximated by the LEM-based two-state analytical model in which the differential free energy, ΔG°, of the folded and unfolded states is given by the relationship:

\[ ΔG° = -RT\ln K = ΔG_0 - m[\text{urea}] \] (2)
where $K = \alpha/(1 - \alpha)$ is the folded/unfolded equilibrium constant; $\alpha$ is the fraction unfolded; $\Delta G_0$ is the differential free energy of the folded and unfolded states in the absence of urea; and $m$ is the so-called “m-value”, a proportionality coefficient.

Fitting the data in Figure A-7 with Eq. (2) yields the values of $\Delta G_0$ and $m$ of $2.5 \pm 0.4$ kcal mol$^{-1}$ and $0.57 \pm 0.09$ kcal mol$^{-1}$M$^{-1}$, respectively. A more conventional way of determining the differential free energy, $\Delta G_0$, is based on the thermodynamic parameters (the melting temperature, $T_M$, and enthalpy, $\Delta H_M$) of the UV melting profile; $\Delta G_0 = \Delta H_M(1 – T/T_M)$. At $20$ mM NaCl and in the absence of urea, $T_M$ and $\Delta H_M$ are equal to $49.6 \pm 0.1 \, ^\circ C$ and $38.1 \pm 0.6$ kcal mol$^{-1}$, respectively. With these numbers, we calculate at $25 \, ^\circ C$ a $\Delta G_0$ of $2.9 \pm 0.6$ kcal mol$^{-1}$ in good agreement with the urea-unfolding-based value of $2.5 \pm 0.4$ kcal mol$^{-1}$. The agreement is significant and suggests that the LEM analysis of urea-induced unfolding profiles of G-quadruplexes provides an alternative way of characterizing G-quadruplex stability.

**Figure A-7.** The dependence of the molar ellipticity of the Tel22 at $295$ nm on the urea concentration at $20$ mM NaCl. Experimental data were approximated with Eq. (2) (solid line).
Our determined m-value of 0.57±0.09 kcal mol⁻¹M⁻¹ is within the range of similar values determined for other nucleic acid structures. For example, the urea m-values for the formation of a number of 12-meric DNA and RNA duplexes range between 0.58 and 0.98 kcal mol⁻¹M⁻¹.

The nature of the urea-induced unfolded state. The CD spectrum of the urea-induced unfolded state of Tel22 at 20 mM NaCl and 8 M urea with a positive peak at 258 nm and a negative peak at 278 nm (see Figure A-2a) is similar to that the unfolded state Tel22 adopts in the atmosphere of tetrabutylammonium ions in the absence of stabilizing Na⁺ ions. Figure 3 compares the CD spectra of the urea-induced unfolded state of Tel22 at 20 mM NaCl and 8 M urea and the heat-induced unfolded state at 20 mM NaCl and 95 °C in the absence of urea. Inspection of Figure 3 reveals that the urea-induced unfolded state is more structured compared to the heat-induced state at 95 °C. The latter does not display the characteristic peaks at 258 and 278 nm. Significantly, an increase in temperature causes a gradual (non-cooperative) conversion of the CD spectrum of the urea-induced unfolded state of Tel22 into that of the heat-induced unfolded state. The inset in Figure 3 presents the temperature dependence of the molar ellipticity at 257 nm. Inspection of the inset reveals that the urea- and heat-induced unfolded states are not separated by a cooperative transition.

Given these results, we conclude that the denaturing effect of urea on Tel22 is milder compared to that of temperature. At room temperature, the urea-induced unfolded state of Tel22 displays a significant amount of residual base stacking. This result is in agreement with the data from the Record lab who, based on a comparative analysis of the experimental and predicted m-values for duplex DNA and RNA, have concluded that the separated single strands are 60 to 90 % stacked. A qualitatively similar inference has been drawn from comparing the UV absorbance spectra of single stranded DNA and the respective mononucleotides. In agreement with our data, it has been suggested that the degree of single strand unstacking and the related exposure of previously buried solvent-accessible surface area is essentially independent of urea concentration. However, the stacking decreases as the temperature increases. Incomplete denaturation of DNA in the presence of urea contrasts the denaturing action of urea on proteins. As a general
rule, urea-induced unfolded states of proteins are more fully disordered compared to the heat-induced denatured states which may retain significant amounts of secondary structural elements

