BIOPHYSICAL STUDIES ON HUMAN TELOMERIC G-QUADRUPLEX DNA: CHARACTERIZATION OF LIGAND AND PROTEIN INTERACTIONS

by

JASON SCOTT HUDSON

DAVID E. GRAVES, CHAIR
HERBERT C. CHEUNG
AARON L. LUCIUS
DONALD D. MUCCIO
PETER E. PREVELIGE

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

BIRMINGHAM, ALABAMA

2010
BIOPHYSICAL STUDIES ON HUMAN TELOMERIC G-QUADRUPLEX DNA: CHARACTERIZATION OF LIGAND AND PROTEIN INTERACTIONS

JASON SCOTT HUDSON

CHEMISTRY

ABSTRACT

G-quadruplex DNAs are higher order nucleic acid structures that have been implicated in the regulation of a number of vital biological processes, including gene transcription and telomeric extension. Guanine rich sequences that are capable of forming G-quadruplex structure are found in the telomeres of chromosomes. The presence of the G-quadruplex structure prevents the binding of telomerase and inhibits the telomeric extension that would result in the immortality of cancer cells. This dissertation describes research that is focused on the biophysical characterization of the G-quadruplex and interactions with a ligand and protein.

The anticancer compound, actinomycin D, is demonstrated to bind to both the Na\(^{+}\) and K\(^{+}\) conformations of telomeric G-quadruplex DNA by an end stacking mechanism with favorable binding enthalpies. Complex formation with actinomycin D resulted in changes to both the Na\(^{+}\) and K\(^{+}\) conformations to ligand-bound complexes that had similar structural features and stabilities. End stacking of actinomycin D with G-quadruplex results in a 2:1 G-quadruplex DNA to drug complex that maximizes the stacking interaction of the planar phenoxazine ring on the terminal G-tetrad.

In an effort to understand the fundamental properties of G-quadruplex structure and stability, the influence of the adenine bases in the three loops on the energetics and stability of the G-quadruplex were investigated. Adenine and guanine stacking
interactions provide important contributions to the overall stability and energetics of the structures. The systematic mutation of adenine to thymine in the three loop sequences results in significant changes in the thermal stability and folding enthalpies.

The nucleoprotein, UP1, is demonstrated to recognize and bind with high affinity to G-quadruplex DNA. Upon initial formation of the UP1-G-quadruplex complex the G-quadruplex structure unfolds, allowing the binding of a second UP1 to the exposed TAG site and resulting in a final unfolded G-quadruplex sequence bound with two UP1 molecules. UP1 binding studies with loop mutations revealed that UP1 recognition and initial binding occurs at loop 2. G-quadruplexes with loop 2 mutations revealed a reduction in binding affinity and unfolding by UP1. These results demonstrated that UP1 specifically recognizes G-quadruplex DNA structural motif and the binding energetics are coupled to the unfolding of the G-quadruplex structure.

Keywords: G-quadruplex DNA, Isothermal Titration Calorimetry (ITC), Differential Scanning Calorimetry (DSC), Circular Dichroism (CD), Unfolding Protein 1 (UP1)
DEDICATION

This work is dedicated to my parents, Danny and Cathy Hudson. My achievements are a result of their encouragement and unwavering support.
ACKNOWLEDGMENTS

I owe my deepest gratitude to my research advisor and friend, Dr. David Graves, for the opportunity he has provided me. From our days at Ole Miss together to the present, his guidance has been invaluable. I must thank him for allowing me to work in his laboratory while providing me with the resources and intellectual freedom to pursue research that was of interest to me, even if it resulted in a trip down the rabbit hole. I also have to thank him for adding levity to the stressful times with his humor and countless number of one-liners that I will undoubtedly use on some unsuspecting individual somewhere along the way. I am appreciative of the numerous letters of recommendations he has written for me, having the freedom to walk into his office at any time, the conferences I have been able to attend, the conversations that had nothing to do with science, reigning me in when I got lost, the research meetings over lunch or at the Barking Kudu, and teaching me to apply common sense to complex problems. Simply put, thanks Boss.

I am grateful to the members of the research committee Dr. Muccio, Dr. Lucius, Dr. Prevelige, and Dr. Cheung, for taking the time to serve on my committee and for reviewing the research that comprises this dissertation.

I value the casual conversations and friendship provided to me by the current graduate students and from those who have preceded me in graduation. A special thanks to Dr. Samuel Owens, Dr. Dustin Myrex, and soon-to-be Dr. Jason Freeman for being good guys to have with me in the trenches.

A majority of this work could not have been completed without our collaboration with Lei Ding of Dr. Ma’s laboratory. Working with him has been a pleasure. In addition, Dr. Edwin Lewis, at Mississippi State was instrumental in the data fitting procedures used for a large portion of the research. I could not have done it without him.

Lastly, I have to thank my mother, father, sister, and the rest of the family for their confidence in me. I hope they consider this work as their accomplishment as well, because without them I don’t think it would have been possible. I must thank my wife, Ashley, for her support and understanding. Her patience during the long nights in the lab and assurance during trying times has not gone unnoticed.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>14</td>
</tr>
<tr>
<td>Deoxyribonucleic Acids (DNA)</td>
<td>15</td>
</tr>
<tr>
<td>Helical DNA Structure and Stability</td>
<td>16</td>
</tr>
<tr>
<td>Telomeric DNA</td>
<td>17</td>
</tr>
<tr>
<td>The G-quadruplex Structural Motif</td>
<td>19</td>
</tr>
<tr>
<td>DNA Binding Ligands</td>
<td>22</td>
</tr>
<tr>
<td>Actinomycin D</td>
<td>25</td>
</tr>
<tr>
<td>Protein Interactions with Nucleic Acids</td>
<td>28</td>
</tr>
<tr>
<td>Ribonucleoprotein A1</td>
<td>29</td>
</tr>
<tr>
<td>Rationale for Research Project and Dissertation Overview</td>
<td>31</td>
</tr>
<tr>
<td>Outline of the Dissertation</td>
<td>32</td>
</tr>
<tr>
<td>II. MATERIALS AND METHODS</td>
<td>34</td>
</tr>
<tr>
<td>Buffer Preparation</td>
<td>34</td>
</tr>
<tr>
<td>Oligonucleotide Preparation</td>
<td>35</td>
</tr>
<tr>
<td>Characterization of G-quadruplex Folding Topology</td>
<td>36</td>
</tr>
<tr>
<td>Isothermal Titration Calorimetry</td>
<td>38</td>
</tr>
<tr>
<td>General Thermodynamic Principles</td>
<td>38</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>39</td>
</tr>
<tr>
<td>Experimental Considerations</td>
<td>40</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>43</td>
</tr>
<tr>
<td>Differential Scanning Calorimetry</td>
<td>44</td>
</tr>
<tr>
<td>General Thermodynamic Principles and Instrumentation</td>
<td>44</td>
</tr>
<tr>
<td>-----------------------------------------------------</td>
<td>----</td>
</tr>
<tr>
<td>Experimental Considerations</td>
<td>46</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>47</td>
</tr>
</tbody>
</table>

| III. THE INTERACTION OF ACTINOMYCIN D WITH HUMAN TELOMERIC G-QUADRUPLEX DNA | 49 |

| IV. BIOPHYSICAL CHARACTERIZATION OF THE BINDING AND UNFOLDING OF HUMAN TELOMERIC G-QUADRUPLEX DNA BY UNFOLDING PROTEIN 1 (UP1) | 76 |

| V. INFLUENCE OF LOOP SEQUENCE MUTATIONS ON THE STRUCTURE AND STABILITY OF HUMAN TELOMERIC G-QUADRUPLEX DNA | 109 |

| VI. THE INFLUENCE OF LOOP SEQUENCE MUTATIONS ON THE BINDING AND UNFOLDING OF HUMAN TELOMERIC G-QUADRUPLEX DNA BY UNFOLDING PROTEIN 1 (UP1) | 140 |

| VII. CONCLUSIONS | 176 |

| LIST OF REFERENCES | 181 |
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>THE INTERACTION OF ACTINOMYCIN D WITH HUMAN TELOMERIC G-QUADRUPLEX DNA</td>
<td></td>
</tr>
<tr>
<td>1 Thermodynamic properties associated with the interactions of actinomycin D with the Na(^+) and K(^+) structural isoforms of G-quadruplex DNA</td>
<td>70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOPHYSICAL CHARACTERIZATION OF THE BINDING AND UNFOLDING OF HUMAN TELOMERIC G-QUADRUPLEX DNA BY UNFOLDING PROTEIN 1 (UP1)</td>
<td></td>
</tr>
<tr>
<td>1 Thermodynamic parameters derived from the non-linear least squares fit of the isothermal titration calorimetry binding studies in the forward direction</td>
<td>101</td>
</tr>
<tr>
<td>2 Thermodynamic parameters derived from the best fit line of the isothermal titration binding studies in the reverse direction</td>
<td>102</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFLUENCE OF LOOP SEQUENCE MUTATIONS ON THE STRUCTURE AND STABILITY OF HUMAN TELOMERIC G-QUADRUPLEX DNA</td>
<td></td>
</tr>
<tr>
<td>1 Nomenclature for the mutated G-quadruplex sequences</td>
<td>132</td>
</tr>
<tr>
<td>2 Summary of DSC and CD thermal denaturation experiments</td>
<td>133</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFLUENCE OF LOOP SEQUENCE MUTATIONS ON THE BINDING AND UNFOLDING OF HUMAN TELOMERIC G-QUADRUPLEX DNA BY UNFOLDING PROTEIN 1 (UP1)</td>
<td></td>
</tr>
<tr>
<td>1 Telomeric G-quadruplex loop sequence mutations</td>
<td>162</td>
</tr>
<tr>
<td>2 Thermodynamic parameters derived from the forward titration ITC binding studies of the G-quadruplex sequences with single loop mutations</td>
<td>163</td>
</tr>
<tr>
<td>3 Thermodynamic parameters derived from the forward titration ITC binding studies of G-quadruplex sequences with multiple loop mutations</td>
<td>164</td>
</tr>
</tbody>
</table>
Thermodynamic parameters derived from reverse titration ITC binding studies......165

Thermodynamic parameters derived from the ITC binding studies of the UP1 binding domains with the wild type G-quadruplex DNA... .................................................................166
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figures</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Complementary base pairing as proposed by Watson and Crick. 15</td>
</tr>
<tr>
<td>2</td>
<td>The end replication problem. 18</td>
</tr>
<tr>
<td>3</td>
<td>The planar G-tetrad formed by reverse Hoogsteen hydrogen bonds. 20</td>
</tr>
<tr>
<td>4</td>
<td>Strand orientations of G-quadruplex structure. 21</td>
</tr>
<tr>
<td>5</td>
<td>G-quadruplex DNA structures formed by the human telomeric sequence. 21</td>
</tr>
<tr>
<td>6</td>
<td>Examples of small molecules that interact with nucleic acids and their primary mode of interaction. 23</td>
</tr>
<tr>
<td>7</td>
<td>Spatial representation of actinomycin D rotated by 90 degrees. 26</td>
</tr>
<tr>
<td>8</td>
<td>Molecular modeling representation of UP1 rotated by 90 degrees. 30</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Base conformations of the guanines in the G-tetrads formed by the Na⁺ and K⁺ G-quadruplex DNA structures. 37</td>
</tr>
<tr>
<td>10</td>
<td>Circular dichroism spectra of the unimolecular G-quadruplexes formed by the human telomeric sequence. 38</td>
</tr>
<tr>
<td>11</td>
<td>Simulated binding isotherms with various c values. 41</td>
</tr>
<tr>
<td>12</td>
<td>Raw ITC binding isotherm for 40 injections of a single stranded oligonucleotide into its complementary strand. 43</td>
</tr>
<tr>
<td>13</td>
<td>The integration of the raw ITC data (black squares) for the titration of complementary oligonucleotides. 44</td>
</tr>
<tr>
<td>14</td>
<td>The determination of the thermodynamic parameters of a typical DSC thermogram. 46</td>
</tr>
</tbody>
</table>
THE INTERACTION OF ACTINOMYCIN D WITH HUMAN TELOMERIC G-QUADRUPLEX DNA

1 Representative CD spectra of the Na$^+$ (black) and K$^+$ (red) of structural isoforms of quadruplex DNA.................................................................71

2 Changes to the CD spectra of the Na$^+$ form (panel A) and the K$^+$ form (panel B) of quadruplex DNAs are observed upon binding of actinomycin D.........................72

3 The induced CD spectra of actinomycin D complexed with the Na$^+$ (red) and K$^+$ (blue) forms of quadruplex DNAs .................................................................73

4 Representative data from ITC experiments showing the titration of actinomycin D with the Na$^+$ (panel A) and the K$^+$ (panel B) forms of quadruplex DNAs. ITC experiments were performed at 25°C ...............................................................74

5 DSC thermograms of the K$^+$ form of quadruplex DNA (panel A, no actinomycin D and panel B, with actinomycin D) and the Na$^+$ form of quadruplex DNA ............75

BIOPHYSICAL CHARACTERIZATION OF THE BINDING AND UNFOLDING OF HUMAN TELOMERIC G-QUADRUPLEX DNA BY UNFOLDING PROTEIN 1 (UP1)

1 Antiparallel structures of the G-quadruplex of the 12 nucleotide sequence of the human telomere..................................................................................................................103

2 Circular dichroism spectroscopy (CD) melts of the Tel-12 (Na$^+$) (squares), Tel-22 (Na$^+$) (circles), and Tel-22 (K$^+$) (triangles).........................................................104

3 Isothermal titration calorimetry (ITC) binding isotherms for the titration of Tel-22 (Na$^+$) and Tel-12 (Na$^+$), panels A and B (top). ..............................................................105

4 Isothermal titration calorimetry (ITC) binding isotherm for the forward titration of Tel-22 (K$^+$)-quadruplex DNA (panel A)..................................................................106

5 The isothermal titration binding isotherm (ITC) for the titration of Tel-22(SS) with UP1 (panel A). ..................................................................................................................107

6 Isothermal titration calorimetry (ITC) binding isotherms for the titration of UP1 with Tel-22 (Na$^+$) and Tel-12 (Na$^+$), panels A and B. ................................................108
INFLUENCE OF LOOP SEQUENCE MUTATIONS ON THE STRUCTURE AND STABILITY OF HUMAN TELOMERIC G-QUADRUPLEX DNA

1 Cartoon representing the chair (left) and basket (right) conformations of the Na\(^+\) human telomeric G-Quadruplex. .................................................................134

2 CD wavelength scans for the human telomeric G-quadruplex d[AGGG(TTAGGG)\_3] (red line) and the loop mutated sequences. .................................................................135

3 Composite CD melting curves of the human telomeric sequence and loop mutated structures in 100 mM NaCl BPES buffer at pH 7.0. .................................................................136

4 A DSC thermogram of the wild type human telomeric G-quadruplex DNA in 100mM NaCl BPES buffer at pH 7.0. .................................................................137

5 Representative DSC thermograms for the single base mutations of the G-quadruplex DNA. ........................................................................................................138

6 Representative DSC thermograms for the multiple loop mutations of the G-quadruplex. ........................................................................................................139

INFLUENCE OF LOOP SEQUENCE MUTATIONS ON THE BINDING AND UNFOLDING OF HUMAN TELOMERIC G-QUADRUPLEX DNA BY UNFOLDING PROTEIN 1 (UP1)

1 The sequence and structural representation of telomeric G-quadruplex DNA .......167

2 Circular dichroism studies of the unfolding of wild type G-quadruplex DNA and sequences with single loop mutants and multiple loop mutants. .........................168

3 ITC binding isotherms for the forward titrations of the single loop mutations of the telomeric G-quadruplex sequence .................................................................169

4 ITC binding isotherms for the forward titrations of the multiple loop mutations of the telomeric G-quadruplex sequence .................................................................170

5 ITC binding isotherms for the reverse titrations of the single loop mutations of the telomeric G-quadruplex sequence .................................................................171

6 ITC binding isotherms for the reverse titrations of the multiple loop mutations of the telomeric G-quadruplex sequence .................................................................172
Circular dichroism studies of the unfolding of wild type G-quadruplex DNA after the interaction with the individual nucleic acid binding domains of UP1......................173

ITC binding isotherms for the forward titrations of the wild type telomeric G-quadruplex with binding domain 1 (panel A) and binding domain 2 (panel B)........174

Cartoon representing a three event model for the binding and unfolding of telomeric G-quadruplex DNA by UP1.........................................................175
CHAPTER I
INTRODUCTION

The motivation to understand the molecular basis for heredity led scientists in the late 19th and early 20th centuries to embark on studies that would ultimately reveal the molecule that was the repository for all genetic information, DNA. Based on X-ray diffractions studies conducted by Maurice Wilkins and Rosalind Franklin, James Watson and Francis Crick developed a model for DNA that consisted of two strands of polymeric nucleotides linked by a sugar-phosphate backbone that were hydrogen bonded together by complementary nucleic acid bases and formed a helical structure termed the double helix (1). The discovery of DNA, the subsequent elucidation of the structure, and revelation of the genetic code linked chemistry and biology and led to the development of new scientific disciplines such as molecular biology and molecular genetics.

Since the discovery of DNA, there have been a significant number of breakthroughs and discoveries which have greatly influenced modern medicine. However, questions still remain regarding many biological processes that involve nucleic acids and their correlation to a variety of disease states. Interactions of small molecules and biological macromolecules with nucleic acids remains of high interest and has led to the scientific pursuits and investigations enclosed in the following dissertation.
Deoxyribonucleic Acids (DNA)

Deoxyribonucleic acids are composed of monomeric nucleotides that are linked together as a biopolymer by phosphate functional groups coupled to the 5'- and 3'-hydroxyl of a deoxyribose sugar. Individual nucleotides are classified based on their ring systems being either a purine or a pyrimidine and further categorized based on their substituent functional groups. The purine nucleotides are either adenine (A) or guanine (G), and the pyrimidines are either cytosine (C) or thymine (T). Based on Erwin Chargraff’s finding that the distribution of A:T and G:C nucleotides in DNA was 1:1 (2), Watson and Crick developed their model of duplex DNA so that A was hydrogen bonded with T and G was hydrogen bonded to C. The hydrogen bonding between the bases was responsible for holding the individual strands of DNA together. This classical base pairing is denoted as Watson-Crick base pairing and is depicted in Figure 1.

Figure 1. Complementary base pairing as proposed by Watson and Crick. The wavy lines represent the attachment to the ribose sugar of the phosphate backbone.
The X-ray diffraction data collected by Maurice Wilkins and Rosalind Franklin revealed critical aspects related to the symmetry and helical nature of DNA (3, 4). These features allowed Watson and Crick to develop a structural model of DNA deemed the double helix. The characterization of the DNA double helix led to more probing questions regarding the biological implications for nucleic acids.

**Helical DNA Structure and Stability**

The primary structure of DNA is formed when each nucleoside is linked via a phosphodiester bond from the 5'-hydroxyl group of the ribose sugar to the 3'-hydroxyl group of a neighboring nucleotide. Hydrogen bonding of the purine and pyrimidine bases of one single stranded DNA chain to a complementary sequence of another DNA chain in an anti-parallel arrangement results in the formation of the double helix. Double helical DNA structures are characterized by features such as rise, helix twist, tilt angle of the bases, and groove depth and width. Variations in these structural features give rise to unique helical conformations and demonstrate the dynamic nature of the DNA structure. For example, A-DNA and B-DNA are both composed of right handed helices with variations in groove widths and depths.

There are primarily four types of forces that contribute to the stability of the DNA double helix: (1) hydrophobic interactions (2) base stacking interactions (3) hydrogen bonding and (4) electrostatic interactions. The complementary base pairs forming the interior of the DNA create a hydrophobic environment. The negatively charged phosphates groups along the sugar-phosphate backbone are exposed to the solvent. The stacking interactions and hydrogen bonding characteristics between the AT and GC base
pairs are a significant stabilizing force that effectively holds the two complementary strands of DNA together. Base stacking interactions and hydrogen bonding interactions are relatively weak but are additive and significant contributors to the overall stability of the duplex DNA. Therefore, the base composition and length of the DNA chain has a considerable effect on the stability of helical DNA. Electrostatic forces contribute to the stability of the double helix through interactions of the negatively charged phosphates oxygen atoms of the sugar-phosphate backbone with positively charged counter-ions such as Na\(^+\), K\(^+\), and Mg\(^{2+}\) ions.

**Telomeric DNA**

Chromosomes consist of a long molecule of double stranded DNA that is packaged by coiling DNA around a histone core. The ends of the chromosomes create a unique biological problem: they are difficult to replicate due to the end replication problem (5,6). The end replication problem originates from the semi-conservative nature of eukaryotic replication. The two strands of DNA are separated and each of the parent strands is used as a template for a new strand of DNA called the daughter strand. The synthesis of the daughter strands is conducted by DNA polymerase and results in two strands of DNA from one parent strand. The problem lies within the directionality of the synthesis of the daughter strands by DNA polymerase. DNA polymerase replicates the parent strands in the 5' to 3' direction and requires an initial RNA primer with a free 3'-OH group to begin replication. The leading strand can be synthesized continuously whereas the lagging strand requires multiple RNA primers with a free 3'-OH group to copy the DNA. When the primers on the 3' ends of the parent strands are removed a gap
is created and the DNA is progressively shortened on the 3' end with each round of replication. Figure 2 illustrates the end replication problem and the resulting 3' overhang.

Figure 2. The end replication problem.

To prevent the loss of genetic information, chromosomes have protective ends called telomeres that are composed of a non-coding tandemly repeated sequence, 5'- (TTAGGG)$_n$-3'. At the 3' end of the telomere, there is a single stranded overhang of the purine rich repeat sequence, d(TTAGGG)$_n$, which is approximately 100-150 bases in length. In germline cells and in the cells of young individuals the telomeres are of maximal length, but with every round of cell division the telomeres become progressively shorter. After many rounds of replication, the telomeric region becomes too short to continue dividing while still maintaining the genetic integrity of the chromosome. At this stage, cells enter senescence and undergo apoptosis (7, 8). Nature has a way to bypass the progressive shortening using the enzyme telomerase. Telomerase has an RNA template complementary to the 3' overhang sequence and the gap can be filled in by
reverse transcription. Telomerase is active in the rapidly dividing cells of newborns and young children but is inactive in adults. Up-regulation of telomerase activity in adults is only observed in tumor cells and is considered to be a contributing factor to the observed immortality of cancer cells. The extension of the telomeres by telomerase in cancer cells allows for indefinite replication, leading to tumor growth and disease progression (9-12).

Higher order DNA structures that are formed by the (TTAGGG)_n repeat sequence have been discovered at the telomeric ends of the chromosome and are called G-quadruplexes (13). The presence of G-quadruplexes at the telomeric ends of the chromosome and their equilibrium with single stranded DNA suggests that these higher order DNA structures serve a regulatory function for biological processes such as telomerase mediated extension of the telomere (14-16). The emergence of the role of telomeres and telomerase in cancer progression has made these structures an attractive target for novel drug development that can selectively target and stabilize the G-quadruplex structures; thus inhibiting the extension of the telomeric sequence and suspending cell immortality (17-19).

The G-Quadruplex Structural Motif

In 1962, Davies et al., reported that guanosine 5’-monophosphate can form G-tetrads that are linked by Hoogsteen hydrogen bonds and stack into planar arrays of successive tetrads (20,21). Figure 3 shows the planar tetrad composed of four guanines. The circle in the center of the tetrad is a coordinated metal cation such as K^+, Na^+, and Mg^{2+}. 
Figure 3. The planar G-tetrad formed by reverse Hoogsteen hydrogen bonds.

When multiple guanine nucleotides are linked by the phosphate backbone, such as the telomeric sequence (TTAGGG)\textsubscript{n}, G-quadruplex structures are readily formed under biological conditions. G-quadruplex structural motifs can be characterized by their strand orientations (parallel or antiparallel) and are sensitive to changes in loop sequences, terminal bases, and selection of monovalent cation. As shown in Figure 4, a variety of structures can be drawn that represents a tetramolecular, biomolecular, and unimolecular G-quadruplexes with unique topologies and strand polarities. The tetramolecular structure is a parallel strand orientation, whereas the bi- and unimolecular quadruplexes possess antiparallel strand polarities and are in equilibrium with multiple folding topologies that are influence by cations and sequence effects.
Figure 4. Strand orientations of G-quadruplex structure. (A) parallel tetramolecular G-quadruplex DNA. (B) unimolecular antiparallel G-quadruplex DNA. (C) bimolecular antiparallel G-quadruplex DNA.

The human telomeric repeat is a purine rich 22 nucleotide sequence of 5’-AGGGTTAGGGTTAGGGTTAGGG-3’ that readily forms stable G-quadruplex structures. The structure is composed of three stacked tetrads that are connected by three TTA loop regions. Monovalent cations such as Na⁺ and K⁺ coordinate with the G-tetrads and significantly influences the stability and the folding topology (22, 23). Rendering of the molecular model of the G-quadruplex structural isoforms are represented in Figure 5.