**Urea-salt phase diagram.** Figure 8 presents the three-dimensional plot of the combined effect of NaCl and urea on the transition temperature, $T_M$. The $T_M$-surface shown in Figure A-8 separates the folded and unfolded conformations of Tel22. Inspection of Figure 8 along with Figure A-5 reveals two interesting features. Firstly, at all NaCl concentrations studied in this work, an increase in the concentration of urea brings about thermodynamic destabilization of the G-quadruplex as reflected in a nearly linear decrease in $T_M$ (see Figure A-5). This observation is consistent with the reports of nonspecific detrimental effect of urea on the thermal stability of nucleic acid structures including an RNA pseudoknot and hairpin and polymeric DNA and RNA duplexes. Secondly, at each experimental urea concentration, an increase in the concentration of sodium ions leads to an increase in $T_M$ fully or partially offsetting the destabilizing influence of urea (see Figure A-5). While a similar behavior has been reported for RNA structures, the molecular nature of the stabilizing action of cations in a G-quadruplex and other (canonical) nucleic acid structures is very different. In G-quadruplexes, the stabilizing action of cations originates predominantly from their specific binding inside the central cavity, while the nonspecific external binding (counterion condensation) may be very slightly stabilizing, insignificant, or even strongly destabilizing. In contrast to G-quadruplexes, canonical nucleic acid structures, e.g. duplexes, are stabilized by counterion condensation.
Figure A-8. The three-dimensional dependence of the transition temperature, $T_M$, on the concentrations of NaCl and urea.

As mentioned above, transition enthalpy data shown in Figure A-6 exhibit a tendency to increase with temperature. The apparent slope, $\Delta \Delta H_M/\Delta T_M$, of $0.14 \pm 0.04 \text{ kcal mol}^{-1}\text{K}^{-1}$ may represent a change in heat capacity, $\Delta C_P$, associated with the heat-induced denaturation of Tel22. An alternative explanation may reflect the fact that $\Delta C_P$ is markedly influenced by a change in DNA solvation following the unfolding transition $^{74-78}$. An increase in urea leads to increasing replacement of water by urea within the hydration shell of DNA with the resulting alteration of solute-solvent interactions. An alteration of the nature of solute-solvent interactions in the vicinity of the folded and unfolded states of Tel22 with an increase in urea concentration may result in the observed temperature dependence of the transition enthalpy, $\Delta H_M$, with an apparently positive $\Delta C_P$. 
Figure A-6. Compilation of transition enthalpies, $\Delta H_M$, of Tel22 measured at various combinations of NaCl and urea concentrations plotted against corresponding transition temperatures, $T_M$.

**Interactions of urea with the folded and unfolded states of Tel22.** The observed deleterious effect of urea on the stability of Tel22 is in line with the previous reports of urea destabilizing the secondary and tertiary structures of a wide range of DNA and RNA constructs as manifested, for example, in a decrease in their melting temperatures, $T_M$ \cite{37, 43-45, 47-49}. The destabilizing effect of urea on nucleic acid structures is due to its favorable interactions with the rings and functional groups of nucleic acid bases which become exposed to solvent in the unfolded state \cite{37}. Significantly, urea interacts favorably with all components of nucleic acids with a local/bulk microscopic partition coefficient, $K_P$, ranging between 1.09 to 1.37 \cite{37}. If normalized per solvent-accessible surface area, urea interacts most favorably with the heterocyclic aromatic rings of all nucleic acid bases and with the methyl group of thymine \cite{37}. While this assessment is based on studies of small analogs of nucleic acids and has been applied to DNA and RNA duplex-to-
single strand transitions \(^37\), it is plausible to expect a similar molecular basis for the effect of urea on G-quadruplex stability.

A thermodynamically rigorous way of quantifying accumulation or exclusion of a cosolvent around a macromolecule in question is given by preferential interaction parameter \(\Gamma_{23} = (\delta m_3/\delta m_2)_{T,P,\mu_3}\), where \(m_2\) and \(m_3\) are the molalities of the cosolvent and the macromolecule, respectively; while \(\mu_3\) is the chemical potential of the cosolvent \(^40,79\). The difference in preferential interaction parameter between the folded and unfolded macromolecular (protein or DNA) states, \(\Delta\Gamma_{23}\), is related to cosolvent-induced modulation of the folded/unfolded equilibrium constant, \(K\), according to \(\Delta\Gamma_{23} = (\delta \ln K/\delta \ln a_3)_{T,P,m_2}\), where \(a_3\) is the activity of the cosolvent \(^79-82\). The differential preferential interaction parameter, \(\Delta\Gamma_{23}\), is related to the transition-induced changes in excess numbers of solvent, \(\Delta n_1\), and cosolvent, \(\Delta n_3\), around the solute as follows:

\[
\Delta\Gamma_{23} = \Delta n_3 - (N_3/N_1)\Delta n_1
\]  

(3)

where \(N_1\) and \(N_3\) are the numbers of mole of the principal solvent (water) and cosolvent in solution \(^40,80,83-86\).

Urea is very slightly excluded from the surface of double-stranded DNA but significantly accumulated at the surface exposed in DNA and RNA unfolding, in particular, the amide-like surfaces of guanine, cytosine, and, especially, thymine and uracil bases \(^47\). There are no data on the preferential interactions of urea with G-quadruplexes. Nonetheless, it might be expected that, by analogy, urea is also very slightly excluded (or very slightly accumulated) from the surface of the G-quadruplex conformation while being strongly accumulated around the unfolded state.

The m-value and differential interaction parameter, \(\Delta\Gamma_{23}\), are related via the following relationship \(^36\):

\[
\]
\[ m \approx RT \Delta \Gamma_{23}/[\text{urea}] \quad (4) \]

The urea m-value of a protein is proportional to a change in solvent-accessible surface area, \( \Delta S_A \), accompanying the unfolding transition \(^{87}\). Comparison of Eqs. (3) and (4) reveals that the ratio \( m/\Delta S_A \) correlates with and reflects the net cosolvent excess or depletion around 1 \( \text{Å}^2 \) of the newly exposed solute surface. Hence, \( m/\Delta S_A \) reflects the complex of differential solute-water and solute-cosolvent interactions that determine the specific effect of a given cosolvent on the stability of the macromolecule in question. It is instructive to compare the ratio \( m/\Delta S_A \) for the unfolding of the Tel22 G-quadruplex with that determined for other DNA and RNA denaturation events. A change in solvent accessible surface area, \( \Delta S_A \), accompanying the unfolding of the \( \text{Na}^+ \)-stabilized Tel22 G-quadruplex has been estimated to be 1230 \( \text{Å}^2 \)\(^{25}\). With this estimate and our determined \( m \)-value of 0.57\( \pm \)0.09 kcal mol\(^{-1}\)M\(^{-1}\), we calculate the ratio \( m/\Delta S_A \) of 0.46 cal mol\(^{-1}\)M\(^{-1}\)Å\(^{-2}\).

Survey of the literature reveals a large variation of the reported values of \( m/\Delta S_A \) for DNA and RNA unfolding reactions. Shelton et al.\(^{44}\) have estimated an average value of urea \( m/\Delta S_A \) of 0.099 cal mol\(^{-1}\)M\(^{-1}\)Å\(^{-2}\) for five self-complementary RNA duplexes of varying lengths and tRNA\(^{\text{Phe}}\). While this estimate is close to that observed for protein unfolding \(^{87}\), it is much smaller than 0.46 cal mol\(^{-1}\)M\(^{-1}\)Å\(^{-2}\), our estimate for the unfolding of the Tel22 G-quadruplex. On the other hand, Lambert and Draper\(^{45}\) have determined, for a number of RNA structures including a hairpin, the tar-tar\(^*\) RNA kissing loop complex, the tetraloop-receptor (TLR) structural motif, and the aptamer domain of the adenine riboswitch (A-riboswich), higher values of urea \( m/\Delta S_A \) ranging between 0.48 (A-riboswich) and 0.65 (hairpin) cal mol\(^{-1}\)M\(^{-1}\)Å\(^{-2}\) (close to our estimate for Tel22). Using the urea \( m \)-values determined for 12-meric DNA and RNA duplexes reported by Guinn et al.\(^{37}\) and their calculated changes in solvent-accessible surface areas assuming 50% stacking of the single stranded conformation, we evaluate the values of \( m/\Delta S_A \) ranging between 0.24 and 0.40 cal mol\(^{-1}\)M\(^{-1}\)Å\(^{-2}\) (again, close to our estimate). Hence, the value of \( m/\Delta S_A \) we determined for the Tel22 G-quadruplex is within the range of similar values determined by Lambert and Draper\(^{45}\) and Guinn et al.\(^{37}\), while being larger than those reported by Shelton et al.\(^{44}\). As noted by Lambert and Draper\(^{45}\), the smallness of the value of \( m/\Delta S_A \) reported by Shelton...
et al.\textsuperscript{44} may be due to an unrealistically large estimate of the change in solvent-accessible surface area, $\Delta S_A$, used by these authors in their analysis (the unfolded state has been modeled by the fully unstacked conformation). The use of more realistic (smaller) values of $\Delta S_A$ would bring $m/\Delta S_A$ to higher values (similar to our estimate for Tel22). Given the similarity of urea $m/\Delta S_A$ ratios determined for the Tel22 G-quadruplex and other DNA and RNA structures\textsuperscript{37,45}, we propose that the molecular basis of urea-induced destabilization of Tel22 is similar to that of other nucleic acid species.