Figure 5. G-quadruplex DNA structures formed by the human telomeric sequence. (A) Na⁺ basket structure; PDB:143D. (B) K⁺ hybrid structure; PDB: 2JPZ. (C) K⁺ propeller structure; PDB:1KF1.
The Na\textsuperscript{+} form of the G-quadruplex is comprised of antiparallel strand polarities, and based on NMR data, exists as a “basket” structure with two lateral and one diagonal TTA loops (24). The K\textsuperscript{+} form of the telomeric G-quadruplex has more complex topological features. Patel and co-workers utilized NMR spectroscopy to report that the K\textsuperscript{+} G-quadruplex formed a structure characterized by a mixed parallel/antiparallel strand polarity with a three dimensional structure similar to the Na\textsuperscript{+} structural isoform (25-27). In contrast, Neidle and co-workers reported the x-ray crystal structure of the K\textsuperscript{+} G-quadruplex and reported a structure that folded into an all parallel arrangement with external loops deemed the “propeller” (28).

**DNA Binding Ligands**

Nucleic acids interact with a large number of molecules such as water, metal ions, heterocyclic small molecules, and proteins. These interactions make available a multitude of roles that serve functional and regulatory purposes. Nucleic acids provide attractive targets for drug development primarily due to their direct influence over gene expression and protein synthesis. Broad classes of therapeutic compounds have been developed for the chemotherapeutic treatment of cancer and viral diseases. Both synthetic and naturally occurring DNA binding ligands have been extensively studied and provide us with most of our current understanding of reversible binding interactions with nucleic acids. Small molecules that interact with nucleic acids are typically classified by their mode of binding and include electrostatic, exterior, and base specific interactions.
Examples of small molecules and their mode of interactions may be seen below in Figure 6.

DNA is a weak acid and is negatively charged due to the ionization of the phosphate groups at physiological pH. Electrostatic interactions are generally non-specific and require a positively charged ligand. The simplest form of electrostatic interactions would be DNA binding with cations such as Na⁺ or K⁺. Electrostatic interactions are important for helical DNA structure and stability, and the presence of a salt that dissociates in aqueous solution to produce positively and negatively charged ions has a direct influence on the thermal equilibrium between single and double stranded
DNA. The concentration of bulk water in solution is 55.5 M and results in extensive hydrogen bonding interactions with nucleic acids. In combination with counter-ions, water molecules orient themselves on the exterior of nucleic acids and influence the conformation and stability of helical DNA.

Small molecules that interact with the major and minor grooves of B-DNA are denoted as groove binders. The major and minor grooves have different physical and chemical properties such as electrostatic potential, hydrogen bonding capability, and variations of the width and depth of the major and minor grooves. These properties impart a binding specificity for small molecules that are groove binding ligands. Sequence specificity of groove binding ligands is due to the hydrogen bonding between the ligand and the exterior of the bases that compose either the major or minor groove. Size and shape of the ligand also plays an important role in the specificity for either the major or the minor groove primarily due to the narrow steric constraints of the minor groove relative to the wider major groove. The curvature in the shape of a groove binding ligand is important in order to maximize the interactions between the small molecule and cylindrical nature of helical DNA.

Netropsin is perhaps one of the better known groove binding ligands and the structural characteristics are depicted in Figure 6. The crystal structure of netropsin binding to a 12-nucleotide duplex DNA revealed that the netropsin bound near AT rich regions of the minor groove (29). The structural influences of netropsin binding to the 12-nucleotide duplex DNA were limited to a widening of the groove with minimal impact on the helical and conformational characteristics of the duplex. Distamycin is a pentapeptide that possesses a curvature that is geometrically complementary to the minor
groove of helical DNA. Similar to netropsin, dystamycin makes significant non-covalent interactions with AT sequences of minor grooves (30).

Intercalators comprise a large class of DNA binding ligands that insert heterocyclic planar chromophores between adjacent base pairs (31). Compounds that intercalate are typically composed of planar aromatic ring systems that maximize the favorable stacking interactions between adjacent base pairs. The structure of the DNA double helix changes significantly upon the intercalation of a compound between adjacent base pairs. The helical DNA unwinds, bends, or kinks in order to accommodate a planar molecule between the base stacks.

*Actinomycin D*

The first DNA binding agent to be used in the treatment of cancer was actinomycin D. Actinomycin D was discovered in the 1940’s by Selman Waksman while studying the antibacterial properties of soil borne bacteria, and molds of *Streptomyces antibioticus* and *Streptomyces chrismallus* (32). Waksman’s research noted that this and related compounds had potent bactericidal properties. Initial attempts to use actinomycin D as compound for the treatment of cancer were unsuccessful due to the toxicity of the drug. However, further examination revealed that actinomycin D was highly effective against childhood cancers. Actinomycin D was demonstrated to be particularly effective in the treatment of Wilms’ tumors, which are malignant tumors of the kidney’s developing almost exclusively in children under the age of six (33, 34). Actinomycin D consists of a heterocyclic phenoxazone ring substituted with two identical cyclic pentapeptide side chains that are composed of hydrophobic
amino acid residues (Figure 7). The planar ring system and the pentapeptide side chains allow for both intercalative and groove binding interactions with duplex DNA.

![Figure 7. Structural representation of actinomycin D rotated by 90 degrees.](image)

The interactions of actinomycin D with nucleic acids have been examined for over sixty years. These studies have provided a model for the interaction of actinomycin D with duplex DNA. In duplex DNA, actinomycin D binds preferentially to the 5'-dGpC-3' sequence (35). The actinomycin D-DNA complex is formed by insertion of the planar phenoxazone ring between 5'-GpC-3' base pairs and orients the cyclic pentapeptide side-chains within the minor groove of the DNA above and below the intercalation site. With bimodal binding features of the interaction with duplex DNA, actinomycin D has both intercalative and groove binding characteristics. In 1992, Takasugawa and co-workers reported the first crystal structure of actinomycin D complexed with DNA.
containing an eight nucleotide double stranded DNA sequence, d(GAAGCTTC)\textsubscript{2}, and revealed the structural features of the actinomycin D-DNA complex (36). The intercalation of actinomycin D involved base stacking interactions and induced a bending coupled with a significant distortion of the helical DNA. Also noted were key hydrogen bond contacts between the carbonyl oxygen of the threonine residue of the pentapeptide side chains with the 2-amino group of guanine in the minor groove. Based on these results, as well as further studies utilizing both X-ray crystallography and NMR spectroscopy, the binding of the drug to duplex DNA was driven by an assortment of stabilizing non-covalent interactions including: base stacking, hydrogen bonding, and hydrophobic effects (37-40).

More recently the binding of actinomycin D has been shown to exhibit greater diversity in both base sequence selectivity as well as a propensity to bind different types of DNA structures (41-46). Utilizing photo-affinity approaches and the photo-reactive analogue of actinomycin D, 7-azidoactinomycin D, Graves and co-workers demonstrated that the drug would bind additional high affinity sites apart from the traditional 5'-XGCY-3' sequence motif (47, 48). The 5'-T(G)\textsubscript{n}T-3' motif where \( n = 1, 2, 3, \) or 4 guanine residues revealed a strong correlation between the number of guanine residues and the binding affinity. The 5'-TGGGT-3' sequence had the highest observed DNA-binding affinity (~10\textsuperscript{6} M\textsuperscript{-1}) for the sequences that were investigated (44). In addition to binding non-traditional sequences (i.e. non 5'-GpC-3' steps), actinomycin D was shown to interact with non-duplex DNA structures as well; actinomycin D exhibited strong affinities for DNA hairpins and single stranded DNAs (49, 50). In 2009, we reported that actinomycin D binds to and stabilizes the G-quadruplex structure formed from the
human telomeric sequence. This is the focus of Chapter III of this dissertation (46).

These results were quickly supported by another research group who demonstrated the binding of actinomycin D to an alternate G-quadruplex structure whose sequence is found within the promoter region of the c-myc oncogene (51).

After more than six decades it continues to be apparent that actinomycin D is a DNA binding ligand with fascinating and unique properties. For example, the majority of DNA intercalators are inhibitors of DNA replication. In contrast, actinomycin D was reported to be a transcriptional inhibitor by Sobell and the mechanistic details are still being debated (52). It is highly probable that future studies focused on the biophysical properties of actinomycin D-nucleic acid interactions will reveal additional novel interactions and biological activities with DNA sequences and structures.

**Protein Interactions with Nucleic Acids**

The intermolecular forces that govern protein interactions with nucleic acids are similar to those discussed previously with small molecules. Common non-covalent interactions include: electrostatic interaction between charged elements of the protein and the DNA, hydrogen bonds between binding domain amino acids and DNA bases, and hydrophobic interactions involving non-polar amino acid residues. These interactions fall into categories of specific and non-specific interactions, yet both are critical for biological function.
Ribonucleoprotein A1

The DNA binding protein that is a focal point of this dissertation is a member of the heterogeneous ribonucleoprotein (hnRNP) family. hnRNP A1 is a constituent of a large class of proteins that are closely associated with DNA transcription (55). This protein has both DNA and RNA binding capabilities and shares a highly conserved nucleic acid binding motif that is commonly found in hnRNP proteins. hnRNP proteins are closely associated with messenger RNA, and it is thought that the binding of hnRNP A1 facilitates the packaging of pre-mRNA, influence mRNA splicing sites, and is potentially involved in transport of mRNA to the cytoplasm (56). Recent studies by Chabot, et al. suggest that the two nucleic acid binding domains of hnRNP A1 play a role in the recruitment of telomerase to the telomeric regions of the chromosome. Recruitment of telomerase could be achieved by binding the telomeric DNA with one binding domain and the second binding domain interacting with the RNA transcript located within telomerase. This would bring the two participants in telomere elongation in close proximity to each other for the initiation of reverse transcription by telomerase (57, 58). However, a single stranded telomeric DNA template is necessary for the binding of telomerase and elongation of the telomere. At the single stranded regions where telomerase binds is the (TTAGGG)$_n$ sequence that has the propensity to form G-quadruplex structure.

Our studies have identified a protein that selectively binds and destabilizes the G-quadruplex structure. One of these proteins is a truncated amino acid sequence of hnRNP A1 that retains the two nucleic acid binding domains named Unfolding Protein 1 (UP1) (59). UP1 is a shortened version of hnRNP A1 by removal of the glycine rich
portion of C-terminal region of the hnRNP A1 protein resulting in a 196 amino acid residue UP1 molecule. In 1997, Xu reported a crystal structure of UP1 describing the structural features as compared with the parent hnRNP A1 protein. These studies demonstrated a highly conserved folding topology between the hnRNP A1 and UP1 (60). The crystal structure derived from the protein data bank (PDB ID: UP1) can be seen below in Figure 8.

![Figure 8](image.png)

**Figure 8.** Molecular modeling representation of UP1 rotated by 90 degrees. Image reproduced from PDB ID: UP1

The studies conducted by Chabot, *et al.* provided the first evidence that UP1 could bind telomeric sequences, and that UP1 was closely associated with telomerase and could induce telomeric extension in vitro (57).
Rationale for Research Project and Overview of Dissertation

As of 2010, approximately 40% of all anticancer agents used in clinical settings have nucleic acids as their primary target. Hence, the discovery of novel DNA binding agents that demonstrate sequence and structural specificity has proven very effective in the development of new drugs for the treatment of cancer. Over the past decade, the discovery of the G-quadruplex structural motif and the presence of G-quadruplex forming sequences in the human genome have led to a dramatic increase in interest in their biophysical properties and potential as targets for drug development. There is a need for fundamental biophysical characterization of the factors which influence G-quadruplex structure and stability, as well as the elucidation of the thermodynamic driving forces behind small molecule and protein interactions. This dissertation encompasses these goals through research that include the fundamental biophysical properties of G-quadruplex stability, the interactions of an anticancer agent (actinomycin D) with the human telomeric G-quadruplexes, and the binding and subsequent unfolding of human telomeric G-quadruplex structures by the protein UP1.

The earlier results reported by the Graves’ laboratory regarding the binding of actinomycin D to atypical binding sites led to our investigation of the potential interactions of actinomycin D with the G-quadruplex structural motif. These studies are the focus of Chapter III and include a combination of calorimetric and circular dichroic methods that revealed the energetic and structural implications of the novel interaction of actinomycin D with the human telomeric G-quadruplex structure. Collaborative work with the laboratory of Dr. Jinbaio Ma, conducting isothermal titration calorimetry on single stranded DNA binding to a human nucleoprotein, UP1, led to the expansion of our
research to include a detailed examination of UP1 and its interactions with the human
telomeric G-quadruplex DNA is described in Chapter IV. This chapter investigates the
ergetic and structural aspects of the binding and unfolding of unimolecular and
bimolecular telomeric G-quadruplex structures. Chapter V focuses on the effect of
sequence mutations in the loops on the structure and stability of the human telomeric G-
quadruplex. Calorimetric and circular dichroism techniques were employed to
understand the influence of \((TTAGGG)_n\) to \((TTTGGG)_n\) sequence mutations on the
structure and stability of the G-quadruplex DNA formed in Na\(^+\) buffer solutions. Chapter
VI brings together the results of the previous two chapters and characterizes the effect of
the loop sequence mutations on the recognition, binding and destabilization of telomeric
G-quadruplex DNA by the UP1 protein. In addition, the influence of two DNA binding
domains in the UP1 proteins ability to bind and unfold the G-quadruplex structure was
also investigated.

**Outline of the Dissertation**

III. THE INTERACTION OF ACTINOMYCIN D WITH HUMAN TELOMERIC G-
QUADRUPLEX DNA

A. Binding energetics investigated by isothermal titration calorimetry
B. Structural modulation upon complex formation
C. Determination of the mode of binding by circular dichroism
D. Thermal characterization of free and bound complexes

IV. BIOPHYSICAL CHARACTERIZATION OF THE BINDING AND UNFOLDING
OF HUMAN TELOMERIC G-QUADRUPLEX DNA BY UNFOLDING PROTEIN
1 (UP1)

A. The binding and destabilization of bimolecular and unimolecular G-
quadruplex DNA by UP1
B. Structural analysis of G-quadruplex unfolding by circular dichroism
C. Calorimetric determination of the binding energetics and stoichiometry
V. INFLUENCE OF LOOP MUTATIONS ON THE STRUCTURE AND STABILITY OF HUMAN TELOMERIC G-QUADRUPLEX DNA

A. Thermal characterization of G-Quadruplex DNA structure by DSC
B. Effect of loop sequence mutation on the thermal stability and unfolding
C. Parsing of the thermodynamic contributions to G-quadruplex stability
D. Correlation of biphasic melting profiles and loop sequence mutations

VI. INFLUENCE OF LOOP SEQUENCE MUTATIONS ON THE BINDING AND UNFOLDING OF HUMAN TELOMERIC G-QUADRUPLEX DNA BY UNFOLDING PROTEIN 1 (UP1)

A. Investigation of the basis for recognition of G-quadruplex DNA by UP1 by mutations of the loop sequence
B. The effect of loop mutations on the energetics and unfolding of G-quadruplex structure by UP1
C. Investigation of the synergistic aspects of the DNA binding domains of UP1 and G-quadruplex unfolding

VII. CONCLUSIONS
CHAPTER II
MATERIALS AND METHODS

This section provides a general description of the preparation and characterization of the G-quadruplex forming oligonucleotides and the biophysical methods used throughout the dissertation. More detailed experimental methods that are specific to each of the individual projects are included in their respective chapters.

Buffer Preparation

Sodium or potassium phosphate buffer at pH 7.0 was used for the experiments conducted in chapters III and IV. The concentration of sodium or potassium phosphate was 0.01 M with the ratios of mono- and dibasic forms to yield a pH of 7.0 were calculated using the Henderson-Hasselbach equation. Solid sodium chloride (NaCl) or potassium chloride (KCl) was added to the buffer solutions to provide a final concentration of 0.1 M. Ethylenediaminetetraacetic acid (EDTA) was at a final concentration of 0.001 M to act as a preservative. The pH was measured and adjusted as needed to 7.0, and the buffer solutions were filtered using 0.45 micron filters. The phosphate buffer abbreviation used throughout this dissertation is BPES. The experiments with the UP1 protein in Chapters IV and VI were conducted in Tris-hydrochloride (0.020 M Tris-HCl, 100 mM NaCl or KCl) buffer at pH 8.0. The 0.020 M Tris-HCl was prepared similar to the manner described above and used for dilutions and for solvation of the G-quadruplex oligonucleotides.
Oligonucleotide Preparation

G-quadruplex forming oligonucleotides were purchased commercially from Midland Certified Reagents, Midland, Texas. The lyophilized samples were reconstituted at a concentration of 1 mM in the appropriate buffer and allowed to hydrate for ~24 hours at 4 °C. The oligonucleotide solution was then filtered using 0.45 micron syringe filters to remove any particulate. To ensure G-quadruplex formation the samples were annealed using a MJ Research Mini-cycler with a heating and cooling cycle consisting of a rapid increase in temperature from 25 to 85 °C, followed by a five minute hold time at 85 °C. The sample was slowly cooled at a rate of 0.1 °C per minute and held at 15 °C for 1 hour. The oligonucleotide concentration was determined using a Cary 100 UV-Visible spectrophotometer. Four replicates of 5 µL stock solution was injected into a 1-cm pathlength cell containing BPES buffer and allowed to equilibrate at 90 °C for 3 minutes prior to each UV-Vis measurement to thermally denature the G-quadruplex structure. The absorbance at 260 and 280 nm was recorded to calculate the concentration and assess the purity of the synthetic oligonucleotide. The ratio of the absorbance at 260 and 280 nm was used for the determination of sample purity and all samples had a ratio of 1.8 or greater. The concentration was determined using the relationship described by the Beer-Lambert law represented in Equation 1.

\[ A = \varepsilon * l * c \]  

[1]

A is the total absorbance, \( \varepsilon \) is the molar extinction coefficient \((M^{-1}\cdot cm^{-1})\), \( l \) is the path length of the cuvette (cm), and \( c \) is the concentration in moles per liter \((M^{-1})\). The
extinction coefficients at 260 nm ($\varepsilon_{260}$) used for the 22 nucleotide human telomeric sequence and the loop mutated sequences were calculated using the nearest neighbor method as described in the literature (61).

**Characterization of G-Quadruplex Folding Topology**

Prior to analysis of ligand or protein interactions with G-quadruplex DNAs, it is necessary to examine the structure of the oligonucleotide to ensure proper folding of the G-quadruplex. The most common technique used to characterize G-quadruplex structure is circular dichroism (CD) spectropolarimetry. The CD signal is a result of the differential absorbance of right and left handed circularly polarized light by an asymmetric molecule. The CD signal at each wavelength, ellipticity ($\theta$), is converted to molar ellipticity with the units mdeg·M$^{-1}$·cm$^{-1}$. The CD spectra of a variety of G-quadruplex conformations have been measured by CD and distinguishing spectra are recorded for structures that possess differences in strand polarity. The strand polarities induce specific base orientations relative to the ribose sugar (syn or anti conformations) for proper base pairing for the formation of the G-tetrad (63). For example, base orientations for each tetrad of the unimolecular antiparallel G-quadruplexes formed in Na$^+$ buffer and the unimolecular G-quadruplex formed in K$^+$ buffer are shown in Figure 9. Stacking interactions between adjacent tetrads with alternating base orientations of the unimolecular G-quadruplex gives rise to the characteristic CD signals that are used for assignment of the folding topologies of the Na$^+$ and K$^+$ telomeric G-quadruplexes (64).
Figure 9. Base conformations of the guanines in the G-tetrads formed by the Na\textsuperscript{+} and K\textsuperscript{+} G-quadruplex DNA structures.

The stacking of guanines with opposite polarities in the G-quadruplex structure gives rise to a positive CD signal at 295 nm and the stacking of guanines with similar polarities in the parallel G-quadruplex is responsible for the positive CD signal at 260 nm (65). The antiparallel G-quadruplex formed in Na\textsuperscript{+} BPES buffer is characterized by a positive peak at 295 nm and a negative ellipticity at 265 nm. In contrast, an all parallel G-quadruplex structure formed by a tetramolecular G-quadruplex is characterized by a positive peak at 264 nm and a negative ellipticity at 240 nm. The G-quadruplex formed in K\textsuperscript{+} BPES buffer solution has been characterized by NMR and has both parallel and antiparallel strand orientations. The spectrum of this structure displays a CD signal that is characteristic of both strand orientations with a positive band at 290 nm, a positive shoulder at 268 nm, and a negative band at 240 nm. Characteristic CD spectra for the Na\textsuperscript{+} and K\textsuperscript{+} G-quadruplex conformations of the human telomeric sequence are provided in Figure 10. The CD spectra for these two conformations of G-quadruplex DNA and allow for easy comparisons of the structural properties of the solvated oligonucleotide prior to biophysical studies being performed.
Figure 10. Circular dichroism spectra of the unimolecular G-quadruplexes formed by the human telomeric sequence. The G-quadruplex formed in Na\(^+\) buffer (black line) and K\(^+\) buffer (red line).

**Isothermal Titration Calorimetry**

*General Thermodynamic Principles*

Isothermal titration calorimetry (ITC) experiments provide a direct method for measuring the enthalpy of binding (\(\Delta H_{\text{bind}}\)) for ligand-macromolecular interactions (66,67). The development of high sensitivity microcalorimeters allows for precise measurements of enthalpy changes and the simultaneous determination of the equilibrium constant, (\(K_{\text{eq}}\)), and binding stoichiometry, (\(n\)), for biomolecular interactions. Upon obtaining the equilibrium constant and change in enthalpy, additional thermodynamic values such as the free energy change, (\(\Delta G^o\)), and entropy, (\(\Delta S^o\)), can be easily derived from the thermodynamic expressions in equations [2] and [3].
\[ \Delta G^\circ = -RT \ln K_{eq} \] [2]

\[ \Delta G^\circ = \Delta H^\circ - T \Delta S^\circ \] [3]

Calculation of the change in free energy from the equilibrium constant is accomplished using equation [2] where R is the gas constant (cal/mol·K) and the temperature, T, expressed in Kelvin. The change in entropy is then calculated from the enthalpy change and the Gibb’s free energy relationship in equation [3]. For a reaction to be spontaneous the free energy change must be negative. Hence, exothermic reactions (-ΔH) with positive entropy values are favorable for binding and are enthalpically driven. Alternatively, endothermic reactions (+ΔH) require the entropy term to be large and positive for the reaction to be spontaneous. An additional thermodynamic parameter that can be determined from ITC experiments is the change in heat capacity, (ΔC_p°). The temperature dependence of the enthalpy change at constant pressure is defined as change in heat capacity (ΔC_p°) and is described by equation [4]. Interactions with apparent changes in heat capacities are often seen in coupled equilibrium and/or structural changes upon complex formation (68).

\[ \Delta C_p^\circ = \left( \frac{\delta \Delta H^\circ}{\delta T} \right)_p \] [4]

*Instrumentation*

The VP-ITC (Microcal, Northampton, MA) is an adiabatic calorimeter that uses a differential technique to measure the temperature of a reference and sample cell. The
amount of released heat or absorbed heat from a reaction is measured as titrant is injected into the sample cell while stirring. The cells are constructed of a corrosion resistant Hastelloy® Alloy 276 material and are housed in an adiabatic shield to prevent heat exchange with the surroundings. The difference in temperature between the sample cell and reference cell is measured as differential power necessary to maintain the reference and sample cells at the same temperature. The experiment is under computer control, and the experiment temperature, injection volumes, stirring rate, and resting time between injections are set by the user prior to initiation of the experiment.

Experimental Considerations

Prior to conducting binding experiments using ITC, the concentration of the macromolecule, the number of binding sites, and the binding affinity is estimated in order to achieve a suitable binding isotherm and saturation of the binding sites. The technique is dependent upon the heat given off or absorbed during the reaction. Therefore, the \( \Delta H_{\text{bind}} \) cannot be zero at the chosen experimental temperature. The typical shape of an ITC titration is sigmoidal and in order accurately fit the data with the fitting algorithms integrated into the Microcal VP-ITC software, the slope of the sigmoid must not be too sharp (very strong binding) or too shallow (very weak binding). If an estimate of the equilibrium constant and binding stoichiometry is known a useful parameter for determining the concentration of the macromolecule to be placed in the sample cell is the unit less \( c \) value described by equation [5], where \( M_t \) is the total macromolecule concentration and \( n \) is the stoichiometry of the reaction.
\[ c = K_{eq} \cdot M \cdot n \]  \hspace{1cm} [5]

The ideal \( c \) value is 40 and will produce a sigmoidal isotherm when the calculated total macromolecule is titrated with ligand that is approximately 10 times the concentration of the macromolecule in the sample cell. Figure 11 is a representation of the theoretical shape of estimated binding isotherms as the \( c \) value is varied.

![Simulated binding isotherms with various \( c \) values.](image)

Figure 11. Simulated binding isotherms with various \( c \) values.

The samples are prepared by dilution of a stock concentration of macromolecule and ligand to satisfy the concentration requirements dictated by the calculation of the \( c \) value. The reference and sample cells hold approximately 1.6 mL of solution; however it is necessary to prepare 2 mL to properly load the reference and sample cells. Similarly, the injection syringe has a volume is 300 \( \mu \)L, but a minimum volume of 750 \( \mu \)L is required to load the syringe. Prior to loading the cells and injection syringe, the buffer, macromolecule, and ligand solutions are degassed with stirring for a minimum of 10 minutes. Using the buffer solution for the reference cell and macromolecule solution for the sample cell are then loaded into the respective cells using a 2.5 mL blunt end
Hamilton® syringe. The solutions are slowly introduced into the cells with short bursts of
the syringe toward the end of the filling procedure to remove any air bubbles that may
have lodged in the cells. The injection syringe is loaded with the titrant by the automated
process provided in the Microcal VP-Viewer 2000 software.

The remaining steps are to set the experimental parameters of temperature,
number of injections, stirring rate, injection volume, and resting time between injections.
The number of injections will depend on the injection volume, but typically is set to 30-35
injections of 5-8 µL. Higher stirring rates are needed for more viscous solutions and
time between injections is dependent upon the reaction kinetics. Typical resting time that
allows for the signal to return to baseline prior to the next injection is 200-250 seconds.
However, the resting times can be extended in the event that the kinetics of the binding
interaction is slow. The final step is to input the syringe and cell concentrations to
normalize the data for the fitting of the data. For proper data analysis, an additional
titration should be conducted with the ligand injected into the sample cell containing only
the buffer used for dilution to account for the heat of dilution of the titrant.

The direction of the titration (forward or reverse) is a decision that must be made
during the experimental design process and is dependent upon the nature of the
interaction (69). For simple ligand-DNA or ligand-protein binding interactions the
direction of the titration is forward and the macromolecule is placed in the sample cell
and the ligand is placed in the syringe. Some situations may arise in protein-protein or
protein-DNA interactions where the definition of the ligand is not clear. Additionally,
binding interactions that display high cooperativity, multiple binding sites, or a structural
change upon binding may be present that results in complex binding isotherm shapes. In
these instances an experiment that is a reverse titration may be beneficial to provide sigmoidal binding isotherms that can be easily fit by the provided non-linear least squares binding models integrated into the Microcal VP-ITC software (70).

Data Analysis

The sigmoidal shape of the binding isotherm is due to the initial injections of ligand evolving the maximum heat change and diminishing to baseline as the number of injection increases and the binding sites are saturated, an example provided by Figure 12.

![Figure 12](image)

Figure 12. Raw ITC binding isotherm for 40 injections of a single stranded oligonucleotide into its complementary strand.

The raw data is integrated with respect to time using the concentration of the injectant, and the molar ratio is calculated with the concentration of the macromolecule. The result is a plot that is ΔH of injectant (kcal/mol) versus the molar ratio and the thermodynamic parameters is calculated from a fit of the sigmoidal isotherm. The simplest case for fitting of the data is an experiment that has n-independent and identical binding sites. However, more complex binding models are integrated into the Origin 7.0
software or unique binding models can be written and utilized to fit the data. Figure 13 provides an illustration of the integrated raw data for duplex formation and fit with a single site binding model.

Figure 13. The integration of the raw ITC data (black squares) for the titration of complementary oligonucleotides. The red line represents the best fit line of a single site binding model.

Differential Scanning Calorimetry

GENERAL THERMODYNAMIC PRINCIPLES AND INSTRUMENTATION

Differential scanning calorimetry (DSC) is similar to the previously discussed ITC method with key differences in how the data is collected. Whereas ITC conducts binding experiments at a constant temperature, the DSC measures the heat evolved during denaturation as a function of temperature (71). The Microcal VP-DSC instrument
is similar to the VP-ITC depicted in Figure 9 with a sample and a reference cell encased in a thermal shield. The thermal shield is surrounded by a heating jacket that increases or cools the sample and reference cell at the same rate and the difference it heat energy uptake is recorded between the two cells as the macromolecule denatures. The reference and sample cell are pressurized to allow heating to 120 °C. The DSC directly measures the excess heat capacity, $C_p$, as the macromolecule denatures from a folded to an unfolded state. At constant pressure the enthalpy change, $\Delta H$, as the temperature is increased is related to the heat capacity, $C_p$, as described by equation [6]. The area under the curve is the calorimetric enthalpy of the transition.

$$\Delta H = \int_{T_1}^{T_2} C_p \, dT \quad [6]$$

Additionally, the change in entropy, $\Delta S$, can be determined from a DSC experiment using equation [7] at constant pressure.

$$\Delta S = \int_{T_1}^{T_2} \frac{C_p}{\tau} \, dT \quad [7]$$

Once these two thermodynamic quantities are known the free energy change, $\Delta G^0$, is determined by equation [2] and the thermodynamic parameters of the denaturation of the macromolecule can be calculated. The melting temperature, $T_m$, of the macromolecule is also determined from DSC thermograms and is defined by the mid-point or the transition or the maximum of the observed peak. The $T_m$ is a useful comparison for the stability of nucleic acids structures and represents the temperature at which 50:50 percent
distributions of the fractions in the folded and unfolded state. Cartoons representing examples of the determination of $\Delta H$, $\Delta S$, and the $T_m$ are illustrated in Figure 14.

![Figure 14](image)

Figure 14. The determination of the thermodynamic parameters of the transition for a typical DSC thermogram.

*Experimental Considerations*

Experimental design for DSC is relatively simple. The basic requirement is that the concentration of the macromolecule be sufficient to provide an accurate signal to measure. For nucleic acids, this is typically a concentration of 100-150 $\mu$M. Prior to starting the experiments the cells must be thoroughly cleaned to prevent contamination and errors in the differential heat measurement. The cells should be vigorously washed with 250 mL of deionized H$_2$O, or with Contrad-70 detergent followed by a 300 mL of deionized H$_2$O. The buffer that is used for the dilutions should be degassed for 10 minutes and loaded into the sample and reference cells. The temperature range and the scan rate are determined by the user, and the first series of buffer/buffer scans are
initiated. The first scan of a DSC experiment has no thermal history and is a “dummy” scan that is not reproducible. However, the subsequent scans should be highly reproducible and suitable for the baseline subtractions methods used in the data analysis. After the buffer/buffer scans are complete, the sample solution is degassed and loaded into the sample cell with the buffer solution remaining in the reference cell. Multiple scans should be conducted to assess the reversibility of the denaturation and to achieve the best thermogram to be used for data analysis.

Data Analysis

Data analyses of the DSC thermograms were conducted using the Microcal Origin 7.0. The general procedure for data analysis is to subtract the reference scan, normalize the data, baseline correction, and integration of the area under the transition curve or a fit to the appropriate model. Subtraction of the reference scan is accomplished by importing the buffer and sample scans and physically subtracting the reference scan from the sample scan. The normalization of the excess heat capacity (cal/°C) to molar heat capacity (kcal/mol/°C) must then be conducted prior to analysis. This is accomplished by inputting the concentration of the macromolecule into the Origin 7.0 software after subtraction of the reference scan. Accuracy of the concentration determination is essential for calculation of $\Delta H_{\text{cal}}$ from the area under the transition curve. The baseline correction procedure is critical. The user must be aware of the influence the iterative baseline correction methods may have on the transition (23). The Microcal software provides three options for baseline correction: progressive, cubic, and linear procedures that are highly dependent on the pre- and post-transition baselines. It is a best practice to
use the baseline correction procedure that has the least influence on the shape of the transition and consistently treat multiple sets of data with the same method. It should be noted that the baseline correction procedure will render the $\Delta C_p = 0$, which is only suitable for transitions that have minimal changes in heat capacity upon denaturing. The determination of the $T_m$ and $\Delta H_{cal}$ is then accomplished by determining the transition peak maximum and integrating the area under the curve from the baseline. The microcal VP-DSC software has integrated 2-state and non-2-state fitting models that may be used for data analysis. These models take into account the van’t Hoff enthalpy ($\Delta H_{vh}$) that is determined from the shape of the transition curve. Sharp transitions will have small $\Delta H_{vh}$ values whereas broad transition will have large $\Delta H_{vh}$ values. For a simple 2-state process the $\Delta H_{cal} = \Delta H_{vh}$ and there are no significant intermediates present in the unfolding process. The 2-state model is rarely suitable for nucleic acids because there are significant enthalpic contributions from base stacking interactions, loop interactions, and helix-coil transition in the unfolding process. The non-2-state model allows for fitting these transitions, and the magnitudes of the $\Delta H_{cal}$ and $\Delta H_{vh}$ values can be useful in determining the nature of the transition. For example, if $\Delta H_{vh} < \Delta H_{cal}$ then the unfolding process includes structural intermediates and if the $\Delta H_{vh} > \Delta H_{cal}$ the unfolding process includes either intermediates or precipitation of the sample solution. Erroneous $\Delta H_{vh}$ and $\Delta H_{cal}$ values can occur with the non-2-state model in situations where the $\Delta C_p \neq 0$. For nucleic acid transition it is best to remove the bias imposed by a particular model and directly integrate the area under the curve to provide the $\Delta H_{cal}$ and examine the y-intercept of the pre- and post-transitional baselines for $\Delta C_p$. 

48
CHAPTER III
INTERACTIONS OF ACTINOMYCIN D WITH HUMAN TELOMERIC G-
QUADRUPLEX DNA

by

JASON S. HUDSON¹, SONJA C. BROOKS¹, AND DAVID E. GRAVES¹²

¹Department of Chemistry, University of Alabama at Birmingham, Birmingham, AL 35294
²Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL 35294

Biochemistry, 2009, 48 (21), pp 4440–4447

Copyright 2009 by The American Chemical Society

Used by permission

Format adapted and errata corrected for dissertation
Abstract:

The G-quadruplex structural motif of DNA has emerged as a novel and exciting target for anticancer drug discovery. The human telomeric G-quadruplex consists of a single strand repeat of d[AGG(TTAGGG)₃] that can fold into higher-order DNA structures. Small molecules that selectively target and stabilize the G-quadruplex structure(s) may serve as potential therapeutic agents and have garnered significant interest in recent years. In the work presented here, the anticancer agent, actinomycin D, is demonstrated to bind to and induce changes in both structure and stability to both the Na⁺ and K⁺ forms of the G-quadruplex DNA. The binding of actinomycin D to the G-quadruplex DNAs are characterized by intrinsic association constants of approximately 2 x 10⁵ M⁻¹ (strand), 2:1 molecularity, and are shown to be enthalpically driven with binding enthalpies of approximately -7 kcal/mol. The free Na⁺ or K⁺ forms of the quadruplex structures differ in melting temperatures by approximately 8°C (60 and 68°C, respectively), whereas both forms, when complexed with actinomycin D are stabilized with melting temperatures of approximately 79°C. The induced CD signals observed for the actinomycin D-G-quadruplex complexes may indicate that the phenoxazone ring of actinomycin D to be stacked on the G-tetrad rather than intercalated between adjacent G-tetrads. Complex formation with actinomycin D results in changes to both the Na⁺ or K⁺ structural isoforms to “ligand-bound” complexes having similar structural properties and stabilities.
Introduction

The DNA structural motif known as the G-quadruplex has recently emerged as a novel and exciting target for the discovery and design of new classes of anticancer agents (1-3). DNA sequences that can form G-quadruplex structures are found extensively throughout the genome and are located in biologically relevant regions. These sequences and their corresponding quadruplex structures were first observed to exist in telomeric regions of DNA, located at the terminal ends of chromosomes (4-6). More recently, G-quadruplex forming sequences have been mapped to the promoter regions of a number of genes and oncogenes (7-9). It is postulated that G-quadruplex structures may serve important biological functions in the regulation of gene expression (10-12). Hence, these findings have led to an increased interest in the structural and functional features of G-quadruplex structural motif and offers potentially novel targets for the development of small molecules that could selectively target and stabilize the quadruplex structure.

The G-quadruplex consists of stacked G-tetrads connected by lateral, diagonal, or external loops and has been shown to possess a wide range of structural polymorphism within guanine rich sequences that exhibit the motif. Burge and coworkers and Dai and coworkers recently published extensive reviews of the topologies of quadruplex DNA structures (13, 14). The observed polymorphism(s) found for G-quadruplex structures include differences in molecularity, strand orientation, loop characteristics, and structural isomers based on selected cation buffer conditions. The structural stability of quadruplex DNAs have received considerable attention and are reported in a detailed review by Chaires and coworkers (15).
Common features of small molecules that have been found to interact with quadruplex DNA structures include heterocyclic planar aromatic ring systems that have the ability to “cap” the terminal ends of quadruplex structures. Hurley and coworkers have utilized the porphyrin ring system as a structural foundation of perhaps the most well known quadruplex binding ligand, telomostatin (3, 16, 17). In recent years, there has been considerable effort devoted toward the discovery of additional DNA binding ligands that would selectively target quadruplex DNA structures. A thorough review of these studies are reported by Monchaud et al (18).

Efforts in our laboratory have focused on the structural and energetic properties describing the binding of actinomycin D to a variety of nucleic acid structural motifs. Actinomycin D consists of a heterocyclic phenoxazone ring that serves as the intercalative portion of the ligand and two cyclic pentapeptide side chains that have extensive interactions within the minor groove of duplex DNA. In duplex DNA, actinomycin D has historically been demonstrated to bind preferentially to the 5’-dGpC-3’ sequence (19); however, more recently the binding of actinomycin D has been shown to exhibit greater promiscuity in both the base sequence selectivity as well as the DNA structure (20-24). As early as 1994, our laboratory demonstrated the ability of actinomycin D to bind to duplex DNA containing G-tracts which led to our hypothesis that actinomycin D may also interact with the G-quadruplex DNA structure(s) and possibly potentiate biological response(s).

In the work presented here, we report such an interaction between actinomycin D and the human telomeric sequence, d[AGGG(TTAGGG)₃]. This 22-mer deoxyoligonucleotide forms an intramolecular G-quadruplex structure that can be used to
probe the binding of ligands including actinomycin D. The studies reported here provide key insights into the structural and thermodynamic features that characterize the binding of actinomycin D to the G-quadruplex DNA motif.

Materials and Methods:

Sample Preparation: The human telomeric 22-mer of sequence 5’-AGGGTTAGGGTTAGGGTTAGGG-3’ was purchased from Midland Reagents, Midland, TX and purified by reverse phase chromatography. Stock DNA solutions were prepared by dissolving in K-BPES (0.01 M KH$_2$PO$_4$, 0.01 M K$_2$HPO$_4$, 0.001 M EDTA, 0.1 M KCl) or Na-BPES (0.01 M NaH$_2$PO$_4$, 0.01 M Na$_2$HPO$_4$, 0.001 M EDTA, 0.1 M NaCl) buffer at pH 7 for 24 hours at 4°C. The deoxyoligonucleotide samples were heated to 90°C for 5 min followed by cooling to 4°C at a cooling rate of 0.1°C per min using a MJ Research Minicycler to ensure formation of the G-quadruplex structure which was checked by circular dichroism spectroscopy. The DNA concentrations were determined by UV-visible spectroscopy using a Cary 100 UV-Visible spectrophotometer (Varian, Inc.) at 90°C and a molar extinction coefficient of $\varepsilon_{260} = 228,500$ M$^{-1}$cm$^{-1}$ (25). Actinomycin D was purchased from Sigma-Aldrich and used without further purification. The drug was in the appropriate BPES buffer, dissolved overnight at 4°C and filtered through a Millipore 0.45 µM syringe filter prior to use. Actinomycin D concentrations were determined by UV-visible spectroscopy using a Cary 100 UV-visible spectrophotometer (Varian, Inc.) at 25°C using the extinction coefficient $\varepsilon_{440} = 24,500$ M$^{-1}$cm$^{-1}$ (23).
**Isothermal Titration Calorimetry Studies:** ITC experiments were performed using a Microcal VP-ITC at 25°C. All samples were degassed prior to use. The sample cell was filled to capacity ~1.6mL with quadruplex DNA (10 µM, (strand)), and actinomycin D (200 µM) was injected in 25 aliquots of 10 µL each with 250 seconds resting time between injections. Three replicate titrations were performed and the resulting data were integrated and analyzed using the single-site binding model and fit by a non-linear least squares fitting algorithm (Microcal Origin 7.1 software) to yield the relevant thermodynamic parameters.

**Differential Scanning Calorimetry Studies:** DSC experiments were performed with a Microcal VP-DSC. The concentrations of the DNA and actinomycin D were 125 µM (strand) and 250 µM ligand. Experiments were conducted over the temperature range of 15-100°C with a scan rate of 90°C/h. Multiple heating and cooling scans were conducted in order to determine reversibility. Analyses of DSC data were performed with Microcal Origin 7.1 software in which all data were analyzed with a two transition non-two-state model.

**Circular Dichroism Studies:** CD experiments were conducted using an AVIV 400 circular dichroism spectrophotometer (Aviv, Inc). All experiments were conducted at 25°C in 1 cm pathlength rectangular cells. The data were collected from 220-350 nm at every 1 nm with a bandwidth of 3 nm. The titration of quadruplex DNA with actinomycin D was performed by injecting 37.5 µL 0.5 mM actinomycin D stock solution into a cell containing 4 µM (strand) quadruplex DNA. Each injection represents a 0.5
molar equivalent of the DNA concentration. After each injection, the solution was stirred and allowed to equilibrate for 10 minutes prior to measuring the CD spectrum. Contributions from the free actinomycin D and buffer solution were subtracted from each subsequent measurement. To measure the induced CD, the samples were prepared at saturating conditions (all actinomycin D is in bound form). The final concentration of actinomycin D and quadruplex DNA is 10 μM and 100 mM (strand), respectively. The CD spectra were collected from 335 to 500 nm at every 1 nm using a bandwidth of 3 nm.

**Results**

*CD Studies:* CD spectroscopy was used to examine the structural characteristics of the structural features of quadruplex DNA (26). Representative CD spectra for the Na⁺ and K⁺ structural conformations of the quadruplex structures can be observed in Figure 1. The Na⁺ isoform displays a characteristic antiparallel G-quadruplex with maxima at 295 and 245 nm and a minimum at 265 nm (27, 28). In contrast, the K⁺ structural isoform exhibits markedly different CD spectral characteristics due to the mixed hybrid (both parallel and antiparallel strand orientations) with two maxima at 290 and 265 nm, a trough at 277 and a minimum at 238 nm as previously described by Yang and coworkers (25, 29). The mixed parallel and anti-parallel orientations of the phosphate backbone results in a CD spectrum with both parallel and anti-parallel characteristics and shows a maximum at 295 nm, representing the antiparallel portion, and the shoulder at 265 nm and minimum at 240 nm corresponding to the parallel portions of the quadruplex structure.
CD titration experiments were used to investigate actinomycin D interactions and any subsequent structural perturbations to the quadruplex structures that may result from complex formation. Figure 2A shows the incremental titration of actinomycin D into the Na\(^+\) form of the G-quadruplex. As observed from these data, titration of actinomycin D into the quadruplex DNA results in significant changes to the CD spectrum; including the observation of an isoelliptical point at 260 nm, which suggests the presence of both free and bound quadruplex DNA in solution. With increasing actinomycin D binding, the CD signal of the quadruplex DNA is shown to exhibit significant shifts in the maximum from 295 nm to 280 nm, as well a complete loss of positive ellipticity to a negative ellipticity at 245 nm. Similarly, the titration of actinomycin D into the K\(^+\) form of the G-quadruplex as monitored by CD is shown in Figure 2B. An isoelliptical point is observed at 265 nm, again suggesting the presence of both free and bound quadruplex DNA in solution. As more actinomycin D binds to the K\(^+\) form, the trough at 277 nm disappears and becomes a maximum at 277 nm. The minimum at 238 nm exhibits a red shift to 245nm and becomes more pronounced. Although the Na\(^+\) and K\(^+\) forms of the quadruplex DNAs have very distinctive CD spectral characteristics, the CD spectra of both quadruplex forms when complexed with actinomycin D converge to very similar CD spectra with maxima ~280 nm and minima ~245-250 nm indicating similarities in structural characteristics.

In order to investigate the mode of binding, CD experiments were designed to discern whether induced CD signals could be observed upon formation of the actinomycin D-G-quadruplex complex. As observed in Figure 3, actinomycin D is shown to have a very weak CD signal (black line) due to the symmetric nature of the
molecule. Upon complex formation with duplex DNA, the intercalation of the phenoxazone ring of actinomycin D between the d(GpC) step results in the induction of two negative CD signal at 375 and 440 nm (i.e. induced CD signals). The presences of these signals are indicative of an intercalative mode of binding. In the case of actinomycin D binding to the Na\(^+\) form (red line) and K\(^+\) form (blue line) quadruplex DNA structures, induced CD signals are observed at both 375 nm and 440 nm; however, while the 375 nm band is negative, the 440 nm band is positive. A positive band at 440 nm indicates that the complexes formed by actinomycin D with both forms of quadruplex DNA are not purely intercalative, but involve end-stacking on the terminal G-tetrad of the quadruplex structures.

*Isothermal Titration Calorimetry Studies:* Thermodynamic and stoichiometric properties associated with the formation of actinomycin D complexes with quadruplex DNAs were determined by isothermal titration calorimetry. Figure 4 shows raw ITC data (panels A and B) as well as the integrated heats of complex formation for both the Na\(^+\) and K\(^+\) quadruplex DNAs, respectively, upon titration with the DNA binding ligand. The thermodynamic parameters derived from these data are summarized in Table 1. The time between injections was 250 seconds. The geometry of the injection peak revealed that with this resting time, the reaction returned to equilibrium prior to subsequent injections. From these studies, we observe that formation of the actinomycin D-G-quadruplex complex(es) for both the Na\(^+\) and K\(^+\) forms are enthalpically driven with binding enthalpies (ΔH\(_{\text{obs}}\)) approximating -7 kcal/mol (-7.3 kcal/mol for the Na\(^+\) form and -6.7 kcal/mol for the K\(^+\) form). The DNA binding affinities were determined to be 2.1 x
$10^5 \text{M}^{-1} \text{(strand)}$ for interaction of actinomycin D to the Na$^+$ quadruplex structure. Similarly, the binding affinity of actinomycin D to the K$^+$ quadruplex form was found to be $2.3 \times 10^5 \text{M}^{-1} \text{(strand)}$. Interestingly, stoichiometries of approximately 0.5(actinomycin D per quadruplex) were observed for both the Na$^+$ and K$^+$ quadruplex complexes with actinomycin D, indicating that one drug molecule could effectively bridge two quadruplex DNA structures to form a 2:1 complex, supporting our hypothesis of “end-stacking” to the terminal G-tetrad.

*Differential Scanning Calorimetry Studies:* Thermal melting profiles were obtained for both the free and bound Na$^+$ and K$^+$ forms (Figure 5A and 5C, respectively) of the G-quadruplex DNA structures using differential scanning calorimetry. These studies are used to discern the thermodynamic stabilities of the free and actinomycin D complexed quadruplex structures and are reported in Table 1. Using Origin 7.1, thermograms for the free and complexed DNA structures were analyzed using a two transition non-two state model. These thermograms reveal that both the Na$^+$ and K$^+$ forms of these DNA quadruplexes exhibit two peaks. Replicate heating and cooling scans were conducted in order to demonstrate reversibility of the denaturation of the quadruplex structures. The melting transitions of both the free and bound forms of the Na$^+$ and K$^+$ quadruplex DNAs were shown to be highly reversible (data not shown). The observation of two melting transition peaks has been previously reported by a number of investigators for G-quadruplex structures with base sequences similar to the d[AGGG(TTAGGG)$_3$] used in this study (25, 30, 31). Figure 5A shows the thermogram for the free form of the K$^+$ quadruplex. Panel B shows the concomitant thermogram of the K$^+$ quadruplex DNA
when complexed with actinomycin D. In Fig. 5A, the free form of the quadruplex DNA is shown to consist of two transitions; the first transition exhibits a melting temperature at 60°C and the second at 68°C. Upon binding of actinomycin D, both transitions are observed to stabilize to higher temperatures of 64°C and 79°C, respectively. For the Na⁺ form of the quadruplex DNA, panels C and D in Figure 5 depict the thermograms of the free and bound forms, respectively. As observed for the K⁺ form, two transitions are evident with melting temperatures of 40°C and 60°C, respectively; however, upon actinomycin D binding, both transitions are stabilized to higher temperatures, 65°C and 79°C, respectively.

Discussion

In this work, we demonstrate the binding of actinomycin D to both the structurally distinct Na⁺ and K⁺ forms of quadruplex DNA structural motifs. These binding studies utilized a variety of spectroscopic and calorimetric methods to probe the structural and thermodynamic features associated with complex formation. Upon binding of actinomycin D, significant structural changes to both the Na⁺ and K⁺ forms of the DNA quadruplexes are indicated by the change in maxima and minima as well as losses in ellipticity as observed from the CD spectra. Prior to addition of actinomycin D to the Na⁺ or K⁺ forms of quadruplex DNA solutions, the two structurally distinct species exhibit unique CD signals. However, upon binding actinomycin D, both species are demonstrated to undergo significant changes as evidenced by their CD spectra, and at saturation both species are shown to converge to very similar CD signals, indicating similarities in the structures of the bound quadruplex DNA species. Although
considerable care must be taken in the interpretation of CD spectra from ellipticity wavelengths to unequivocal structural determinations (26), the results observed in this study are indicative that upon formation of a complex with actinomycin D, the two structurally distinct quadruplex species converge into common bound conformations that have markedly similar CD spectra. DNA quadruplexes possess the ability to exist in many isomeric forms that have considerable variation in loop orientation, the number of tetrads, and strand molecularity. Chaires and coworkers have stated that there are at least 26 possible structures that can be formed; however, only 6 have being observed thus far in vitro (15). Current solution structures for the quadruplex exist for the anti-parallel Na$^+$ form as solved by NMR (32). Additionally, the NMR structure of the K$^+$ form of a mixed parallel anti-parallel structure has been reported (33), as well as an X-ray diffraction structure that features a more compacted G-quadruplex consisting of edgewise loop orientations (34).

The characteristic CD signals for different quadruplex structures are thought to arise from alterations of the glycosidic bond angles of guanine and sugar phosphate backbone. The CD spectra presented in Figure 1 are characteristic for the Na$^+$ and K$^+$ forms and reveal typical maxima and minima for the anti-parallel and mixed hybrid parallel structures (25, 26, 33). The observed shift in CD spectra for both the Na$^+$ and K$^+$ forms of the quadruplex from the 295 nm (characteristic of anti-parallel structure) to a lower wavelength centered around 280 nm as well as the observation of negative ellipticity at 245 nm is characteristic of a parallel stranded quadruplex. From the CD data presented here, complex formation with actinomycin D is shown to drive a change in the quadruplex structure. We postulate that the pronounced changes in CD signals, for both
the Na\(^+\) and K\(^+\) forms of the quadruplex DNAs, and convergence to relatively similar spectra arise as a result of the quadruplex DNA structurally rearranging to more favorably accommodate the binding geometry of the actinomycin D. Such phenomena have been observed previously wherein DNA binding ligands were shown to direct the folding of the G-quadruplex into an alternate structure. In these studies, Hurley and coworkers demonstrated the accelerated assembly of intermolecular quadruplex DNA structures by TMPyP4 (35). Similarly, Hurley and coworkers reported the binding of telomostatin to the G-quadruplex in the absence of Na\(^+\) and K\(^+\) ions that resulted in a structural change to the anti-parallel form of the G-quadruplex (36). Recent studies reported by Chaires et al (37) examining binding kinetics of Na\(^+\) and K\(^+\) and the ligand to G-quadruplex DNA reveal a rapid phase complete within 5 milliseconds followed by two slower phases of 40-50 seconds and 600-800 seconds. Chaires attributes these relaxation times to a rapid cation binding phase followed by slower kinetics of quadruplex refolding. Binding of the TMPyP4 was postulated to facilitate a conformational switch from the Na\(^+\) antiparallel structure to the K\(^+\) hybrid. Based on the evidence presented in their manuscript, a complete refolding of the quadruplex DNA requires extended equilibration times. The CD data presented in Figure 2C indicates that the binding of actinomycin D the Na\(^+\) form of the G-quadruplex structure induces a structural change to the quadruplex DNA whose CD-spectrum more closely resembles that of the K\(^+\) isoform. These data are supportive of Chaires’ hypothesis of a “conformational switch” that is induced by a DNA binding ligand. However, additional structural data are required to discern the exact nature of the actinomycin D-G-quadruplex complexes. Of particular interest and unique to the study that is presented here is that actinomycin D is shown to
bind to both the Na\textsuperscript{+} and K\textsuperscript{+} forms of the quadruplex and directs both quadruplex structures to converge to unique “complexed” structures that have very similar CD characteristics as shown in Figure 2C. The thermodynamic binding parameters for binding of actinomycin D to both the Na\textsuperscript{+} and K\textsuperscript{+} structural isoforms of the G-quadruplex DNAs are quite similar, regardless of the initial starting structural form. From ITC studies, similar binding affinities (2.1 and 2.3 \times 10^5 M^{-1} strand) for the binding of actinomycin D to the Na\textsuperscript{+} and K\textsuperscript{+} structural isoforms of G-quadruplex DNAs, respectively, are observed. The binding of actinomycin D to both the Na\textsuperscript{+} and K\textsuperscript{+} forms of the quadruplex DNA were shown to exhibit favorable binding energies (\Delta G \sim -7.3 kcal/mol). As shown in Table I, parsing of the energetic components reveal that the binding of actinomycin D to both G-quadruplex structures to be enthalpy driven with \Delta H \sim -7 kcal/mol. The binding stoichiometries indicate that upon complex formation, actinomycin D is complexed with two quadruplex structures. Literature reports of binding stoichiometries for small molecule interactions with quadruplex DNA are highly varied; however, terminal stacking of heterocyclic planar ligands on the plane of the G-tetrad have been observed for several co-crystal complex structures (38). Indeed, in the case of daunomycin, the anthracycline ring is shown to stack on the terminal G-tetrad and may offer insights into the mode of binding for actinomycin D to the quadruplex DNA (39).

In an effort to discern structural features associated with the actinomycin D-quadruplex DNA complex, experiments were designed to probe the presence of an induced CD signal. In duplex DNA, actinomycin D has been demonstrated to intercalate between adjacent dGpC steps, resulting in an induced CD signal, with negative bands
observed at 375 and 440 nm (40). In the case of actinomycin D binding to both the Na\(^+\) and K\(^+\) forms of quadruplex DNAs, induced CD signals for both complexes are observed. As described above, negative induced CD bands are observed at 375 nm. In contrast to intercalative binding which exhibits an additional negative band at 440 nm, we observe an induced signal at 440 nm; however, this band is now positive. These observations of induced CD signals reveal an atypical binding interaction (i.e., nonintercalative) and support our premise of “end-stacking” of the phenoxazone to the terminal G-tetrad of the quadruplex DNA. There are several examples of mixed (negative and positive) induced CD signals including the binding of actinomycin D to non-dGpC containing sequences (24) as well as stacking with d(GTP) (41).

The structural stabilities of the quadruplex DNA with and without actinomycin D were probed by differential scanning calorimetry (DSC). Figures 5A and 5C reveals thermograms of the K\(^+\) and Na\(^+\) structural isoforms of the G-quadruplex DNA, respectively, in the absence of actinomycin D. These melting profiles consist of two transitions for both the K\(^+\) and Na\(^+\) forms with the predominant melting transition at 60\(^\circ\)C for the Na\(^+\) and for the K\(^+\) form, a more stable transition observed at 68\(^\circ\)C. These values are in excellent agreement with previously published results as summarized by Lane et al. (15). Upon binding of actinomycin D to the Na\(^+\) and K\(^+\) forms of G-quadruplex DNAs, DSC thermograms of both the ligand-DNA complexes are observed to converge with comparable melting transitions at 79\(^\circ\) C. These data provide further evidence that actinomycin D binds to different G-quadruplex structural isoforms (Na\(^+\) and K\(^+\) forms) and the resulting ligand-DNA complexes that are formed are similar in structure and stability. The presence of two peaks for the melting transitions may be indicative of the
existence of different structural isomers in solution (30) or the presence of a stable intermediate that is formed in the unfolding pathway. In 2007, Sheardy and Chaires identified the presence of a stable unfolding intermediate in the human telomeric sequence in K\(^+\) solution using the application of singular value decomposition (SVD) to quadruplex unfolding studies (30). The studies presented here suggest that upon binding actinomycin D, both the major and minor DNA components observed in the DSC thermograms form complexes with the ligand and both components are converted to more energetically stable complexes. In the case of the K\(^+\) quadruplex DNA, the lower melting transition component is stabilized by 4\(^\circ\)C, from 60 to 64\(^\circ\)C; however, for the Na\(^+\) form, the lower melting transition component is stabilized from 39 to 65\(^\circ\)C. Although our data do not differentiate whether this lower melting transition component is a structural isomer or stable unfolding intermediate of the predominant Na\(^+\) and K\(^+\) quadruplex DNA structure, it is clear that the lower melting transition component for both the Na\(^+\) and K\(^+\) quadruplex forms are stabilized upon complex formation with actinomycin D. Similarly, the dominant melting transition components observed for both the Na\(^+\) and K\(^+\) quadruplex forms are shown to be significantly stabilized upon complex formation with actinomycin D. The melting temperatures of the K\(^+\) and Na\(^+\) G-quadruplex structural isoforms are determined to be 68\(^\circ\)C and 60\(^\circ\)C, respectively. DSC thermograms obtained in the presence of actinomycin D reveal both major melting transition components to exhibit significant stabilization; the K\(^+\) form is shifted to 79\(^\circ\)C (an 11\(^\circ\)C stabilization) and the Na\(^+\) form to 79\(^\circ\)C (a 19\(^\circ\)C stabilization). These data are indicative of the strong propensity of actinomycin D to form stable interactions with quadruplex DNAs resulting in the formation of bound complexes whose thermal
stabilities are markedly enhanced from native quadruplexes and converge to very similar bound states.

In the data reported here, we present a novel interaction between the antibiotic actinomycin D and both the Na\(^+\) and K\(^+\) structural isoforms of quadruplex DNA. The CD and calorimetric data demonstrate that actinomycin D forms strong stable complexes with both cationic forms of the quadruplex DNA in a thermodynamic and structurally similar manner. Both the CD and DSC data reveal that initial quadruplex DNA structures that are unique for either Na\(^+\) or K\(^+\) environments undergo modifications to accommodate the formation of the actinomycin D complex and the resulting complexes that are formed are similar in both structure and enhanced stability. Although the exact structure of the actinomycin D-quadruplex DNA complex(es) remain unknown, the CD data is suggestive of a more parallel stranded quadruplex. The observation of a 2:1 (quadruplex to ligand) binding ratio as well as the induced CD signals upon complex formation provide considerable insights into a proposed binding mechanism for actinomycin D to the quadruplex DNA. Our studies are consistent with a stacking of the phenoxazone ring of actinomycin D to a terminal G-quartet of the quadruplex DNA. We speculate that this interaction is further stabilized by the interaction of the pentapeptide sidechains of actinomycin D with the shallow grooves of the face of the quadruplex DNA formed by the sugar phosphate backbone. It is plausible for the ligand to recruit a second quadruplex structure that would effectively stack its terminal G-quartet to the opposite face of the phenoxazone ring and allow the second pentapeptide sidechain to orient in the grooves of the second quadruplex DNA, resulting in the 2:1 binding stoichiometry.
In 2005, Jovin and coworkers examined the binding of actinomycin D to single stranded DNA and demonstrated several key elements of actinomycin D binding that are pertinent for the binding of this ligand to quadruplex DNA (42); the binding is directed by the stacking of the planar phenoxazone ring with guanine residues. This stacking interaction can occur from either face of the phenoxazone ring, and serves as an anchor for structural rearrangements of the DNA strands which appear to be quite mobile, allowing the DNA to find the most stable conformation around the actinomycin D anchor. The principle driving force for the binding of actinomycin D to a single-stranded deoxyoligonucleotide was postulated to be through an initial stacking and hydrogen bonding between the phenoxazone chromophore and the guanine residue. Jovin observed that the stacked actinomycin D-guanine complex served as an anchor that allowed rearrangements of the DNA strands to orient in an energetically stable conformation with the two pentapeptide sidechains of the actinomycin D. Similarly, we propose that the binding of actinomycin D to quadruplex DNA is facilitated through the stacking of the phenoxazone chromophore to the terminal G-quartet of the quadruplex DNA. This stacked interaction between the ligand and DNA serves as an anchor, allowing the DNA strands to reorient with respect to the rigid cyclic pentapeptide side chains of the actinomycin D and results in an actinomycin D-quadruplex DNA complex whose quadruplex DNA structure has changed to a more energetically stable conformation to accommodate the actinomycin D sidechains.
REFERENCES


Table 1. Thermodynamic properties associated with the interactions of actinomycin D with the Na\(^+\) and K\(^+\) structural isoforms of quadruplex DNA.

<table>
<thead>
<tr>
<th></th>
<th>ITC</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(n)</td>
<td>(K_{\text{eq}})</td>
<td>(\Delta G^o)</td>
<td>(\Delta H^o)</td>
</tr>
<tr>
<td>(Na^+) Quadruplex</td>
<td>0.5 ± 0.03</td>
<td>2.1 ± 0.3 x 10(^5)</td>
<td>-7.3 ± 0.1</td>
<td>-7.3 ± 0.1</td>
<td>-0.07</td>
</tr>
<tr>
<td>(K^+) Quadruplex</td>
<td>0.6 ± 0.02</td>
<td>2.3 ± 0.1 x 10(^5)</td>
<td>-7.3 ± 0.2</td>
<td>-6.7 ± 0.1</td>
<td>0.60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>DSC</th>
<th>(Tm_1^a)</th>
<th>(\Delta H_1^b)</th>
<th>(\Delta H_{v1}^c)</th>
<th>(Tm_2^d)</th>
<th>(\Delta H_2^e)</th>
<th>(\Delta H_{v2}^f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Na^+) Quadruplex</td>
<td>39</td>
<td>1.3 x 10(^3)</td>
<td>3.9 x 10(^4)</td>
<td>60</td>
<td>2.4 x 10(^4)</td>
<td>5.2 x 10(^4)</td>
<td></td>
</tr>
<tr>
<td>(Na^+) Quadruplex + ActD</td>
<td>65</td>
<td>4.9 x 10(^3)</td>
<td>1.5 x 10(^4)</td>
<td>79</td>
<td>1.5 x 10(^4)</td>
<td>9.6 x 10(^4)</td>
<td></td>
</tr>
<tr>
<td>(K^+) Quadruplex</td>
<td>60</td>
<td>5.6 x 10(^3)</td>
<td>4.5 x 10(^4)</td>
<td>68</td>
<td>1.6 x 10(^4)</td>
<td>6.9 x 10(^4)</td>
<td></td>
</tr>
<tr>
<td>(K^+) Quadruplex + ActD</td>
<td>64</td>
<td>7.1 x 10(^3)</td>
<td>6.3 x 10(^4)</td>
<td>79</td>
<td>1.1 x 10(^4)</td>
<td>9.8 x 10(^4)</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)\(Tm_1\) is defined as the lower temperature transition; \(^b\)\(\Delta H_1\) is defined as the calorimetric enthalpy for the low temperature transition; \(^c\)\(\Delta H_{v1}\) is defined as the van’t Hoff enthalpy of low temperature transition; \(^d\)\(Tm_2\) is defined as the high temperature transition; \(^e\)\(\Delta H_1\) is defined as the calorimetric enthalpy for the high temperature transition; \(^f\)\(\Delta H_{v2}\) is defined as the van’t Hoff enthalpy of high temperature transition.
Figure 15. Representative CD spectra of the Na⁺ (black) and K⁺ (red) of structural isoforms of quadruplex DNA. CD spectra were recorded on an Aviv 400 CD spectrophotometer at 25°C in 1-cm pathlength cuvettes in 0.01 M (Na or K) phosphate buffer, pH 7.0 and 0.001 M EDTA. The deoxyoligonucleotide concentrations for both DNAs were 12.5 µM (strand). An averaging time of 4 sec at each nm was used over the range of 220 to 320 nm.
Figure 2. Changes to the CD spectra of the Na\textsuperscript{+} form (panel A) and the K\textsuperscript{+} form (panel B) of quadruplex DNAs are observed upon binding of actinomycin D. Experiments were performed at 25°C in 0.01 M (Na or K) phosphate buffer, pH 7.0 and 0.001 M EDTA. Half-mole equivalents of actinomycin D were titrated into a 1-cm CD cell containing 12.5 µM (strand) of quadruplex DNA. After each 37.5 µL injection of ligand, the drug-DNA solution was mixed and allowed to equilibrate for 10 minutes prior to measuring each spectrum. An averaging time of 4 sec at each nm was used over the range of 220 to 320 nm. Contributions from free actinomycin D and buffer were subtracted from each subsequent measurement. Scvitzky-Golay smoothing was applied to the raw spectral data. Figure 2C shows the CD spectra of the Na\textsuperscript{+} (black) and K\textsuperscript{+} (red) forms of G-quadruplex DNAs at the endpoints of titration (saturation) with actinomycin D.
Figure 3. The induced CD spectra of actinomycin D complexed with the Na\(^+\) (red) and K\(^+\) (blue) forms of quadruplex DNAs. These induced CD signals were obtained by mixing actinomycin D and quadruplex DNA at 25°C (in appropriate Na\(^+\) or K\(^+\) buffers as described in the legend of Figure 1) and under conditions ensuring that all of the actinomycin in solution is complexed to the DNA. This figure shows the CD spectrum of free actinomycin D (black), actinomycin D bound to the Na\(^+\) form of the quadruplex DNA (red), and actinomycin D bound to the K\(^+\) form of the quadruplex DNA (blue). Induced CD signals are observed at 380 nm (negative) and 440 nm (positive). An averaging time of 4 sec at each nm was used over the range of 340 to 500 nm.
Figure 4. Representative data from ITC experiments showing the titration of actinomycin D with the Na\(^+\) (panel A) and the K\(^+\) (panel B) forms of quadruplex DNAs. ITC experiments were performed at 25\(^\circ\)C (in appropriate Na\(^+\) or K\(^+\) buffers as described in the legend of Figure 1). All solutions were degassed using the Microcal thermovac 10 min prior to loading. Injections of 10 µL of stock actinomycin D solution (200 µM) into quadruplex DNA solution (10 µM strand) with 250 seconds resting time between injections. The top panels represent the raw heats of binding generated with each addition of actinomycin D and bottom panels are the integrated heats in kcal/mol of injectant. Thermodynamic parameters were derived from the single site binding model incorporated into the Microcal Origin 7.1 software.
Figure 5. DSC thermograms of the K$^+$ form of quadruplex DNA (panel A, no actinomycin D and panel B, with actinomycin D) and the Na$^+$ form of quadruplex DNA (panel C, no actinomycin D and panel D, plus actinomycin D). Experiments were performed in appropriate (Na or K) 0.01 M phosphate buffer, pH 7.0 and 0.001 M EDTA. Quadruplex DNA concentrations were 125 μM for both the Na$^+$ and K$^+$ forms of quadruplex DNAs. Actinomycin D concentrations were 250 μM. Heating temperatures ranged from 15 to 100 °C with a scan rate of 1.5 °C per min. Buffer-buffer baseline scans were subtracted from DNA and ligand-DNA scans. Progressive baseline subtraction procedures were used to provide the corrected DSC scans for data analyses. The two transition non-two state model (Microcal Origin 7.1) was used to analyze the resultant DSC scans.
CHAPTER IV

BIOPHYSICAL CHARACTERIZATION OF THE BINDING AND UNFOLDING OF HUMAN TELOMERIC G-QUADRUPLEX DNA BY UNFOLDING PROTEIN 1 (UP1)

by

Jason S. Hudson\textsuperscript{1}; Lei Ding\textsuperscript{2}; Jinbiao Ma\textsuperscript{2}; Edwin Lewis\textsuperscript{3} and David Graves\textsuperscript{1,4}

\textsuperscript{1}Department of Chemistry, University of Alabama at Birmingham, Birmingham, AL 35294
\textsuperscript{2}Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL 35294
\textsuperscript{3}Department of Chemistry, Mississippi State University, Mississippi State, MS 39762
\textsuperscript{4}Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL 35294

In preparation for submission to \textit{Biochemistry}

Format adapted for dissertation
Abstract

Recent studies have provided growing evidence that the formation of higher order nucleic acids structure(s) within the genome may regulate critical cellular processes such as genetic transcription of certain oncogenes and maintenance of the telomere at the ends of the chromosome. The specific recognition of G-quadruplex structures by proteins provides evidence for a functional role for the formation of G-quadruplex structure, in vivo. It has been previously reported that the nucleoprotein hnRNP A1 and its proteolytic derivative, UP1, bind to and destabilize G-quadruplex structures formed by the telomeric repeat d(TTAGGG)_n and may be involved in the recruitment of telomerase to the telomeres. In this study, we present a thermodynamic characterization of the binding and unfolding of unimolecular and bimolecular G-quadruplexes formed from the human telomeric repeat sequence, d[AGGG(TTAGGG)]_3. Isothermal titration calorimetry (ITC) studies have provided a detailed energetic profile of the UP1-G-quadruplex DNA interaction and provide insights into the unfolding mechanism of the G-quadruplex structure. This protein-DNA interaction was found to be enthalpically driven with large and negative binding enthalpies (ΔH = -36 and -43 kcal/mol) for the Na^+ forms of the bimolecular and unimolecular G-quadruplexes, respectively. In contrast, a less favorable binding enthalpy of ΔH = -19 kcal/mol for the K^+ form of the unimolecular G-quadruplex was observed. The ITC studies revealed three distinct enthalpic contributions to the interaction of UP1 with the Na^+ forms of G-quadruplex DNA. The initial interaction is characterized by a binding affinity of 8.5 x 10^8 M^{-1}(strand); markedly higher than that observed for the binding of UP1 to single stranded DNA (4.1 x 10^6 M^{-1}(strand)) for a comparable but non-quadruplex forming sequence. Circular dichroism (CD)
spectroscopy reveals that both the unimolecular and bimolecular Na$^+$ forms of G-quadruplex DNAs could be efficiently unfolded by UP1. Our ITC titration experiments have provided the basis for an energetic explanation for the apparent differences in the unfolding capability of the Na$^+$ and K$^+$ G-quadruplexes by the UP1 protein. The results presented here demonstrate that UP1 binds with high affinity to G-quadruplex DNA structural motif and that the binding energetics are closely coupled to the unfolding of the G-quadruplex structure.
**Introduction**

The human genome contains highly conserved repetitive G-rich sequences that are located in strategic regions of chromosomes and are important for biological functions such as transcriptional promoter regions and at the telomeric ends of eukaryotic chromosomes (1, 2). The human telomeric DNA is a single stranded region located at the terminal ends of chromosomes with a 3’-end overhang containing tandem repeats of d(TTAGGG)_n. This non-coding overhang is composed of ~150-200 nucleotides and protects genomic DNA from end fusion and for maintenance of chromosomal integrity during replication (3). There is evidence to suggest that telomeres serve as a biological clock, determining the lifespan of the cell and the loss of telomeric DNA upon replication ultimately leads to apoptosis when the telomeres become critically shortened (4-6).

Telomeric DNA has been implicated in cancer progression due to the up-regulation of telomerase activity and subsequent lengthening of telomeres leading to cellular immortality as observed in malignant cells (7-9). The nucleotide sequences that compose telomeric DNA have been shown to readily form G-quadruplex structural motifs in vitro and provide novel targets for the development of new classes of anticancer agents (10-12). The formation of higher order DNA structures within telomeric sequences lend confidence to the idea that these structures serve regulatory roles for telomere extension and maintenance (13-15).

Several proteins and enzymes such as TRF2, POT1, and telomerase have been shown to associate with telomeric DNA and exert a number of biological functions (16). Telomerase is an enzyme composed of a reverse transcriptase and telomeric RNA transcript whose function serves to elongate the repetitive sequence at the telomeric ends.
of the chromosomes (17). In mature somatic cells, telomerase is down-regulated and inactive, however it is well documented that in many cancer cells telomerase activity is up-regulated, resulting in the lengthening of the repetitive sequences leading to the characteristic immortalization of the cell during cancer progression. Telomerase requires that the telomeric region be in the single stranded conformation for binding and elongation to occur. The presence of the folded G-quadruplex structural motif inhibits this interaction (18). The discovery of proteins (19) and helicases (20) that specifically recognize and destabilize the G-quadruplex structural motif raises strong implications for the dynamic nature and function of G-quadruplex structures within the genome. In contrast to helicases, destabilizing proteins unfold G-quadruplex DNAs in a non-enzymatic manner and do not require ATP hydrolysis for activity. Among these, hnRNP A1 is a member of a class of nucleoproteins whose function has been reported to be involved in RNA transport and alternative splicing and is closely associated with DNA polymerase transcripts (21, 22). The hnRNP A1 contains two nucleic acid binding domains that strongly interacts with both RNA and DNA sequences. Unwinding protein 1 (UP1), is a 196 amino acid proteolytic product of hnRNP A1 that retains the two nucleic acid binding domains. UP1 has been shown to bind and destabilize G-quadruplex structures and potentially serve as a DNA chaperone that is responsible for the unfolding of the G-quadruplex structure into single stranded DNA to facilitate the binding of telomerase for lengthening of the telomere (23). Further studies have demonstrated that increased hnRNP A1 expression and/or introduction of exogenous UP1 results in elongation of the telomeres (24). Fakuda, et al. reported UP1 to bind to and destabilize G-quadruplex structures in mouse minisatellite repeats as well as human telomeric DNA
This work was further supported by Shamoo and coworkers in 2003. The structure of UP1 complexed with single-stranded DNA was reported by Ding and coworkers in 1999 and provides a detailed image of the UP1-DNA complex wherein short oligonucleotides (12-mers) consistent in sequence to telomere regions were complexed in a single stranded antiparallel conformation with two UP1 proteins. The two nucleic acid binding domains, BD1 and BD2, are composed of residues that participate in the formation of non-covalent interactions including hydrogen bonding, stacking, van der Waals, and electrostatic interactions to approximately four nucleotides d(TAGG) in the telomeric DNA sequence. It is clear that UP1 actively functions as a G-quadruplex destabilizing protein; however the nature through which UP1 recognizes the G-quadruplex structure and the unfolding mechanism(s) remain unknown. The work presented here provides a detailed biophysical examination of the binding of UP1 with the human telomeric (Tel22) G-quadruplex structural motif (Na$^+$ and K$^+$ forms). Similarly, we examine the interactions of UP1 with the Na$^+$ form of the bimolecular (Tel12) G-quadruplex that is formed by the repeat sequence of human telomeric DNA. The studies presented here examine the energetics associated with complex formation between UP1 and human telomeric G-quadruplex DNA by isothermal titration calorimetry (ITC) and follow the unfolding process upon complex formation by circular dichroism (CD) spectroscopy. Our results provide a direct method for binding equilibrium measurements that results in an energetic characterization of the interaction and provide key insights into the recognition, affinity, and destabilization of G-quadruplex structures by UP1.
Materials and Methods

Oligonucleotides and Sample Preparation

The 22-nucleotide human telomeric sequence 5'-AGGGTTAGGGTTAGGGTTAGG-3' (Tel-22), the 22-mer non-G-quadruplex forming control sequence (where the third G is mutated to an A) 5'-AGGGTTAGGGTTAGGGTTAGG-3' (Tel-22ss), and the G-quadruplex forming 12-mer 5'-TTAGGGTTAGGG-3' (Tel-12) were obtained from Midland Certified Reagents, Midland, TX. DNA samples were prepared by dissolving in Tris buffer (0.020M Tris-HCl, pH 8.0, 0.1M NaCl), followed by an annealing process wherein the DNA solution was heated to 85°C and slowly cooled to 4°C at a rate of 1.0°C/min using an MJ Research thermocycler. Stock concentrations of DNAs were determined by UV absorbance at 85°C. Molar extinction coefficients for the sequences were Tel-22 (ε_{260} = 228,500 M⁻¹ cm⁻¹), Tel-22ss (ε_{260} = 240,900 M⁻¹ cm⁻¹) and Tel-12 (ε_{260} = 122,400 M⁻¹ cm⁻¹) as determined by the nearest neighbor method (28, 29). The formation of the G-quadruplex structural motif and topologies were evaluated for both sequences at a concentration of 5 µM (quadruplex) in 1-cm pathlength cells using an Aviv 400 circular dichroism spectropolarimeter. Wavelength scans from 325-225 nm were examined for positive ellipticity signal at 295 nm, which is the characteristic signature for antiparallel G-quadruplex DNA structure (30).

Preparation of UP1

The cDNA of the UP1 (1-196aa of hnRNP A1) was cloned into the vector pET28-SMT3 to result in an N-terminal SUMO tag and 6x His tag and expressed in Escherichia
coli BL21-Gold(DE3). Cells were grown in LB media for 3 hours at 37°C until optical density of 0.6 at 600nm was achieved. Induction using 0.1mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) was performed followed by overnight incubation at 18°C prior to harvesting the cells. The cells were washed with phosphate buffer (PBS) and resuspended in 5mL of binding buffer (20 mM Tris pH 8.0, 0.5 M NaCl, 25 mM imidazole) per 100mL of culture. The cells were disrupted by Emulsiflex-C3 cell homogenizer (Avestin Inc.) and followed by centrifugation for 40min at 16,000 rpm at 4°C. The supernatant was loaded on 5mL HisTrap column (GE-Health) and then eluted with buffer (20 mM Tris pH 8.0, 0.5M NaCl, 500 mM imidazole). The eluted protein was digested with SUMO protease and dialyzed against 20 mM Tris pH 8.0, 0.5 M NaCl overnight, followed by treatment with a second pass through the HisTrap column. The protein was further purified by passing over a Heparin column to exclude proteins that complex with nucleic acids. The protein was then fractionated by gel filtration using Superdex 75 16/60 column (GE-Health). The concentration of protein was determined by the absorbance at 280 nm by UV-visible spectroscopy (Varian).

Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were performed using a Microcal VP-ITC (GE Healthcare, Piscataway, NJ) at 25°C. All samples were thoroughly degassed with stirring prior to use. For the forward titrations, the sample cell was filled to capacity with the G-quadruplex DNA and titrated with UP1 protein and 400 seconds resting time between each injection. Replicate titrations were performed and the resulting data were corrected for heat of dilution effects. The resulting binding isotherms
were non-sigmoidal and could not be fit with the binding models incorporated into the Microcal Origin VP-ITC software to calculate the energetics of the interaction. It was apparent from the shape of the isotherm that multiple events were present in the reaction and either a two-event or three-event binding model was necessary to achieve a suitable fit for the binding isotherms. The fitting algorithms were developed by combination of the mass balance and equilibrium constant expressions expanded to n-independent site binding expressions. Lewis and co-workers previously reported the details regarding the development of the fitting algorithm to describe complex ITC binding isotherms and is the basis for our fitting procedure (31).

When multiple events that have enthalpic contributions to the binding isotherm are present, it is useful to reverse the direction of the titration. We conducted reverse titrations by placing UP1 in the sample cell and titrating with the G-quadruplex DNA. The binding isotherms were sigmoidal and could be fit using the single-site binding model included in the Origin 7.1 software. It should be noted that the resulting thermodynamic parameters of the reverse titration fits are the average of all enthalpic events that occur during the reaction. However, these titrations are useful to reduce the uncertainties and confirm the binding stoichiometry.

ITC titrations for the control 22-mer sequence (does not form the G-quadruplex structure) were conducted in both the forward and reverse direction. The binding isotherms were observed to be sigmoidal in both directions and were fit with the single-site binding model used to fit the reverse titrations of the G-quadruplex DNA.
CD Titrations

Examination of changes in structural features of the human telomere G-quadruplex resulting from UP1 binding were performed using circular dichroism spectropolarimetry (CD). CD titrations were conducted using an AVIV 400 circular dichroism spectropolarimeter (Aviv, Inc.). Experiments were performed in 1 cm path length rectangular cells at 25° C. Wavelength scans were collected every 1 nm from 225 to 325 nm at with a 2 second averaging time. Initial concentrations of the 12-mer and 22-mer quadruplex DNAs were 5 µM (strand) and titrated with the UP1 protein. After each injection the sample was allowed to equilibrate for 10 minutes with constant stirring. Contributions of the buffer were subtracted from each wavelength scan and the data normalized to molar ellipticity (deg cm²/mol) to account for dilution effects of the initial DNA concentration in the cell.

Results

Characterization of the G-quadruplex Structure

Prior to investigating the UP1 interaction with G-quadruplex DNA, it was necessary to structurally characterize the deoxyoligonucleotide sequences. Circular dichroism (CD) is a useful tool to determine the presence of the G-quadruplex structural motif (32). G-quadruplexes exhibit characteristic positive ellipticity at wavelengths between 280 and 300 nm, primarily due to the stacking interactions of the guanine bases in the G-tetrads. The folding topology and strand orientations (parallel/antiparallel) of the G-quadruplex structure can be determined by the positive ellipticity between 280 and 300 nm. A G-quadruplex structure that has complete parallel strand orientations will
have a positive ellipticity at 280 nm, whereas fully antiparallel strand orientations will have a positive ellipticity at 295 nm. In K\(^+\) buffer conditions, a mixture of orientations is observed with positive maxima at both 280 and 295 nm.

The resulting CD spectra of the sequences selected for this study and cartoons depicting possible strand orientations are shown in Figure 1. The DNA sequences prepared in Na\(^+\) buffer are expected to form bimolecular and unimolecular antiparallel G-quadruplex structure (Figures 1A and 1B) and displays characteristic CD spectra, Figures 1D and 1E. In contrast, the mixed parallel/antiparallel conformation for the 22-mer sequence prepared in K\(^+\) buffer is shown in Figure 1C and 1F. The bimolecular Tel-12 (Na\(^+\)) and unimolecular Tel-22 (Na\(^+\)) shown in Figure 1A and 1B are characterized by ellipticities with maxima at 295 and 245 nm and a minimum at 265 nm, suggesting that both of the G-quadruplex structures are in antiparallel conformations. The CD spectrum for the unimolecular Tel-22 (K\(^+\)) shown in Figure 1F is consistent with a mixture of parallel and antiparallel contributions, with a positive ellipticity at 295 nm and a shoulder peak of positive ellipticity at 265 nm. There are multiple conformations that could give rise to the strand orientation as illustrated in Figure 1 panel A, B, and C. Both structures may be present and there are no determining factors to distinguish them based on CD measurements alone.

The control deoxyoligonucleotide (Tel-22ss) is unable to fold into the G-quadruplex structure. The third guanine in each of the four repeats was mutated to adenine in order to prevent the formation of the G-tetrad and the formation of G-quadruplex structure. The CD spectrum was analyzed and revealed an unstructured DNA
with no positive ellipticity at 280 or 295 nm wavelengths of the spectrum (data not shown).

In addition to structural characterization of the G-quadruplex forming sequences, the thermodynamic stabilities were examined by CD melting studies. The molar ellipticities for each of the G-quadruplexes were monitored as a function of temperature and the resulting changes in the CD signal at 295 nm revealed sigmoidal melting curves and the melting temperatures, $T_m$, were derived from the midpoint of these transitions and are shown in Figure 2. The Na$^+$ form of the 22-mer antiparallel unimolecular G-quadruplex exhibits a $T_m$ of 59° C in 100 mM NaCl, 0.01 M Tris buffer at pH 8. In contrast, the Na$^+$ form of the 12-mer bimolecular antiparallel G-quadruplex is less stable, exhibiting a modest $T_m$ of 39° C in 100 mM NaCl, 0.01 M Tris buffer at pH 8. The Tel-22 (K$^+$) G-quadruplex was markedly more stable than the both Na$^+$ forms. The Tel-22 (K$^+$) G-quadruplex exhibited a $T_m$ of 68° C in 100 mM KCl buffer, 0.01 M Tris at pH 8.0; approximately 10° C higher than the unimolecular Na$^+$ counterpart.

The Binding and Unfolding of Telomeric G-Quadruplex DNA by UP1

Previous studies have reported that UP1 binds and unfolds G-quadruplex structures (25, 26, 33); however, these studies provide limited information regarding the energetics associated with the interaction beyond an estimated equilibrium constant. The studies presented here examine the fundamental thermodynamic properties associated with the binding and unfolding of telomeric G-quadruplex DNA by directly measuring the binding enthalpy ($\Delta H_{\text{bind}}$) upon complex formation. We have utilized isothermal titration calorimetry (ITC) to investigate the enthalpic properties of these interactions and
derive relevant thermodynamic parameters that allow us to more fully understand the
binding mechanism(s) associated with complex formation and unfolding.

ITC titrations were conducted by placing the G-quadruplex DNA in the sample
cell and UP1 in the injection syringe. The results of the binding of UP1 to the
unimolecular Tel-22 (Na⁺) and Tel-12 (Na⁺) G-quadruplexes are shown in Figure 3A and
3B and summarized in Table I. Figures 3A and 3B depict the raw ITC data and insets
represent the integration of the measured heats associated with complex formation for
each injection. The bottom panels (3C and 3D) are the CD spectra for the titrations of
UP1 into the DNA solution at specific molar ratios that coincide with the ITC titration.
The binding isotherm shown in Figure 3A reveals a complex binding process and is
indicative of a multi-event process. The solid line drawn through the data points
represents the best fit to the data by a three event model and results in three distinct
enthalpic components of the interaction. At low UP1 concentrations (initial injections),
K₁ is estimated to be \(8.5 \times 10^8\) M\(^{-1}\) (strand). Two additional equilibrium constants, K₂
and K₃ are determined to be \(8.4 \times 10^7\) M\(^{-1}\) (strand) and \(6.4 \times 10^6\) M\(^{-1}\) (strand),
respectively. Upon achieving a 2:1 molar ratio (UP1: G-quadruplex), the binding
isotherm shapes becomes sigmoidal. The overall stoichiometry for binding of UP1 to the
Tel-22 Na⁺ G-quadruplex was determined to be two proteins per one G-quadruplex DNA.
Binding enthalpies for the three events observed for the unimolecular Na⁺ G-quadruplex
were determined to be \(ΔH_1, ΔH_2, \) and \(ΔH_3\) of -45.4, -7.3, and -29.6 kcal/mol, respectively,
indicative of a highly enthalpically favored binding for event 1, markedly reduced
favorable enthalpy for event 2, and favored enthalpically driven binding to event 3.
In an attempt to interpret the complexity of the binding isotherm we conducted CD spectropolarimetry experiments that monitored the structural characteristics of the G-quadruplex by observing the change molar ellipticity at 295 nm of the Tel-22 Na⁺ G-quadruplex as a function of UP1 are shown in Figure 3C. The decrease in the ellipticity at 295 nm is indicative of a UP1-mediated unfolding of the G-quadruplex. A complete unfolding of the quadruplex is achieved by a molar ratio of 2:1 (UP1 to G-quadruplex). Based on these results, we conclude that the third binding event observed in the ITC binding isotherm must correspond to the binding enthalpy for the interaction of UP1 with the unfolded G-quadruplex sequence. Therefore, the enthalpic contributions to the first two events must be a combination of the initial binding interaction (first observed event) and subsequent unfolding of the G-quadruplex structure (second observed event). The G-quadruplex structure is unfolded at ratios above to and the sigmoidal portion of the binding isotherm must be the binding of the second UP1 molecule to the unfolded G-quadruplex sequence.

The binding of UP1 to the bimolecular (Tel-12 Na⁺) G-quadruplex is shown in Figure 3B, with the inset representing the integration of the measured heats of complex formation for each injection. This binding is typified by a well resolved sigmoidal binding isotherm that does not reveal multiple enthalpic components. The data could be fit with a single site model resulting in a binding affinity of $3.0 \times 10^7 \text{ M}^{-1}$ and a stoichiometry of 2.0 (UP1: G-quadruplex) as summarized in Table I. The binding enthalpy ($\Delta H$) for the UP1-bimolecular G-quadruplex interaction is -36 kcal/mol and indicative of a very favorable enthalpy driven process. The concomitant CD spectra for this titration is shown in Figure 3D and reveals the unfolding of the G-quadruplex
structure to occur at markedly lower UP1 concentrations than needed for the Tel-22 Na\(^+\) G-quadruplex.

The binding of UP1 to the more stable Tel-22 K\(^+\) G-quadruplex results in a binding isotherm with different features from either of the Na\(^+\) G-quadruplexes, as shown in Figure 4A. Integration of the raw injection peaks results in a near hyperbolic binding isotherm (inset in Figure 4A) that is characterized by only 2 events. The unfolding of the K\(^+\) G-quadruplex by CD experiments is shown in Figure 4B and reveals that even at a molar ratio of 4.0 (UP1:G-quadruplex), the peak at 295 nm is still present, indicating that G-quadruplex structure is not fully unfolded by UP1. We were able to fit the hyperbolic ITC binding isotherms by a two event binding model, in contrast the Tel-22 Na\(^+\) G-quadruplex that required three events for a suitable fit. The K\(_1\) and K\(_2\) were estimated to be 2.8 x 10\(^7\) M\(^{-1}\) and 1.3 x 10\(^6\) M\(^{-1}\), respectively. The enthalpy of the first event was significantly reduced for the K\(^+\) G-quadruplex with a less favorable value of \(\Delta H_1 = -23.4\) kcal/mol. There was not a significant deviation in the second enthalpy, \(\Delta H_2 = -7.8\) kcal/mol, and the binding stoichiometry remained at two UP1 molecules per one G-quadruplex.

The control deoxyoligonucleotide (Tel-22-ss) was designed by replacing the d(TTAGGG) repeat with d(TTAGGA) resulting in a deoxyoligonucleotide that is comparable in sequence but cannot fold into a G-quadruplex structure under Na\(^+\) or K\(^+\) buffer conditions. Binding of UP1 to this single-stranded deoxyoligonucleotide is shown in Figure 5 and shows that UP1 binds to the single-strand DNA in a single event with a reduced binding affinity of 4.1 x 10\(^6\) M\(^{-1}\) and a stoichiometry of two UP1 molecules per one strand of DNA. Interestingly, the binding enthalpy for the interaction of UP1 to the
single-stranded DNA was comparable to that of the Na$^+$ G-quadruplex forming sequence with a $\Delta H = -37$ kcal/mol.

Under conditions where multiple binding sites or multiple equilibria exist, unusual binding isotherms shapes are often encountered in the forward titrations. In such cases, the ITC experiment can be designed so that the ligand and target molecules are reversed and in this case, DNA in the injection syringe is titrated into UP1 in the sample cell, deemed a “reverse” ITC titration (38). For simple binding interactions, it is expected that the shape of the binding isotherm should remain consistent after reversing the direction of the titration. However, in more complex cases as observed in Figure 3, a binding isotherm consisting of multiple binding events may be simplified into a sigmoidal binding isotherm that represents the approximate average of all events as shown in Figure 6. In Figure 6A, the Tel-22 Na$^+$ G-quadruplex DNA is titrated into UP1. Having a large excess UP1 in the sample cell results in driving the multi-event reaction to completion, and the binding and unfolding of all G-quadruplex DNA that is added for each injection. This occurs until the UP1 concentration becomes the limiting component and fails to satisfy the 2:1 binding stoichiometry. At this point small endothermic peaks become present and we observe the energetics of the G-quadruplex unfolding. The binding constant, $K_{eq}$ of $3.3 \times 10^7$ M$^{-1}$ (strand) is estimated by fitting of a single-site model to the data shown in Figure 6A that represents the reverse titration of Tel-22 Na$^+$. This is an approximate value for an average of the association constants derived from the forward ITC titration. Similarly, the $\Delta H$ of binding (-35.8 kcal/mol) for the reverse ITC titration is in good agreement with the approximate average of each of the individual components estimated from the forward ITC titration.
Figure 6B represents the reverse titration experiment for the bimolecular Tel-12 Na\(^+\) G-quadruplex. The binding of UP1 to the Tel-12 (Na\(^+\)) had similar results to those derived from the forward titration and retained the sigmoidal characteristics. The equilibrium constant was 4.1 \(\times\) 10\(^{-7}\) M\(^{-1}\) and the \(\Delta H\) of binding was -38.6 kcal/mol. The K\(^+\) G-quadruplex that appeared hyperbolic in the forward direction was performed in the reverse direction and resulted in a sigmoidal isotherm that was fit with the single-site model used for the Tel-12 Na\(^+\) and Tel-22 Na\(^+\) sequences (data not shown). The binding affinity was comparable to the other G-quadruplex sequences (K\(_{eq}\) = 2.6 \(\times\) 10\(^{-7}\) M\(^{-1}\)) but the binding enthalpy was significantly less favorable (\(\Delta H\) = -18.9 kcal/mol). The results for the reverse titrations are provided in Table 2.

**Discussion**

In the work presented here we have reported the energetics of binding for the unfolding protein UP1 interacting with a bimolecular G-quadruplex (Tel-12-Na\(^+\)), two conformations of a unimolecular G-quadruplex (Tel-22 Na\(^+\) and Tel-22 K\(^+\)), and a single stranded nucleic acid sequence (Tel-22ss). Previously published studies have demonstrated that the UP1 binds to d(TTAGGG)_n sequences that form G-quadruplex structures with high affinity and destabilizes the G-quadruplex structure. Fakuda, *et al.* reported that UP1 could bind and unfold G-quadruplex structures formed from sequence d(GGCAG)_n and for those of the human telomeric sequence d(TTAGGG)_n (25). Shamoo, and coworkers further examined the basis for purine recognition by the binding domains of UP1 and found that the substitution of 2-aminopurine, nebularine, or 7-deazaquanine for the first guanine residue in the sequence d(TAGGG)_n was poorly tolerated and greatly
reduced the binding affinity (34). These results, in conjunction with the crystal structure provided by Ding, et al., led to the proposal that the consensus binding sequence for UP1 to nucleic acids is d(nYAGn), where Y is either a thymine or cytosine residue, and is stabilized by hydrogen bonding, base stacking, and hydrophobic effects.

Currently there is limited information regarding the energetic forces that are present in the binding and unfolding of the telomeric G-quadruplex by UP1. The purpose of this manuscript is to provide the thermodynamic aspects of the G-quadruplex and UP1 interaction. The sequences and G-quadruplex structures used for our study all contain the d(TAG) binding sites but differ in their molecularity and folding topologies. We characterized the G-quadruplex structures and stabilities by CD spectroscopy and determined that the bimolecular 12 nucleotide sequence and both unimolecular 22 nucleotide sequences formed G-quadruplex structures and had consistent thermal stabilities and folding topologies observed by other research groups (35,36). The stability was influenced by the molecularity and buffer solution cation resulting in increasing melting temperatures of Tel-12-Na⁺ < Tel-22-Na⁺ < Tel-22-K⁺.

Forward titrations of the UP1-G-quadruplex interaction were performed where small aliquots of protein were injected into a large excess of nucleic acid. The ITC binding isotherms for these experiments were composed of multiple enthalpic contributions and revealed a complex binding isotherm for the unimolecular G-quadruplexes. The “hook” shape that was observed for the Tel-22-Na⁺ sequence at molar ratios less than 2.0 (UP1 to G-quadruplex) suggested the occurrence of two distinct events composed of enthalpic contributions to the heat of the reaction that are of opposite sign (exothermic and endothermic). The observation of endothermic peaks (+ΔH) in the
early stages of the titration for Tel-22-Na\(^+\) (peaks 4-9) suggests that there is an entropically favored component related to either the binding and/or unfolding of quadruplex structure by UP1. The complex shape of the binding isotherm that we observe is a result of the measurement being the sum of the two enthalpies (+\(\Delta H\) and – \(\Delta H\)) for the binding of UP1 and unfolding of the G-quadruplex. From our CD spectropolarimetry experiments we are aware that the unfolding of the G-quadruplex reaches completion when a molar ratio of 2:1 UP1 to G-quadruplex is achieved. However, the isotherm revealed an additional binding event still present that is characterized by a sigmoidal shape. Our results indicate that the G-quadruplex structure is unfolded at this ratio and that the third event is the binding of a second UP1 molecule to the unstructured DNA. Based on this observation we have denoted the three events that are apparent in the interaction to be (1) the initial binding interaction of UP1 and the G-quadruplex, (2) the contribution of energetics of unfolding the G-quadruplex structure, and (3) the binding of the second UP1 molecule to the unfolded G-quadruplex sequence. The observed binding affinity for the first event was estimated to be a very high affinity interaction (\(K_1 = 8.5 \times 10^8 \text{ M}^{-1}\)) and the third event was an order of magnitude lower in affinity than the first (\(K_3 = 6.4 \times 10^6 \text{ M}^{-1}\)), suggesting a preference for binding the G-quadruplex structure as opposed to the single stranded conformation. The binding enthalpy was also more favorable for the first event in comparison to the second, (\(\Delta H_1 = -45.2 \text{ kcal/mol and } \Delta H_3 = -29.6 \text{ kcal/mol, respectively}\)).

The forward titration for the Tel-12 Na\(^+\) G-quadruplex did not have any apparent complexity and resulted in a sigmoidal binding isotherm. The binding affinity was 3.0 x \(10^7 \text{ M}^{-1}\) and the binding enthalpy was favorable, \(\Delta H_{\text{bind}} = -36 \text{ kcal/mol}\). The lack of any
complexity may be due to the lower stability of the Tel-12 Na\(^+\) G-quadruplex. Our melting studies showed that the Na\(^+\) conformations of the Tel-12 and Tel-22 G-quadruplexes differed by almost 20° C. The lower thermal stability of Tel-12 would indicate a lower energy barrier and a lower observed enthalpic contribution of unfolding the bimolecular structure. Under these circumstances, the isotherm would be dominated by the very large and exothermic enthalpy of binding and the endothermic contribution of unfolding would not be apparent.

The isotherm observed for Tel-22-K\(^+\) was relatively hyperbolic in shape and required only two events to fit the shape of the isotherm. No endotherms were observed in the raw data, but after inspection of the integrated data it was noted that there was a deviation in the hyperbolic shape that appeared to be the development of a “hook” at approximately the same molar ratio as seen for Tel-22-Na\(^+\) G-quadruplex. The CD spectropolarimetry studies revealed that Tel-22-K\(^+\) never achieved complete unfolding of the G-quadruplex even up to molar ratio of four UP1 molecules per G-quadruplex. The binding affinity for the first event was lower than that observed for either of the Na\(^+\) G-quadruplexes (K\(_1\) = 2.8 x 10\(^7\) M\(^{-1}\)) and complex formation was characterized by a reduction in binding enthalpy (ΔH\(_1\) = -23.4 kcal/mol). The K\(^+\) G-quadruplex is a very stable structure having a melting temperature of 68 °C and the binding energetics may not be sufficient to destabilize the G-quadruplex upon complex formation. In addition, the K\(^+\) conformation has different structural properties than the antiparallel Na\(^+\) G-quadruplexes. The antiparallel structures of the Na\(^+\) conformation have either lateral or diagonal loops in contrast to the K\(^+\) structure that has a mixed strand polarity of both parallel and antiparallel characteristics. The third loop of the K\(^+\) G-quadruplex is an
edgewise loop and located on the sides of the G-tetrad core, whereas lateral and diagonal loops are above or below the G-tetrads. This may alter recognition of the G-quadruplex by UP1 and reduce the binding energetics and in turn, the unfolding capability. This idea is supported by a recently published report that makes a similar observation for the binding and unfolding of the K+ conformation (37). The authors examined the bound conformation of the potassium form of the G-quadruplex with a single molecule fluorescence technique and observed that the oligonucleotide was still in a compacted state after complex formation and not fully unfolded.

The control sequence that was mutated to promote a single stranded conformation was utilized to investigate if there was a preference for binding to the G-quadruplex structure. We estimate a binding constant that was significantly lower (4.1 x 10^6 M^-1) than that observed for any of the G-quadruplex forming sequences. The binding enthalpy was still very favorable (ΔH_{bind} = -37 kcal/mol) and the binding isotherm was sigmoidal with no evidence of unfolding energetics. The recognition of the d(TTAGGG) sequence by UP1 was proposed by Shamoo, et al. (34), and it was noted that base composition, stereochemistry, and steric properties of the bases that make up the 5\'-TAG-3\' nucleic acid sequence are critical for UP1 binding and even slight modifications are detrimental to the binding affinity. The NMR solution structure of the G-quadruplex DNA reported by Wang and Patel (28) has adenine and guanine stacking interactions between the loop sequence and the next guanine in the sequence that forms one the four guanines in the G-tetrad. The steric properties of the bases in the single stranded conformation may be such that strong interactions of UP1 with the d(nTAGn) sequence are not formed and reduces
the binding affinity and binding enthalpy for the control single stranded sequence used in our study.

The work we have presented here represents the first examination of the energetic characteristics of the binding and unfolding of G-quadruplex DNA by the protein UP1. Our results provide insight into the enthalpic contributions to the binding and unfolding interaction. We have found that the binding and unfolding of G-quadruplex structure by UP1 is influenced by both the structure and stability of the oligonucleotide. The unfolding process is non-catalytic and the energy requirement for unfolding is directly coupled to the binding energetics. The recognition aspects and the exact mechanism by which UP1 facilitates quadruplex unfolding remain speculative and difficult to interpret based on these data alone. Additional research is currently underway investigating the effect of loop sequence on the binding energetics and unfolding mechanism.
References


Table 1

Thermodynamic parameters derived from the non-linear least squares fit of the isothermal titration calorimetry binding studies in the forward direction.  

<table>
<thead>
<tr>
<th></th>
<th>$n_{tot}$</th>
<th>Events</th>
<th>$n$</th>
<th>$K_{eq}$ (M$^{-1}$)</th>
<th>$\Delta G^o$ (kcal/mol)$^a$</th>
<th>$\Delta H^o$ (kcal/mol)</th>
<th>-$T\Delta S^o$ (cal/mol K)$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tel-12 (Na$^+$)</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3.0 X 10$^7$ ± 1.1</td>
<td>-10.2 ± 0.1</td>
<td>-36.0 ± 1.4</td>
<td>28.2 ± 1.8</td>
</tr>
<tr>
<td>Tel-22 (Na$^+$)</td>
<td>2</td>
<td>3</td>
<td>0.2</td>
<td>8.5 X 10$^3$ ± 1.9</td>
<td>-12.1 ± 0.2</td>
<td>-45.4 ± 2.1</td>
<td>33.2 ± 2.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.8</td>
<td>8.4 X 10$^3$ ± 3.7</td>
<td>-10.6 ± 0.2</td>
<td>-7.30 ± 1.9</td>
<td>-3.50 ± 1.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.2</td>
<td>6.4 X 10$^5$ ± 0.2</td>
<td>-9.27 ± 0.1</td>
<td>-29.6 ± 0.9</td>
<td>20.3 ± 0.9</td>
</tr>
<tr>
<td>Tel-22 (K$^+$)</td>
<td>2</td>
<td>2</td>
<td>0.2</td>
<td>2.8 X 10$^3$ ± 2.2</td>
<td>-10.4 ± 0.5</td>
<td>-23.4 ± 3.0</td>
<td>13.4 ± 2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.8</td>
<td>1.3 X 10$^5$ ± 0.6</td>
<td>-8.50 ± 0.4</td>
<td>-7.80 ± 1.0</td>
<td>-0.6 ± 1.3</td>
</tr>
<tr>
<td>Tel-22 (ss)</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4.1 X 10$^6$ ± 1.9</td>
<td>-9.01 ± 0.1</td>
<td>-37.0 ± 1.2</td>
<td>98.2 ± 1.5</td>
</tr>
</tbody>
</table>

$^a$ Gibb’s free energy calculated from the relation $\Delta G^o = -RT\ln K$.  
$^b$ Entropy ($-T\Delta S^o$) was calculated from the rearrangement of $\Delta G^o = \Delta H^o - T\Delta S^o$ to $-T\Delta S^o = (\Delta G^o - \Delta H^o)$. 
Table 2

Thermodynamic parameters derived from the best fit line of the isothermal titration binding studies in the reverse direction. \(^a\)The Gibb’s free energy was calculated from the relation \(\Delta G^\circ = -RT\ln K\). \(^b\)The entropy (\(-T\Delta S^\circ\)) was calculated from the rearrangement of \(\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ\) to \(-T\Delta S^\circ = (\Delta G^\circ - \Delta H^\circ)\).

| Sequence | \(n\) | \(K_a\) (mol\(^{-1}\)) | \(\Delta G^\circ\) (kcal/mol)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tel-12 (Na(^+))</td>
<td>0.5 ± 0.01</td>
<td>4.1 (\times 10^7) ± 0.2</td>
<td>-10.5 ± 0.2</td>
</tr>
<tr>
<td>Tel-22 (Na(^+))</td>
<td>0.5 ± 0.01</td>
<td>3.3 (\times 10^7) ± 0.8</td>
<td>-10.2 ± 0.1</td>
</tr>
<tr>
<td>Tel-22 (K(^+))</td>
<td>0.5 ± 0.02</td>
<td>2.6 (\times 10^7) ± 0.8</td>
<td>-10.1 ± 0.1</td>
</tr>
</tbody>
</table>
Figure 1: Panel A: Antiparallel structures of the G-quadruplex of the 12 nucleotide sequence of the human telomere. Panel D: CD spectrum of Tel-12 (Na⁺) in 100 mM NaCl Tris Buffer at pH 8.0. Panel B: Antiparallel conformations of the G-quadruplex formed by the 22 nucleotide sequence of the human telomere in sodium buffer. The chair structure (left) is composed of 3 lateral TTA loops and the basket structure (right) is composed of two lateral and one diagonal TTA loop. Panel E: CD spectrum of Tel-22 (Na⁺) in 100 mM NaCl Tris buffer at pH 8.0. Panel C: Conformations of the G-quadruplex formed by the 22 nucleotide human telomeric sequence in potassium buffer. Hybrid-1 or Hybrid 2 structures are composed of two lateral and one edgewise TTA loop. Panel F: CD spectrum of Tel-22 (K⁺) in 100 mM NaCl Tris buffer at pH 8.0.
Figure 2: Circular dichroism spectroscopy (CD) melts of the Tel-12 (Na\(^+\)) (squares), Tel-22 (Na\(^+\)) (circles), and Tel-22 (K\(^+\)) (triangles). Melting studies were carried out in 0.01 M Tris-HCl, pH 8.0, 0.001 M EDTA, and either 0.1 M NaCl or 0.1 M KCl. Changes in ellipticity were monitored at 295 nm over a temperature range of 10-90 °C. The melting temperatures (Tm) for each of the G-quadruplexes were calculated as the first-derivatives of the transition curves and found to be 39 °C for Tel-12 (Na\(^+\)), 59 °C for Tel-22 (Na\(^+\)), and 68° C for Tel-12 (K\(^+\)).
Figure 3. Isothermal titration calorimetry (ITC) binding isotherms for the titration of Tel-22 (Na\textsuperscript{+}) and Tel-12 (Na\textsuperscript{+}), panels A and B (top). Experiments were carried out in Tris buffer (0.020M Tris-HCl, pH 8.0, and 0.1M NaCl at 25 °C. The insets in panels A and B (top) represent the integration of the raw data to provide the binding enthalpies. The solid lines drawn through the data points represents best fits of a non-linear least squares model from which thermodynamic binding parameters can be estimated as described in Table I. The bottom panels (C and D) show the concomitant CD spectra upon titration of the Tel-22 (Na\textsuperscript{+}) (panel C) and Tel-12 (Na\textsuperscript{+}) (panel D) G-quadruplexes with UP1. In Panel C, the molar ratio ($r$) of UP1 to G-quadruplex spans the range of $r = 0 - 2.5$. In panel D, the molar ratio ($r$) of UP1 to G-quadruplex spans the range of $r = 0 - 1.5$. The black arrows represent the directionality of the ellipticity change at 295 nm as a function of increasing UP1.
Figure 4. Isothermal titration calorimetry (ITC) binding isotherm for the forward titration of Tel-22 (K\(^+\))-quadruplex DNA (panel A). Experiments were performed in Tris buffer (0.020M Tris-HCl, pH 8.0, and 0.1M KCl) at 25 °C. The inset represents the integration of the raw data to provide the binding enthalpy. The solid line drawn through the data points represents the best fit of a non-linear least squares model from which thermodynamic binding parameters are estimated as described in Table I. CD spectra upon titration of the Tel-22 (K\(^+\))-quadruplex DNA with UP1 is shown in panel B (bottom). The molar ratio (r) of UP1 to G-quadruplex DNA spans the range of r = 0 – 4.0. The black arrow represents the directionality of the ellipticity change at 295 nm as a function of the addition of UP1.
Figure 5: The isothermal titration binding isotherm (ITC) for the titration of Tel-22(SS) with UP1 (panel A). Experiments were performed in Tris buffer (0.020M Tris-HCl, pH 8.0, and 0.1M KCl at 25°C. Panel B: The integration of the raw data to provide the binding enthalpy and is represented in panel B. The solid lines drawn through the data points represents best fits of a non-linear least squares model from which thermodynamic binding parameters can be estimated as described in Table I.
Figure 6: Isothermal titration calorimetry (ITC) binding isotherms for the titration of UP1 with Tel-22 (Na$^+$) and Tel-12 (Na$^+$), panels A and B. Experiments were carried out in Tris buffer (0.020M Tris-HCl, pH 8.0, and 0.1M NaCl at 25 °C. The bottom panels C and D represent the integration of the raw data to provide the binding enthalpies. The solid lines drawn through the data points are the best fits of a non-linear least squares model from which thermodynamic binding parameters can be estimated as described in Table II.
CHAPTER V

INFLUENCE OF LOOP SEQUENCE MUTATIONS ON THE STRUCTURE AND STABILITY OF HUMAN TELOMERIC G-QUADRUPLEX DNA

by

Jason S. Hudson\textsuperscript{1} and David E. Graves\textsuperscript{1, 2}

\textsuperscript{1}Department of Chemistry, University of Alabama at Birmingham, Birmingham, AL 35294
\textsuperscript{2}Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL 35294

In preparation for submission to \textit{Biophysical Journal}

Format adapted for dissertation
Abstract

G-quadruplexes are higher order DNA structures that play a role in gene transcription and telomeric maintenance. The structure and stability of the G-quadruplex is under thermodynamic control and influenced by the base composition of the loop sequences. Here, we report the structural influence and energetic contributions to the overall stability of the adenine bases in the loop sequences of the human telomeric G-quadruplex DNA. Utilizing spectroscopic and calorimetric techniques, we have measured the thermal stability and thermodynamic contributions to the stability of human telomeric G-quadruplexes with systematic mutations of the adenines in the telomeric G-quadruplex sequence. Through these experiments we found that the enthalpy of unfolding folding has a direct relationship with the number of adenines present in the G-quadruplex loop sequences. In addition, when all adenines in the loops were mutated to thymine the melting temperature was reduced from 59° C for the wild type sequence to 47° C. The DSC thermograms had biphasic melting profile with two distinct transitions. In contrast, the multiple loop mutations sequences that did not exhibit the biphasic DSC thermograms and were characterized by only one melting transition. We contribute the biphasic unfolding profile and the reduction in the enthalpy of unfolding to the energetic contributions of adenine in the loop sequences. The adenine bases contribute to the overall enthalpy and stability by hydrogen bonding and stacking interactions with the G-tetrads.
Introduction

Guanine rich nucleic acids have been shown to form higher order DNA structures called G-quadruplexes (1, 2). Analysis of the human genome has established the presence of a large number of highly conserved guanine rich sequences that have potential for forming G-quadruplex structures (3-5). These sequences were found in oncogenic promoter sites as well as at the telomeric ends of the chromosome providing evidence for an in vivo biological function for G-quadruplex structures. The potential for G-quadruplex structures to serve regulatory functions has made them attractive targets for the development of novel therapeutic drugs to modulate genetic transcription and telomeric extension (6, 7). The core of G-quadruplex structures are planar G-tetrads that are composed of four guanines held together through Hoogsteen bonding. In vitro, G-quadruplexes may be constructed via intermolecular or intramolecular interactions with significant variations in sequence and molecularity. Oligonucleotides of sufficient length have the innate ability to form unimolecular G-quadruplexes with a high degree of topological variation that is dependent on sequence and cationic buffer conditions (8).

Human telomeric DNA is a non-coding region at the end of the chromosome that consists of a repetitive sequences that are approximately 150-200 kilo-bases in length (9). The terminal end of this sequence has a 3’ single stranded overhang with the tandem repeat of 5’-(TTAGGG)_{n}-3’ with n ranging from 10 to 50. These sequences have been demonstrated to form stable G-quadruplex structures in vitro (10, 11). The formation of G-quadruplex structure has been shown to inhibit the attachment of telomerase to the template telomeric DNA, thus halting the enzymatic extension of the telomere that is often associated with cancer cells (12, 13). The unimolecular human telomeric G-
quadruplex structure has three stacked G-tetrads connected by three -TTA- loops. The nature of the loop connectivity gives rise to alternate conformations that have been observed by both NMR and X-ray crystallography. The most notable structural differences are between the sodium (Na\(^+\)) and potassium (K\(^+\)) forms. Sodium based buffers result in an all antiparallel strand orientation of the G-quadruplex structure. The loop connectivities may vary and be all lateral loops, the “chair” conformation, or have two lateral loops and one diagonal loop, the “basket” conformation. Both of these folding topologies and their loop connectivity are illustrated in Figure 1. Potassium based buffers result in a more complex G-quadruplex structure that is characterized by two lateral loops and one loop that is a chain reversal oriented on the edge of the tetrad core. This results in a mixed strand polarity that has both antiparallel and parallel characteristics. The hybrid structure was described by Phan and Patel by NMR spectroscopy (14). An alternate structure derived from X-ray crystallography was reported by Neidle and coworkers (15). Nevertheless, G-quadruplexes have an amazing ability to form topologically diverse structures that are very stable at biological temperatures and buffer conditions. The stability of G-quadruplexes is primarily due to the formation of the G-tetrads, hydrophobic stacking of successive G-tetrads, cationic coordination with the G-tetrad by Na\(^+\) or K\(^+\), and loop sequence effects (16).

The energetics of G-quadruplex unfolding and the enthalpic (\(\Delta H_{\text{unfold}}\)) and entropic (\(\Delta S_{\text{unfold}}\)) contributions to the thermal stabilities have been reported by several research groups using a variety of techniques. There is a large degree of variation in the reported energetic values for G-quadruplex forming sequences (17). Differences in selected sequences and buffer conditions, as well as the methodologies are the basis for
inconsistencies reported in the literature. Here, we present an analysis of the thermal stabilities and the energetics of the unfolding of the 22 nucleotide human telomeric sequence using circular dichroism spectropolarimetry (CD) and differential scanning calorimetry (DSC). We have constructed a number of G-quadruplexes with loop mutations where the adenine base in the -TTA- sequence is mutated to thymine resulting in a -TTT- loop sequence. The buffer conditions and loop sequence length (3 nucleotides) remain consistent in an effort to determine role of the adenine bases on the energetic contributions within the loop sequences.

**Materials and Methods**

*Oligonucleotide Preparation and Nomenclature for Mutated Sequences*

Synthetic oligonucleotides with the selected A to T base mutations were purchased from Midland Reagents, Midland, TX. The oligonucleotides were solvated in 100 mM NaCl BPES buffer (0.01 M NaH$_2$PO$_4$, 0.01 M Na$_2$HPO$_4$, 0.001 M EDTA) and were allowed to equilibrate overnight at 4 °C. Prior to use, stock solutions were filtered using 0.45 micron Millipore syringe filters and the concentrations were determined by UV-visible spectrophotometry at 90° C to ensure that any residual secondary structure was fully denatured. The molar extinction coefficients were calculated using the nearest neighbor method as previously described (18). The solutions were annealed at a cooling rate of 0.5 degrees/min using a MJ Research Mini-Cycler. G-quadruplex formation was confirmed for the oligonucleotides by collecting CD wavelength scans from 225-325 nm and observing the presence of a positive ellipticity at 295 nm.
Loop mutations were designed to provide both single and multiple mutations in each of the three loops. Table 1 provides the sequence mutations and the nomenclature that will be used for this manuscript to distinguish them. The human telomeric sequence begins with a 5’ adenine and was also mutated to thymine in addition to the loops.

Circular Dichroism Spectropolarimetry

CD melting experiments were conducted using an AVIV 400 CD spectrophotometer (AVIV Biomedical, Inc., Lakewood, NJ) at a strand concentration of 3 µM in a 1-cm pathlength quartz cuvette. The diluted oligonucleotide solutions were allowed to equilibrate with stirring for fifteen minutes prior to each measurement. The thermal difference spectra revealed the largest change at 295 nm for all sequences and the CD signal at 295 nm was monitored as a function of temperature. The temperature program was designed to range from 10º C to 90º C at a rate of 1º C/minute. Measurements were taken at every 1º C after a one minute equilibration time. A buffer melt baseline was subtracted from each spectrum and the data was normalized to molar ellipticity. The melting temperature (Tm) was defined as the midpoint of the transition calculated by the first derivative of the sigmoidal curve. A model dependent van’t Hoff analysis of the melting curves was not performed due to the presence of sloping pre-transition baselines and undetermined heat capacity changes upon melting of the G-quadruplex structures.
Differential Scanning Calorimetry

The heat associated with the unfolding of the G-quadruplex was measured using a Microcal VP-DSC (GE Healthcare, Northampton, MA), at strand concentrations of 175-225 µM. Experiments were conducted with a heating rate of 1° C/min over a range from 10-100° C. A minimum of five heating and cooling cycles was conducted for each sample to assess reversibility of the unfolding. In addition, heating rates were reduced to 0.5° C/min to ensure that thermodynamic equilibrium was reached during the heating cycle. Buffer versus buffer scans were conducted in an identical manner for the subtraction of the reference scan. The baseline uncertainties that are inherent in the calorimetric analysis of G-quadruplex structures have been previously reported and discussed (19) and users should clearly state the method employed to treat the data. The data was analyzed using Origin 7.0 VP-DSC software with the following procedure. The third buffer/buffer scan was subtracted from the raw data and then normalized to strand concentration to provide heat capacity versus temperature plots. These data were then corrected for baseline effects by using the iterative baseline function integrated into the Origin 7.0 VP-DSC software. The baseline procedure was conducted by allowing the software to detect the linear portions of the pre- and post-transitions of the baseline and was only manually adjusted when there were obvious deficiencies in the selected position. The pre- and post-transition baselines were connected using the cubic function and the resulting baseline was subtracted from the data. The user can choose from a number of baseline connecting functions and can inadvertently subtract important energetic contributions from the data. This is particularly true for transitions that exhibit non-zero changes in heat capacity for the folded and unfolded structures or sloping
baselines. For our experiments, the cubic and progress baseline functions resulted in baselines that had a minimal effect on the transition. However, the progress baseline function encountered difficulties in providing a reliable baseline for the G-quadruplex with single loop mutations that exhibited pre-transition baseline slopes. Therefore, the cubic baseline function was consistently applied to all of the DSC scans contained within this study. After baseline subtraction, the transition was deconvoluted using the minimum number of Gaussian curves necessary to achieve a good fit. The total area under the curve was integrated to provide the enthalpy of unfolding (ΔH_{\text{unfold}}) as defined by the thermodynamic relationship $\Delta H_{\text{cal}} = \int \Delta C_p(T) \, dT$. The free energy ($\Delta G^\circ$) of unfolding can be calculated from the Gibb’s free energy equation $\Delta G^\circ = \Delta H^\circ - T \Delta S^\circ$ or alternatively at any reference temperature, $\Delta G^\circ(T) = \Delta H^\circ \left[ 1 - \frac{T}{T_m} \right]$. The midpoint of the transition is defined as the melting temperature ($T_m$) and corresponds to a population of 50% folded and 50% unfolded G-quadruplex structure.

**Results**

*Conformational Analysis and Thermal Denaturation by Circular Dichroism*

The conformation of the wild type and loop mutated G-quadruplex sequences were analyzed by CD spectropolarimetry. The wavelength scans are provided in Figure 2 and are either single loop mutations or multiple loop mutations. The wild type spectrum (red line) is included in both panels for comparison. All sequences exhibited positive ellipticity at 295 nm corresponding to an antiparallel strand orientation, but exhibited differences in the magnitude of the positive ellipticity. There were also variations in the minima at 265 nm and the secondary maxima around 245 nm. The single loop mutations
resulted in only a minimal shift of the wavelength but displayed significant differences in the magnitudes of the maxima and minima. Sequences with multiple loop mutations showed greater deviations from the wild type resulting in a higher magnitude of the positive ellipticity at 295 nm. In addition, there was a shift to higher wavelengths for the minimum at 265 nm and secondary maximum at 245 nm. Interestingly, the minimum did not go below the zero baseline and was no longer a negative ellipticity for the 0123TTT and 13TTT mutated sequences.

Direct interpretation of CD spectra of G-quadruplexes is difficult and any structural information aside from strand orientation is uncertain. The CD spectra represent the average solution conformation and are influenced by the syn/anti base conformation as well as base stacking interactions. Currently no empirical procedure for the direct analysis of G-quadruplex CD spectra exists that will provide the details of the folding topology in its entirety. However, the CD spectra of G-quadruplexes can easily determine the nature of the strand polarity for the folded structure. For the sequences used in this study, all were antiparallel arrangements.

Melting experiments monitoring the change in CD signal at 295 nm were conducted to determine the thermal stability of each loop mutant. Overlays for the single loop mutation and multiple loop mutation sequences are provided in Figure 3. The thermal denaturation curves appear to be monophasic and the T_m was estimated by taking the first derivative of the sigmoidal curve. The wild type G-quadruplex resulted in a melting temperature of 58°C and the majority of the mutants were ± 4°C of the wild type. In contrast, the 0123TTT mutant exhibited a reduction in the thermal stability with a T_m of 44°C, approximately 15°C lower than the wild type G-quadruplex. The results
of this study are summarized in Table 1 and reveal a correlation between the conversion of the adenine bases to thymine and the stability of the G-quadruplex. An unexpected result was obtained for the single loop mutant 2TTT which revealed a slightly enhanced thermal stability relative to the wild type of 62°C. The G-quadruplexes also exhibited differences in the pre-transition baselines. In cases of the wild type sequence and single loop mutants, there were apparent sloping pre-transitions baselines. In contrast, the multiple loop mutations did not exhibit sloping pre-baseline transitions. Sloping baselines are problematic for model dependent van’t Hoff analyses and could potentially induce systematic errors in the calculation of the melting temperature. For this reason, the CD data were not analyzed using a model dependent analysis to determine the thermodynamic contributions to the unfolding process.

**DSC Analysis of G-quadruplex Unfolding**

We utilized differential scanning calorimetry (DSC) to directly measure the heat of unfolding for the G-quadruplex sequences. The advantage of using DSC is that the measurement is model independent and does not rely on van’t Hoff plots to determine the enthalpy of the transition. The thermogram generated for the wild type human telomeric sequence in 100 mM NaCl BPES buffer is shown in Figure 4. The thermal transition is biphasic with a shoulder peak that precedes the larger transition that has a $T_m$ of 59°C. It was necessary to include two Gaussian curves (red lines) to deconvolute the raw data (black line) and achieve a suitable fit for the complete transition. The calorimetric enthalpy was calculated by the integration of the area under the curve and was determined to be 37.4 kcal/mol for the unfolding direction. The thermograms of the
single and multiple loop mutants are shown in Figures 5 and 6, respectively. The melting temperatures and calorimetric enthalpies were calculated in a manner similar to the wild type sequence and their results are reported in Table 2.

Comparisons of the melting temperatures derived from the CD data with the DSC results reveal that the T_m’s are in agreement. The most significant variations in T_m’s arise from DNAs with single loop mutations. These variations may be due to the pre-transitions baselines described for the CD data and shoulder peaks observed in the DSC data. The enthalpy of unfolding of the G-quadruplex sequences allowed determination of the thermodynamic contributions of the adenine bases. The enthalpy associated with unfolding the G-quadruplex was systematically reduced as the number of adenine to thymine mutations was increased. The average energetic contribution of an adenine in the loop was determined to be approximately 1-2 kcal/mol. It is interesting to note that the number of transitions necessary to fit the DSC data was reduced to 1 when the sequences with multiple loop mutations were examined. The DSC data was also in agreement with the CD data for mutations that resulted in the highest and lowest observed thermal stabilities. The 2TTT sequence melting temperature was 62º C and the 0123TTT sequence was destabilized to 47º C.

None of the DSC transitions could be fit with a simple 2-state model that assumes a negligible heat capacity change. Analysis of the raw data prior to baseline correction revealed non-zero heat capacity changes as would be expected for such a significant structural change. The single loop mutation sequences exhibited pre-transition sloping baseline effects that made ΔCp values difficult to determine. In contrast, the multiple loop mutation sequences did not have sloping pre-transition baselines and it was apparent
that the change in heat capacity was not zero with approximate values in the range of 0.7-1.0 kcal/mol. ITC studies have been reported non-zero heat capacity changes for G-quadruplex structures (20). Taking into account the non-zero changes in heat capacity and the apparent non-two-state unfolding mechanism, our analyses was limited to quantitation of the enthalpy of unfolding by integrating the area under the transition in an effort to limit the bias and assumptions embedded in the integrated models for data fitting.

**Discussion**

This work presents spectroscopic and calorimetric investigations on the influence of the adenine bases on the energetics and unfolding of the human telomeric G-quadruplex sequence. Our studies focus on the structural and thermodynamic properties associated with the systematic mutation of adenine to thymine residues in G-quadruplex loop sequences. There has been considerable work published that examines the effects of loop sequence (21-24) and loop length (25-27) on G-quadruplex stability. However, the work described in this manuscript provides an in-depth analysis of the influence of adenine to thymine base substitutions of the human telomeric G-quadruplex DNA.

*Structural Characterization of the Loop Mutated G-Quadruplexes*

Structural analysis of the mutated sequences by CD revealed that all of the selected sequences used for this study form antiparallel G-quadruplex structures as illustrated by a positive maximum band at 295 nm and a negative minimum band at 265 nm. G-quadruplex folding topologies are conveniently characterized by CD
spectropolarimetry where positive ellipticity at 295 nm is generally accepted as an indication of the antiparallel folding topology (28-30). The CD spectra for the loop mutated sequences reveal that G-quadruplex structures retained antiparallel strand orientations. However, subtle differences are observed in the magnitudes and the wavelengths of the maxima and minima. The single loop mutations exhibited differences in magnitudes of the maxima at 295 and 245 nm and minimum band at 265 nm. In contrast, the G-quadruplex sequences incorporated with multiple loop mutations had differences in magnitudes for the maxima and minima, as well as wavelength shifts as seen in Figure 2. Direct interpretation of the CD spectra of G-quadruplexes is ambiguous due to the limited empirical correlation of the CD spectrum with distinct folding topologies. The contributions to the CD spectra are influenced by sequence effects and base stacking interactions of the G-tetrads and loop nucleotides. CD analyses of G-quadruplexes are most suited for classifying the strand polarities (parallel and antiparallel orientations) and observations of changes in G-quadruplex structure. However, CD spectropolarimetry cannot unambiguously distinguish between conformations or structural isoforms of the same strand polarity. When analyzed in conjunction with the thermal unfolding data obtained by DSC, we speculate that the changes in the CD spectral characteristics are likely due to disruption of the stacking interactions by removing the adenine bases from the loop sequences. It is also conceivable that a loss of these stacking interactions may distort the geometry of the G-tetrads which could possibly result in subtle structural deviations relative to the wild type sequences that would give rise to the variations in the CD spectra.
Thermodynamic Profiles for the Unfolding of the G-quadruplexes.

The large hypochromic effects on the CD signal at 295 nm are observed for the denaturation of G-quadruplex structures and have been extensively used to characterize their thermal stabilities (8, 31-35). The data shown in Figure 3 depicts the CD melting experiments conducted on mutated loop sequences with the calculated melting temperatures summarized in Table 2. Systematic mutations consisting of adenine to thymine conversions in each of the loops resulted in changes in the thermal stabilities of the G-quadruplex structures. The majority of mutations resulted in only slightly lowered melting temperatures; however, when all four adenines were converted to thymine (0123TTT) the result G-quadruplex was significantly less stable ($T_m = 44^\circ C$). In contrast, the loop 2 mutant, 2TTT, had a slightly enhanced stabilization with an observed melting temperature of 62 °C. All of the transitions appeared to be monophasic and could be characterized by sigmoidal curves. However, a closer examination of the pre-transition baselines showed that there were differences between the single loop and multiple loop mutation sequences. All of the single loop mutations sequences had sloping pre-transition baselines, in contrast to the multiple loop mutation sequences that had minimal sloping effects observed for the baselines. A number of potential factors may cause sloping pre-transition baselines including a temperature dependence of the solvent, heat capacity effects, loss of base stacking interactions, or the presence of intermediates in the unfolding mechanism. The lack of clearly defined pre- and post transition baselines is problematic for the calculation of the thermodynamic parameters by van’t Hoff analysis and may not represent the true enthalpy of the unfolding process (36,37).
We utilized the model independent method of DSC to directly measure the heat of unfolding and characterize the thermodynamic parameters of the unfolding process. A biphasic transition profile was observed for the wild type human telomeric sequence is shown in Figure 4. Deconvolution of the DSC thermogram required two transitions with melting $T_m$ of 39° C and 59° C for the first and second transitions, respectively. Our research group as well as others have observed the biphasic melting profile for similar G-quadruplex sequences in potassium buffer (38-40). In addition to telomeric sequences, two distinct unfolding events were also observed for the c-MYC promoter region G-quadruplex as reported by Freyer and coworkers (41). The overall transition of the wild type human telomeric G-quadruplex is dominated by the peak at 59° C and the melting temperature derived from the midpoint of the transition is consistent with our CD melting experiments. The unfolding enthalpy was calculated to be 37 kcal/mol. Thermodynamic unfolding parameters determined by DSC for G-quadruplexes in the literature for exact sequences and solution conditions used in this study are limited making direct comparisons challenging. A review of G-quadruplex stability and kinetics by Lane and coworkers (19), has addressed the apparent discrepancies and noted that the basis for the thermodynamic parameters to be a combination of sequence variations, differences in buffer solution conditions, and methodologies employed to determine the parameters. Chaires recently reviewed the published thermodynamic data for the unfolding of human telomeric G-quadruplexes and noted that reported $T_m$ values in the literature varied from 56 to 63.7° C in 100 mM Na$^+$ buffer and the enthalpy values varied dramatically from 38 to 72.7 kcal/mol (17).
The thermal unfolding of the mutated loop sequences revealed a surprising result in that the low temperature transition was reduced in magnitude for the single loop mutations and was not observed at all for the multiple loop mutations. The 0T mutant most resembled the wild type and the additional single loop mutant transition profiles were broadened with only minor features present corresponding to the low temperature transition. All of the multiple loop mutations were fit with a single transition and lacked any evidence of the low temperature transition that was observed for the wild type single loop mutants. Not only did the presence of the adenine bases contribute to the shape of the melting transition, but also contributed to the calculated unfolding enthalpies. Table 2 summarizes the unfolding enthalpies and reveals that as the number of adenines mutated to thymines is increased, the unfolding enthalpy decreases. When all of the adenines were mutated to thymine in the 0123TTT sequence, the enthalpy of the transition was reduced approximately 15 kcal/mol from 37 kcal/mol observed for the wild type to 22 kcal/mol for the 0123TTT sequence. We postulate that the adenine bases are involved in important stacking interactions with adjacent G-tetrads and contribute to the G-quadruplex stability (42). Our DSC analyses of systematic adenine to thymine mutations throughout the loops provide a quantitative measure of those interactions. The losses of adenine stacking explain the reduction in the thermal stability for most of the G-quadruplex sequences that were examined. An outlier would be the 2TTT mutant that exhibited an enhanced thermal stability of 62° C (+ 3° C). This may result from removal of the adenine in the second loop allowing for a more favorable stacking interaction of the 5'-adenine at the beginning of the sequence with the bottom G-tetrad. In addition to stacking interactions, thermodynamic contributions also arise from the loops ionic...
interactions and hydration effects (43). However, these effects are thought to contribute only modestly to the folding enthalpy. Calculation of the free energy at a reference temperature is a more useful measure of G-quadruplex stability as compared with the melting temperature. The calculated free energies at 25° C for each of the sequences are provided in Table 2. The enhanced thermal stability of the 2TTT loop mutation is not as significant when comparing the free energies of unfolding of the wild type with the mutated 2TTT sequence, with both free energies calculated to be 3.8 kcal/mol. The remaining mutated sequences show a systematic reduction in the free energy with the 0123TTT mutant possessing the lowest calculated free energy at 25° C of 1.5 kcal/mol.

The Biophysical Nature of the Unfolding Transition

The molecular basis for the low temperature transitions that are observed for the DSC profiles is a more complex issue and rather controversial within the field. As previously mentioned, direct comparisons of published results are often difficult due to variations in the G-quadruplex sequence, structures, and preparation methodologies. One possible explanation that could describe the biphasic melting profile observed in the DSC studies would be the presence of multiple G-quadruplex conformers on that thermally unfold independent of each other. The sodium form of the human telomeric G-quadruplex can assume either the “chair” and/or “basket” folding topologies in solution (2). It is plausible that an A to T mutation within the loop sequences may influence the folding topology and favor one form over the other conformation. However, if that were the case we would not expect the T_m of the low temperature transition to deviate as much as 15° C as observed for the 2TTT mutation. The T_m of the low temperature transitions
that were apparent in the wild type and single loop mutations appeared to be linked to the high temperature transition and were never clearly resolved with a baseline between the transitions. Further, we are not aware of any published biphasic DSC thermograms of G-quadruplexes that have been able to be resolved by stabilizing one of the transitions relative to the other using variations in salt concentrations, scan rates, or sequence variations. Without a detailed structural analysis of the wild type and mutated sequences, it is difficult to definitively state that the low temperature transition is the chair or basket form and the high temperature transition is the other conformation.

An alternate possibility for the biphasic profile is the presence of intermediates in the unfolding pathway. It has been proposed by Sheardy and coworkers that the human telomeric G-quadruplex does not unfold in a simple two state manner but that there are significantly populated intermediates states between the folded and unfolded states. Analysis of CD spectra collected as a function of temperature for the K⁺ human telomere G-quadruplex DNA were analyzed by singular value decomposition (SVD) and suggested the presence of an intermediate species along the unfolding pathway (38). Fluorescence studies have also been reported that both Na⁺ and K⁺ forms of the G-quadruplex have populated intermediate states or rapid equilibrium between multiple conformers (44,45). From our data, we speculate that the biphasic melting profile of the human telomeric sequence is due to the presence of unfolding intermediates and characterized by the loss the enthalpic contributions of the loop base stacking interactions represented by the low temperature transition. The presence of single loop mutations significantly reduces the magnitude of the low temperature transition, and the multiple loop mutation sequences resulted in a complete loss of the low temperature transition.
Mutation of the terminal 5'-adenine (0T), that is not a loop base, had only a minimal effect on the biphasic profile and closely resembled the wild type. The unfolding of a G-quadruplex structure may involve a mechanism wherein the loop structures begin to denature prior to the denaturation of the G-tetrad core. Disruption of adenine-guanine stacking interactions and hydrogen bonding between the bases in the loops would have significant enthalpic consequences and are observable by DSC. It is important to point out that neither the low or high temperature transitions of the denaturing profile represent the intermediate structure. Instead we postulate that the initial unfolding step is loop structure destabilization (low temperature transition), which is then followed by the breaking of the G-tetrad core of the structure (high temperature transition). In situations where multiple loops had been mutated and only the high temperature transition was observed, the multiple loop mutants could not be fit with a simple two state model. These data suggest that the G-quadruplex unfolds by initial loop denaturation followed by disassembly of the G-tetrads at higher temperatures. The low temperature transition is no longer apparent because the adenine-guanine base stacks that were responsible for the low temperature transition and are no longer present in the sequence.

In summary, the spectroscopic and calorimetric analysis of the unfolding of the human telomeric sequence provides insight into the energetic contributions of the adenine bases that are present in the loop sequences. The presence of adenines in each of the loop sequences correlates to the observed biphasic melting profiles. The structure and stability of the human telomeric G-quadruplex is under dynamic control and this study provides evidence that base sequence modifications within the loop regions of the G-quadruplex markedly influence the stability and energetics of the structure. Additionally, these
results presented here suggest that G-quadruplex thermodynamic stability is a combination of loop sequence effects and G-tetrad formation.
References


Table 1: Nomenclature for the mutated G-quadruplex sequences.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Loop Mutation</th>
<th>Oligonucleotide Nomenclature</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'-AGGGTTAGGGTTAGGGTTAGGG-3'</td>
<td>-</td>
<td>wild type</td>
</tr>
<tr>
<td>5'-TGGGTTAGGGTTAGGGTTAGGG-3'</td>
<td>0</td>
<td>0T</td>
</tr>
<tr>
<td>5'-AGGGTTTGGGTAGGGTTAGGG-3'</td>
<td>1</td>
<td>1TTT</td>
</tr>
<tr>
<td>5'-AGGGTTAGGGTTTGGGTAGGG-3'</td>
<td>2</td>
<td>2TTT</td>
</tr>
<tr>
<td>5'-AGGGTTAGGGTTAGGTGG-3'</td>
<td>3</td>
<td>3TTT</td>
</tr>
<tr>
<td>5'-AGGGTTTGGGTGTGGGTAGGG-3'</td>
<td>1,2</td>
<td>12TTT</td>
</tr>
<tr>
<td>5'-AGGGTTAGGGTTTGGGTAGGG-3'</td>
<td>2,3</td>
<td>23TTT</td>
</tr>
<tr>
<td>5'-AGGGTTTGGGTAGGTGGGTGG-3'</td>
<td>1,3</td>
<td>13TTT</td>
</tr>
<tr>
<td>5'-AGGGTTTGGGTGGTTGGTGG-3'</td>
<td>1,2,3</td>
<td>123TTT</td>
</tr>
<tr>
<td>5'-TGGGTTTGGGTGTGGGTGG-3'</td>
<td>0,1,2,3</td>
<td>0123TTT</td>
</tr>
</tbody>
</table>
Table 2: Summary of DSC and CD thermal denaturation experiments.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Number of Transitions</th>
<th>DSC</th>
<th>CD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>T_{max} (°C)</td>
<td>ΔH_{cal(unfold)} (kcal/mol)</td>
</tr>
<tr>
<td>wtQuad</td>
<td>2</td>
<td>59</td>
<td>37.4</td>
</tr>
<tr>
<td>0T</td>
<td>2</td>
<td>59</td>
<td>36.1</td>
</tr>
<tr>
<td>1TTT</td>
<td>2</td>
<td>55</td>
<td>35.7</td>
</tr>
<tr>
<td>2TTT</td>
<td>2</td>
<td>62</td>
<td>34.3</td>
</tr>
<tr>
<td>3TTT</td>
<td>2</td>
<td>55</td>
<td>34.6</td>
</tr>
<tr>
<td>12TTT</td>
<td>1</td>
<td>59</td>
<td>32.3</td>
</tr>
<tr>
<td>23TTT</td>
<td>1</td>
<td>57</td>
<td>30.0</td>
</tr>
<tr>
<td>13TTT</td>
<td>1</td>
<td>59</td>
<td>32.1</td>
</tr>
<tr>
<td>123TTT</td>
<td>1</td>
<td>59</td>
<td>28.0</td>
</tr>
<tr>
<td>0123TTT</td>
<td>1</td>
<td>47</td>
<td>22.3</td>
</tr>
</tbody>
</table>
Figure 16. Cartoon representing the chair (left) and basket (right) conformations of the Na⁺ human telomeric G-Quadruplex. The blue circles represent the guanines and the red circles represent the adenines.
Figure 17. CD wavelength scans for the human telomeric G-quadruplex d[AGGG(TTAGGG)]₃ (red line) and the loop mutated sequences. Panel A represents the sequences that possess single adenine to thymine base mutations. Panel B represents the sequences with multiple base mutations.
Figure 18. Composite CD melting curves of the human telomeric sequence and loop mutated structures in 100 mM NaCl BPES buffer at pH 7.0. Panel A contains the CD melting curves for the wild type sequence (red line) and the single base mutated sequences. Panel B contains the CD melting curves for the G-quadruplex sequences with multiple loop mutations. Melting temperatures were calculated from the first derivative of the sigmoidal curve. Data was only normalized to molar ellipticity and not converted to fraction folded/unfolding in order to retain the pre-transition baseline sloping characteristics.
Figure 19. A DSC thermogram of the wild type human telomeric G-quadruplex DNA in 100mM NaCl BPES buffer at pH 7.0. The observed transition (black line) is biphasic with a lower temperature transition at 39 °C and higher temperature transition at 59 °C. The observed transition was fit to two transitions (red lines) and the total area under the curve was integrated to provide the enthalpy of unfolding ($\Delta H_{\text{unfold}}$).
Figure 20. Representative DSC thermograms for the single base mutations of the G-quadruplex. The black lines represent the raw DSC data and the red lines represent the number of transitions. (Panel A) Mutation of the 5'-A (0T); (Panel B) Mutation of loop 1 (1TTT); (Panel C) Mutation of loop 2 (2TTT); (Panel D) Mutation of loop 3 (3TTT).
Figure 21. Representative DSC thermograms for the multiple loop mutations of the G-quadruplex. The black lines represent the raw DSC data and the red lines represent the number of transitions. (Panel A) wild type; (Panel B) Mutations of loop 1 and 2 (12TTT); (Panel C) Mutations of loop 2 and 3 (23TTT); (Panel D) Mutations of loop 1 and 3 (13TTT); (Panel E) Mutations of loop 1, 2, and 3 (123TTT); (Panel F) Mutations of all adenines in the sequence to thymines (0123TTT).
CHAPTER VI

THE INFLUENCE OF LOOP SEQUENCE MUTATIONS ON THE BINDING AND UNFOLDING OF HUMAN TELOMERIC G-QUADRUPLEX DNA BY UNWINDING PROTEIN 1 (UP1)

by

Jason S. Hudson¹; Lei Ding²; Jinbiao Ma²; Edwin Lewis³ and David Graves¹.⁴

¹Department of Chemistry, University of Alabama at Birmingham, Birmingham, AL 35294

²Department of Biochemistry, University of Alabama at Birmingham, AL 35294

³Department of Chemistry, Mississippi State University, Mississippi State, MS 39762

⁴Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL 35294

In preparation for submission to Biochemistry

Format adapted for dissertation
Abstract

Unfolding Protein 1 (UP1) is a proteolytic derivative of heterogeneous ribonucleoprotein A1 (hnRNP A1) that retains the two nucleic acid binding domains that are highly conserved in the hnRNP A1 family. UP1 has a high affinity for telomeric DNA repeat sequences d(TTAGGG)_n and destabilizes the G-quadruplex structure formed by the sequence. We have systematically mutated the G-quadruplex TTA loop sequences to TTT in an effort to investigate the fundamentals of G-quadruplex recognition and unfolding by UP1. The research presented here utilizes isothermal titration calorimetry (ITC) to determine the binding energetics and circular dichroism spectropolarimetry (CD) to monitor the unfolding of G-quadruplex structure. Our studies reveal that the middle loop (loop 2) to be critical for recognition and binding of UP1. Mutations of adenine to thymine in the loop 2 sequence significantly reduced the binding affinity and enthalpy of binding as well as the degree of unfolding of the G-quadruplex structure. These results suggest that the adenine-guanine base stacking of the G-quadruplex structure play a role in the recognition factor and binding of UP1. In addition, we have also investigated the interactions of each individual DNA binding domains (BD1 and BD2) of UP1 with the telomeric G-quadruplex. CD studies reveal that when the domains are not physically linked, the degree of unfolding is minimal. ITC binding studies demonstrated BD1 to be responsible for G-quadruplex recognition with a high binding affinity for G-quadruplex structure, whereas BD2 has a much low affinity and less specific binding. These data provide key insights with respect to UP1 loop sequence recognition and correlation of the binding energetics with the unfolding of G-quadruplex structure.
Introduction

Proteins that interact with G-quadruplex DNA’s are of interest in part because of the biological implications for regulation of gene transcription and telomeric extension (1). The immortality of cancerous cells is closely associated with both telomeric length and the extension of the telomere by the enzyme telomerase which results in an indefinite number of cell replication cycles (2, 3). The binding of telomerase to the telomeric ends of the chromosome requires a single stranded DNA template and accessory proteins that assist in G-quadruplex unfolding. Once unfolded, telomerase binds to the template telomeric DNA and initiates reverse transcription to extend the telomere (4, 5). The unfolding of G-quadruplexes is accomplished by proteins that are members of the heterogeneous ribonucleoprotein family (hnRNP). These proteins are non-catalytic and have been demonstrated to unfold a diverse set of G-quadruplex DNA structures. Two of these proteins, hnRNP D and hnRNP A1, have been shown to bind and destabilize the telomeric G-quadruplex and are closely associated with telomeric maintenance and extension (6,7).

Unwinding protein 1 (UP1) is derived from hnRNP A1 through the proteolytic removal of the glycine rich amino acid region of the C-terminus. UP1 retains the two nucleic acid binding motifs of hnRNP A1 (8, 9). In 2002, Fakuda and coworkers reported that UP1 could destabilize the intramolecular G-quadruplex structures that were formed in hypervariable minisatellite sequences, d(GGCAG)₅, and in the human telomeric sequence, d(TTAGGG)₄, (10). The interactions of the two binding domains of UP1, BD1 and BD2, with the 5'-TTAG-3' sequence in short telomeric sequence are thought to be driven by a variety of non-covalent forces, including hydrogen bonding,
stacking interactions, van der Waals interactions, and electrostatic interactions (11).

However, key interactions between the two binding domains with the DNA are the
stacking interaction of a phenylalanine residue between adenine and guanine nucleotides.

Intramolecular G-quadruplexes formed from the human telomeric sequence
(d[AGGG(TTAGGG)₃]) are composed of 5'-TTA-3' connecting loop structures with the
next guanine in the sequence forming the G-tetrad that is located on either top or bottom
“face” of the G-quadruplex structure. The solution based structure of the human
telomeric G-quadruplex reported by Wang and Patel in 1993 showed that three G-tetrads
formed the core of the G-quadruplex and the loop sequences were oriented above and
below the structure (12). Based on these findings we theorized that is may be possible
for UP1 to bind G-quadruplex structure by recognition of the exposed loop sequences.

Our laboratory has recently conducted a biophysical analysis of the energetic and
structural characteristics of UP1 binding to a bimolecular G-quadruplex, d(TTAGGG)₂,
and the unimolecular Na⁺ and K⁺ G-quadruplex conformations of the human telomeric
sequence, d[AGGG(TTAGGG)₃]. Our results indicated that UP1 binds to all of the
selected G-quadruplex structures, but could not efficiently unfold the K⁺ unimolecular G-
quadruplex. The degree of unfolding of the Na⁺ form of the G-quadruplex was nearly
100 percent for both unimolecular and bimolecular G-quadruplexes. In contrast, UP1
induced unfolding of the K⁺ G-quadruplex was only 50 percent effective. By measuring
the enthalpy of the binding interactions of the selected G-quadruplexes with UP1, we
concluded that the increased thermal stability of the K⁺ G-quadruplex and higher energy
barrier for unfolding could not be overcome by the observed binding enthalpy of the
protein-DNA complex formation. Our results revealed the binding energetics of UP1 to
the G-quadruplex structure to be closely coupled with the unfolding process. The binding affinities for UP1 interactions with the G-quadruplex DNAs and a control DNA sequence that does not form a G-quadruplex structure revealed a markedly higher affinity for the G-quadruplex structure, suggesting a unique specificity for the higher order DNA structure that may be of biological significance.

Additional questions remain regarding the mechanistic details associated with UP1 binding and the destabilizing the human telomeric G-quadruplex DNA. The focus of the work presented here is to examine the effects of the base sequence in the loops regions of the G-quadruplex on UP1 recognition and induced unfolding of the Na+ G-quadruplex conformation of human telomeric sequence. We have utilized isothermal titration calorimetry (ITC) and circular dichroism (CD) spectropolarimetry to characterize the influence of adenine to thymine mutations in the loop sequences, (TTA → TTT), on the binding energetics and unfolding of G-quadruplex DNA by UP1. We have also investigated the role of the two UP1 binding domains (BD1 and BD2) in the interaction with G-quadruplex DNA. From these experiments we have determined the differences in affinity for the G-quadruplex structure and synergy that exists between the two binding domains and their role in the unfolding the G-quadruplex structure. The results of this study provide novel insights into the influence of the loop sequences for the recognition, binding, and subsequent unfolding of the human telomeric G-quadruplex by UP1.
Materials and Methods

Oligonucleotide Preparation and Design

Synthetic oligonucleotides were purchased from Midland Certified Reagents, Midland, TX and used without further purification. The wild type human telomeric G-quadruplex sequence is d[AGGG(TTAGGG)₃] and a cartoon depicting the “basket” and “chair” conformations is shown in Figure 1. A series of base sequence mutations (5’-TTA-3’ to 5’-TTT-3’) in the three loops regions of the G-quadruplex were synthesized and are summarized in Table 1. The oligonucleotides were dissolved in Tris buffer (0.020 M Tris-HCl, pH 8.0, 0.1 M NaCl), and were allowed to equilibrate overnight at 4º C. Prior to use the stock solutions were filtered using 0.45 micron Millipore syringe filters and the concentrations determined by UV-Visible spectrophotometry at 90º C. The molar extinction coefficients used for concentration determinations were calculated using the nearest neighbor method as previously described (13). The solutions were annealed using a MJ Research Mini-Cycler by a single heating and cooling cycle with a cooling rate of 0.5º C/min. The formation of the G-quadruplex structure was confirmed for the oligonucleotides by collecting CD wavelength scans from 225-325 nm and observing the presence of a positive ellipticity at 295 nm. Positive ellipticity at 295 nm is considered to be characteristic for the antiparallel G-quadruplex structure (14).

Protein Preparation

The cDNA of the UP1 (1-196 amino acid sequence of hnRNP A1) was cloned into the vector pET28-SMT3 to result in an N-terminal SUMO tag and 6x His tag and expressed in Escherichia coli BL21-Gold(DE3). Cells were grown in LB media for 3 hr at 37º C until OD reached 0.6 absorbance units at 600 nm. Induction with 0.1 mM
isopropyl-β-D-1-thiogalactopyranoside was performed overnight at 18° C before harvesting the cells. Cells were washed with PBS and resuspended in 5 mL of Tris buffer (20 mM Tris pH 8.0, 0.5 M NaCl, 25 mM imidazole) per 100 mL of culture. Cells were disrupted by Emulsiflex-C3 cell homogenizer (Avestin Inc.) and then followed by centrifugation for 40 minutes at 16,000 rpm while the temperature was maintained at 4°C. The supernatant was loaded on 5 mL His-Trap column (GE Health) and eluted with elution buffer (20 mM Tris pH 8.0, 0.5 M NaCl, 500 mM imidazole). The eluted protein was digested with SUMO protease and dialyzed against 20 mM Tris pH 8.0 (0.5 M NaCl) overnight. The elution procedure was repeated and the flow through was collected. It was further purified by a Heparin column to exclude the protein that complexed with nucleic acids. The collected protein was fractionated by gel filtration using a Superdex 75 in a 16/60 column (GE Health). The concentration of protein was determined by UV-vis spectrophotometry (Varian, Inc).

**Isothermal Titration Calorimetry**

Isothermal titration calorimetry (ITC) experiments were performed using a Microcal VP-ITC (GE Healthcare, Piscataway, NJ) at 25° C. All samples were thoroughly degassed with stirring prior to use. For forward titrations G-quadruplex DNA was placed in the sample cell and titrated with UP1 protein. The time between injections was 400 seconds. Replicate titrations were performed and the resulting data were corrected for heat of dilution. The resulting binding isotherms were non-sigmoidal and could not be analyzed with the standard binding models incorporated into the Microcal Origin VP-ITC software used to calculate the energetics of the interaction. It was apparent from the shapes of the isotherms that multiple events were occurring and a three
event binding model was necessary to achieve a suitable fit for the binding isotherms. The three-event binding algorithm was developed by combination of the mass balance and equilibrium constant expressions expanded to n-independent site binding expressions. Lewis and co-workers previously reported the details regarding the development of the fitting algorithm to describe complex ITC binding isotherms and is the basis for the model used for our fitting procedure (15).

When multiple events having enthalpic contributions to the binding isotherm are present, it is useful to reverse the direction of the titration. We conducted reverse titrations by placing UP1 in the sample cell and titrating with G-quadruplex DNA. The binding isotherms generated from the reverse ITC experiments were sigmoidal and could be fit using a single-site binding model. It should be noted that the resulting thermodynamic parameters of the reverse titration fits represent an average of all events in the reaction, however these titrations are useful to reduce the uncertainties of the binding stoichiometries and binding enthalpies determined from the forward titration fits.

**G-Quadruplex Unfolding Monitored by Circular Dichroism Spectropolarimetry**

Experiments that monitored the unfolding of the G-quadruplex structure were conducted using an AVIV 400 circular dichroism spectrophotometer (Aviv, Inc.). Experiments were performed in 1-cm pathlength rectangular cells at 25°C. The concentration of G-quadruplex in the cell was 3 µM (strand) and a sufficient volume of UP1 was injected to result in a concentration of 6 µM protein in the cell (2:1 molar ratio). Immediately upon injection data collection began and the CD signal was monitored at 295 nm for 400 seconds with constant stirring. The dead time between injection and
initiation of data collection was 1-2 seconds. The kinetic traces of the G-quadruplex signal at 295 nm revealed the degree of unfolding that induced by the binding of UP1 to the G-quadruplex structure. Experiments were identically conducted for the wild type G-quadruplex and the loop mutated sequences, and the complete loss of ellipticity at 295 nm was assumed to indicate the complete unfolding of the G-quadruplex structure. The degree of G-quadruplex unfolding was calculated by determining the percent change in the CD signal at 295 nm and used for the comparison of all G-quadruplex sequences.

Results

Effects of the Loop Sequence Mutations on UP1 Mediated G-quadruplex Unfolding

The loss of the CD signal of G-quadruplex DNA at 295 nm upon interaction with UP1 was recorded for the wild type as well as the mutated loop sequences. Figure 2 provides data for the single loop mutations of the G-quadruplex DNA (Fig. 2A) and the G-quadruplex sequences with multiple loop mutations (Fig. 2B). For reference, the wild type G-quadruplex (wt-hTel) is represented in both panels. Our earlier results had demonstrated sodium form of wild type G-quadruplex to be completely unfolded at a molar ratio of 2:1 (UP1 to G-quadruplex DNA). Experiments were designed such that a single injection of UP1 would achieve a 2:1 molar ratio, and the value of the CD signal at 295 nm after 400 seconds was used to evaluate the degree of unfolding of the G-quadruplex DNA. Relative to wild type, the single loop mutations (1TTT, 2TTT, and 3TTT) exhibited varying influences on the unfolding of the G-quadruplex. After 400 seconds, the CD$_{295\text{nm}}$ signal of the 1TTT loop mutated sequence was near 0 mdeg and was very similar to the wild type sequence. The 3TTT retained only slightly more G-quadruplex structure than the 1TTT with a signal near 1.0 mdeg. The 2TTT loop mutated
sequence did not have the same degree of unfolding and retained approximately 50% of the G-quadruplex structure after the injection UP1.

The G-quadruplex sequences where multiple loops had been mutated (12TTT, 23TTT, 13TTT, and 123TTT) all had considerable effects on the degree of unfolding. The 12TTT and 23TTT G-quadruplexes resulted in a CD$_{295\text{nm}}$ signifying a reduction in the degree of unfolding by approximately 25%. The most severe effect was the sequence where all loops had been mutated with less than 50% of the G-quadruplex structure remaining in the folded state. Interestingly, the 13TTT mutation where the second loop remains a wild type sequence achieved near complete unfolding and had the highest degree of unfolding for all of the multiple loop mutated sequences. Examination of the unfolding data revealed that the G-quadruplex sequences that had the second loop mutated (2TTT, 12TTT, 23TTT, and 123TTT) all had significant reductions in the degree of G-quadruplex unfolding after interacting with UP1.

Effects of the Loop Sequence Mutations on the Binding Energetics

Isothermal titration calorimetry (ITC) was used to determine the thermodynamic parameters for the interaction of UP1 with wild type and G-quadruplex DNAs mutated in each of the loops 1, 2, and 3. Forward titrations were conducted wherein the DNA was placed in the calorimeter sample cell and titrated with UP1. Figures 3 and 4 represent the binding isotherms for the single loop and multiple loop mutations, respectively. The binding isotherm for the wild type G-quadruplex (Fig. 3A) exhibits a complex shape that appears to be composed of multiple events with overlapping enthalpic contributions. The isotherm is non-sigmoidal at molar ratios of less than two and are composed of at least
two enthalpic contributions to the signal that are opposite in sign (exothermic and endothermic). Presumably, the first event to occur is the binding of the UP1 protein and is characterized by a large and negative enthalpy of binding \(-\Delta H_{\text{bind}}\). Initial binding of UP1 to G-quadruplex is highly favored with a large and favorable binding enthalpy. The CD experiments indicate that the G-quadruplex structure is induced to unfold upon complex formation with UP1. Therefore, it is likely the second event is the unfolding of the G-quadruplex and overlaps with the first. G-quadruplex unfolding is an enthalpically unfavorable process and would be characterized by a positive enthalpy of unfolding \(+\Delta H_{\text{unfold}}\). Indeed, the observed positive peaks between injections 4 through 9 suggest that the positive enthalpy terms are due to the unfolding of the G-quadruplex structure induced by the binding UP1. The ITC experiments measure the total enthalpy of the reaction and the overlapping endothermic and exothermic enthalpic components are summed together resulting in the complex shape observed in the binding isotherm up to a molar of 2:1. The binding isotherm becomes sigmoidal at protein to DNA ratios greater than 2:1; consistent with our CD studies which demonstrate that unfolding of the G-quadruplex structure is complete upon reaching a molar ratio of 2:1. However, the wt-hTel binding isotherm clearly depicts a binding event at molar ratios greater than 2:1. This suggests that the second UP1 is binding to the telomeric DNA in an unfolded state at these molar ratios.

Using this information, we utilized a three event fitting algorithm to determine the thermodynamic parameters for each event occurring during the interaction. The second event is likely the enthalpy of unfolding the G-quadruplex and is not considered to be a binding event. The binding isotherms for the wild type sequence and the single loop
mutants are represented in Figure 3. The binding stoichiometry for the wild type and all of the single loop mutated sequences was determined to be two molecules of UP1 to one G-quadruplex. The recognition and binding of the first UP1 molecule to the wild type G-quadruplex (Fig. 3A) exhibits a high affinity ($K_1 = 8.5 \times 10^8 \text{ M}^{-1}$) and favorable enthalpy of binding ($\Delta H_1 = -43.1 \text{ kcal/mol}$). The binding of the second UP1 molecule to the unfolded G-quadruplex structure (third event) exhibits a 100 fold lower binding affinity ($K_1 = 6.3 \times 10^6 \text{ M}^{-1}$) and less a favorable binding enthalpy ($\Delta H_3 = -30.1 \text{ kcal/mol}$). The loop 1 mutation (1TTT) and loop 3 mutation (3TTT) sequences had similar binding energetics to that observed for the wild type G-quadruplex (Fig. 3B and Fig. 3D). In contrast, the loop 2 mutant (2TTT) exhibited a significant deviation from the wild type sequence. The multi-event binding isotherm was not as evident and the isotherm appeared to be more hyperbolic (Fig. 3C). The isotherm was best described by a two event model and the energetics of the initial binding event was more than an order of magnitude lower in affinity ($K_1 = 1.4 \times 10^7 \text{ M}^{-1}$) and the enthalpy ($\Delta H_1 = -19.7 \text{ kcal/mol}$) was approximately half of that observed for the wild type G-quadruplex. Table 2 summarizes the thermodynamic parameters obtained for the interactions of UP1 with the wild type and loop mutant G-quadruplexes.

The binding isotherms for the forward titrations of the multiple loop mutant G-quadruplex sequences with UP1 are depicted in Figure 4 and summarized in Table 3. 12TTT (Fig. 4A) and 23TTT (Fig. 4B) had very similar binding energetics and the calculated stoichiometry was 2.0. The first binding event for 12TTT had a binding constant estimated to be $3.7 \times 10^8 \text{ M}^{-1}$ and the binding enthalpy was $-37.5 \text{ kcal/mol}$. The 23TTT sequence also had a high affinity first binding event, $4.6 \times 10^8 \text{ M}^{-1}$, and slightly
more favorable binding enthalpy, -43.2 kcal/mol. The binding affinity of the second UP1 molecule to 12TTT and 23TTT were identical, \( (K_3 = 2.2 \times 10^6) \). The binding enthalpy was found to be less favorable for the 12TTT mutant \( (\Delta H_3 = -8.6 \text{ kcal/mol}) \) than that of the 23TTT mutant \( (\Delta H_3 = -12.7 \text{ kcal/mol}) \). The 13TTT G-quadruplex (Fig. 4C) had the largest first event binding affinity \( (K_1 = 5.7 \times 10^8 \text{ M}^{-1}) \) and the most favorable binding enthalpy \( (\Delta H_1 = -20.7 \text{ kcal/mol}) \) in comparison to the 12TTT and 23TTT sequences. The binding of the second UP1 molecule to 13TTT was characterized by a binding affinity of \( 7.8 \times 10^6 \text{ M}^{-1} \) and a -17.4 kcal/mol binding enthalpy. The 123TTT G-quadruplex (Fig. 4D) resulted in a weaker binding isotherm that was not capable of being accurately described with our fitting model.

Reverse titrations were conducted in an attempt to reduce the complexity of the binding isotherms and confirm the binding stoichiometry of the interactions of UP1 with single and multiple loop mutants of the G-quadruplex. The reverse titration was conducted by placing the UP1 protein in the sample cell and titrating with DNA. The large excess of UP1 in the sample cell resulted in the multi-step reaction being driven to completion after each injection of DNA. The observed binding isotherms were sigmoidal and could easily be fit with a single site binding model. The reverse direction of the titration results in an isotherm that represents the average enthalpy of all the events that were observed in the forward direction. Figure 5 provides the experimental results of the single loop mutated sequences. Similar to the forward titrations, the reverse titration of UP1 with the wild type G-quadruplex (Fig. 5A) had the highest observed average binding constant \( (K_a = 7.4 \times 10^7 \text{ M}^{-1}) \) and a favorable binding enthalpy \( (\Delta H_1 = -32 \text{ kcal/mol}) \). The single loop sequences had similar trends as observed in the forward titration with
2TTT (Fig. 5C) with the lowest binding affinity ($K_a = 5.6 \times 10^6 \text{ M}^{-1}$), but surprisingly the binding enthalpy was quite favorable ($\Delta H = -35 \text{ kcal/mol}$). The apparent increase in binding enthalpy but low binding affinity is likely due to the limited unfolding that we observed in the CD studies for the 2TTT mutant, and therefore a minimal contribution of positive enthalpy of unfolding. The 1TTT (Fig. 5B) and 3TTT (Fig. 5D) had very similar binding affinities ($K_a = \sim 1.0 \times 10^7 \text{ M}^{-1}$) that were similar in magnitude to that observed for the wild type sequence and favorable binding enthalpies ($\Delta H = -31 \text{ kcal/mol}$ and -36 kcal/mol). The wild type sequence binding stoichiometries and single loop mutated sequences were all 0.5 signifying two molecules of UP1 bound to one molecule of G-quadruplex oligonucleotide.

The reverse titrations were useful for observing the overall energetics of the multiple loop mutated sequences that were poorly resolved isotherms in the forward titrations. Figure 6 shows the collected data that resulted in a sigmoidal shape for all of the binding isotherms. The 12TTT (Fig. 6A) and 23TTT (Fig. 6B) had very similar affinities ($K_a = 3.4 \times 10^6 \text{ M}^{-1}$ and $3.5 \times 10^6 \text{ M}^{-1}$), but a significant difference in binding enthalpies ($\Delta H = -22$ and -13 kcal/mol). 13TTT (Fig. 6C) had the highest affinity ($K_a = 8.8 \times 10^6 \text{ M}^{-1}$) and most favorable binding enthalpy ($\Delta H = -27 \text{ kcal/mol}$) of all the multiple loop mutated sequences. In contrast to the forward titrations, we were able to achieve binding isotherms that could be fit to estimate the binding energetics for 123TTT (Fig. 6D). The affinity was low relative to the wt-hTel ($K_a = 4.1 \times 10^6 \text{ M}^{-1}$) and the binding enthalpy was significantly reduced ($\Delta H = -16 \text{ kcal/mol}$). An intriguing result derived from the reverse titration was the 1:1 binding stoichiometry for the 123TTT
sequence, whereas all other G-quadruplex oligonucleotides were 2:1. The results for the reverse titration experiments are provided in Table 4.

**Effect of UP1 Binding Domain Deletions on the Binding and Unfolding of G-Quadruplex DNA**

We performed identical experiments as outlined above with a UP1 protein that was expressed with only one of the binding domains present. The interaction of G-quadruplex DNA with the individual domains (BD1 or BD2) was investigated to characterize differences in their DNA binding energetics and capacity for induced unfolding. Figure 7 represents the ability to unfold G-quadruplex DNA when the binding domains are not tethered together. It is clear from the ellipticity at 295 nm neither BD 1 or BD2 alone could unfold the wild type G-quadruplex. The ITC forward titrations with the individual domains revealed striking differences in the binding affinities and energetics of complex formation of the individual domains with the G-quadruplex DNA. Figure 8A represents BD1 alone and the data could be described by two events with a high affinity binding constant for the first event \(K_1 = 1.9 \times 10^8 \text{ M}^{-1}\) accompanied by a very favorable binding enthalpy \(\Delta H_1 = -39 \text{ kcal/mol}\). In contrast, the interaction of BD2 alone (Fig. 8B) could only be described by one event and the result was a weak binding affinity \(K_1 = 4.5 \times 10^5 \text{ M}^{-1}\) and less binding enthalpy \(\Delta H_1 = -8.0 \text{ kcal/mol}\). These results indicate distinct differences between the binding of BD1 and BD2 to G-quadruplex DNA and a specific interaction of BD1 with G-quadruplex DNA. The results are summarized in Table 5.
Discussion

These studies provide an in-depth characterization of the energetic and structural features of the interaction of UP1 and human telomeric G-quadruplex DNA. The influence of loop sequence mutations on the energetics associated with complex formation was examined. Systematic mutation of the adenine to thymine in the G-quadruplex loop sequences (TTA → TTT) has significant influence on recognition, binding, and unfolding by UP1. CD experiments were used to monitor G-quadruplex unfolding and revealed that mutation of the second loop (2TTT) has considerable influence on both the binding and the unfolding process. Sequences that retained wild type sequence in the loop 2 (1TTT, 3TTT, and 13TTT) were all observed to unfold to near completion upon binding UP1. In contrast, the G-quadruplex sequences that were a mutated in Loop 2 (2TTT, 12TTT, 23TTT, and 123TTT) revealed a significant reduction in ability of UP1 to induce unfolding. The UP1 mediated unfolding process is accomplished non-catalytically with no requirement for an external source of energy. We postulate that the energy source for unfolding is coupled with the highly favorable binding enthalpy. The enthalpies for the formation of G-quadruplex structure of identical or closely related sequence and buffer conditions of approximately 35-40 kcal/mol have been reported by our laboratory and additional research groups (16-18). In order to destabilize and unfold the structure the binding enthalpy of UP1 must be large enough to overcome the unfavorable unfolding enthalpy of the G-quadruplex structure. Our studies revealed that the sequences with the highest degree of unfolding had binding enthalpies that were close to or exceeded the unfolding enthalpy of the G-quadruplex structure. The reduction in the degree of unfolding for the G-quadruplex DNAs with a loop 2
mutation and less favorable binding energetic suggests that the second loop is a vital component for UP1 mediated unfolding of the G-quadruplex DNA structure.

The energetics of the binding interaction derived from the ITC studies of UP1 with the mutated loop sequences also revealed significant differences for the sequences that had a mutation in the loop 2 sequence. Analysis of the complex shape of the forward titrations due to overlapping enthalpic components was challenging, however upon deconvolution our data revealed a three event binding model that could explain the shape of the binding isotherm. Our previous studies on the interaction of UP1 with the Na+ conformation of human telomeric G-quadruplex DNA used circular dichroism spectropolarimetry to characterize the structural features of the G-quadruplex at distinct molar ratios of UP1 to G-quadruplex DNA and correlate these structural properties to specific regions on the ITC binding isotherm. The G-quadruplex DNA is shown to be completely unfolded at a UP1 to G-quadruplex molar ratio greater than 2.0, and binding at higher ratios must be to the unfolded G-quadruplex sequence.

Analysis of the NMR solution structure of the G-quadruplex sequence reported by Wang and Patel, reveals insights into the basis for a preference of UP1 for binding the loop 2. The basket structure has two lateral loops (loop 1 and 3) and one diagonal loop (loop 2) that are exposed to the surface of the G-quadruplex. The adenine of loops 2 and 3 appear to stack on the adjacent guanine nucleotides that form the top and bottom G-tetrads. Ding and coworkers reported that the phenylalanine amino acid residues in the two binding DNA domains inserted between adenine and guanine bases in the co-crystal structure of UP1 with short oligonucleotide sequences (11). Meyers and coworkers examined basis for purine recognition by the binding domains of UP1 and found that the
substitution of 2-aminopurine, nebularine, or 7-deazaquanine for the guanine residues in the sequence d(TTAGGG)$_2$ was poorly tolerated and greatly reduced the binding affinity (19). They proposed that the presence of the fluorescent probes reduced the stacking interaction of the adenine nucleotide and influenced the insertion of the phenylalanine residue between the base stacks. It is possible that the replacement of guanine with 2-aminopurine, nebularine, or 7-deazaquanine disrupted the folding of the G-quadruplex structure and the sequence remained single stranded. Our studies demonstrate UP1 binding to single-stranded DNA to have a lower binding affinity relative to binding a G-quadruplex structure.

A model for analyzing the experimental data for the forward titrations was designed to quantitatively account for three enthalpic contributions to the isotherm, the initial binding interaction, subsequent unfolding of the G-quadruplex structure, and the binding of the second UP1 molecule. A cartoon representing a potential binding mechanism and the basis of the three event model is postulated in Figure 9. The three event model accurately describes the shape of the isotherm and allowed us to estimate the binding energetic of complex formation. In conjunction with these studies reverse titrations were carried out to provide an alternate measure of the overall binding stoichiometry. The energetics of the third event (binding of the second UP1 molecule) of are directly correlated with the degree of G-quadruplex unfolding after binding the first UP1. Therefore, the energetics of the first binding event was examined for significant differences that could be related to the results of the CD experiments.

A general trend that was noted was that changes in the sequence of loop 2 resulted in markedly lower binding affinities and binding enthalpies. Loop 2 is exposed to the
exterior of the G-quadruplex and it is plausible to consider that the initial binding interaction occurs by inserting the phenylalanine residue of the UP1 binding domains between the adenine-guanine base stack of loop 2. Mutating the adenine of loop 2 to thymine removes the adenine-guanine stacking interaction and appears to inhibit the recognition by UP1, therefore reducing the binding affinity and unfolding capability that we have observed in our experimental results.

Additionally, our experiments with UP1 domain deletions (BD1 or BD2) provide insights into the synergy between the two domains for G-quadruplex unfolding and compelling evidence that BD1 plays the key role in the recognition of G-quadruplex DNA structure. The CD experiments demonstrate that when the two domains are linked the protein is highly effective in unfolding the G-quadruplex. However, either domain alone has minimal capability to unfold the structure. The difference in the binding energetics was striking with BD1 having a high binding affinity \((1.9 \times 10^8 \text{M}^{-1})\) and very favorable binding enthalpy. In contrast, BD2 had a weak binding affinity \((4.5 \times 10^5 \text{M}^{-1})\) and less favorable binding enthalpy. The differences in binding energetics and the necessity for both domains to be tethered in order to unfold G-quadruplex structure suggests that BD1 specifically recognizes the G-quadruplex and places BD2 in close proximity to the secondary binding site allowing non-specific binding to the DNA assisting in the unfolding process. Recently, Nagata and coworkers reported that BD1 has a higher affinity for DNA and BD2 has a higher binding affinity for RNA \((20)\).

In the work presented here, we report the first biophysical investigation on the influence of loop sequence mutations on the binding and unfolding of human telomeric G-quadruplex by UP1. These results provide evidence for specific recognition of G-
quadruplex and binding interactions with loop 2 of the G-quadruplex. The calorimetric analyses of the interactions provide a basis for an energetic interpretation of the G-quadruplex unfolding induced by UP1. These studies also developed a three event binding model that allows analysis the complex non-sigmoidal ITC binding isotherms. Structural data regarding the interaction of UP1 with the full length 22 nucleotide telomeric G-quadruplex is limited, and all of the reported crystal structures have used shortened 12 nucleotide telomeric sequences. The 2:1 (UP1 to G-quadruplex) binding stoichiometry is consistent with the crystal structures as well as recent reports that utilized the full length 22 nucleotide sequence (19, 21). The location of the two UP1 molecules on the oligonucleotide sequence in the bound state for the full length sequence remains unclear. Structural and mechanistic questions remain, however efforts are currently underway to further characterize these interactions and refine our model to provide a more detailed representation of the mechanism through which UP1 recognizes, binds, and unfolds telomeric G-quadruplex DNA.
References

Table 1. Telomeric G-quadruplex loop sequence mutations.

<table>
<thead>
<tr>
<th>Sequence 5' - 3'</th>
<th>Loop Mutation</th>
<th>Oligo Nomenclature</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGGGTTAGGGTTAGGGT</td>
<td>-</td>
<td>wt-hTel</td>
</tr>
<tr>
<td>AGGGTGTAGGGTTAGGG</td>
<td>1</td>
<td>1TTT</td>
</tr>
<tr>
<td>AGGGTTAGGGTTAGGGT</td>
<td>2</td>
<td>2TTT</td>
</tr>
<tr>
<td>AGGGTTAGGGTTAGGGT</td>
<td>3</td>
<td>3TTT</td>
</tr>
<tr>
<td>AGGGTTAGGGTTAGGGT</td>
<td>1,2</td>
<td>12TTT</td>
</tr>
<tr>
<td>AGGGTTAGGGTTAGGGT</td>
<td>2,3</td>
<td>23TTT</td>
</tr>
<tr>
<td>AGGGTTAGGGTTAGGGT</td>
<td>1,3</td>
<td>13TTT</td>
</tr>
<tr>
<td>AGGGTTAGGGTTAGGGT</td>
<td>1,2,3</td>
<td>123TTT</td>
</tr>
</tbody>
</table>
Table 2. Thermodynamic parameters derived from the