### A.6 Conclusions

We presented the first systematic investigation of the combined effect of urea, salt, and temperature on the stability of a G-quadruplex. We derived the urea-salt-temperature phase-diagram of a telomeric DNA G-quadruplex which, in the presence of Na$^+$ ions, adopts the antiparallel conformation. At 20 mM NaCl and room temperature, urea fully denatures the Tel22 G-quadruplex. An increase in salt mitigates the deleterious effect of urea on Tel22. The determined urea $m$-value of Tel22 normalized per change in solvent-accessible surface area, $\Delta S_A$, is similar to those for other DNA and RNA structures while being several-fold larger than that of proteins. Taken together, our results suggest that urea can be used as an analytical tool in thermodynamic characterizations of G-quadruplexes in a manner similar to the use of urea in protein studies. Studies involving a larger selection of G-quadruplexes varying in sequence, topology (parallel, antiparallel, hybrid), and molecularity (monomolecular, bimolecular, tetramolecular) must be conducted to outline the advantages and the limits of the use of urea in G-quadruplex studies. In particular, a deeper understanding of the effect of solvent and cosolvents on the differential stability of the G-quadruplex and duplex conformations is a step towards elucidation of the modulating influence of different types of cosolvents on duplex-G-quadruplex molecular switches triggering genomic events.
A.7 Acknowledgements

This work was supported by grant 203816-2012 from the Natural Sciences and Engineering Research Council of Canada (to T. V. C.) and grant 15T-1F048 from the State Committee of Sciences of the Republic of Armenia (to Y. B. D.).
A.8 Reference


63. Kim, B. G.; Shek, Y. L.; Chalikian, T. V. Polyelectrolyte effects in G-quadruplexes. 

64. Kim, B. G.; Evans, H. M.; Dubins, D. N.; Chalikian, T. V. Effects of salt on the stability 

65. Kim, B. G.; Long, J.; Dubins, D. N.; Chalikian, T. V. Ionic effects on VEGF G-

66. Pace, C. N.; Shaw, K. L. Linear extrapolation method of analyzing solvent denaturation 

67. Makhatadze, G. I. Thermodynamics of protein interactions with urea and guanidinium 

68. Greene, R. F., Jr.; Pace, C. N. Urea and guanidine hydrochloride denaturation of 
   5388-5393.

69. Holbrook, J. A.; Capp, M. W.; Saecker, R. M.; Record, M. T., Jr. Enthalpy and heat 
   capacity changes for formation of an oligomeric DNA duplex: interpretation in terms of 
   coupled processes of formation and association of single-stranded helices. *Biochemistry* 
   **1999**, *38*, 8409-8422.

70. Manning, G. S. Molecular theory of polyelectrolyte solutions with applications to 

71. Record, M. T.; Anderson, C. F.; Lohman, T. M. Thermodynamic analysis of ion effects 
   on binding and conformational equilibria of proteins and nucleic acids. Roles of ion 
   association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* **1978**, 
   *11*, 103-178.

72. Anderson, C. F.; Record, M. T. Polyelectrolyte theories and their applications to DNA. 


82. Courtenay, E. S.; Capp, M. W.; Saecker, R. M.; Record, M. T., Jr. Thermodynamic analysis of interactions between denaturants and protein surface exposed on unfolding:


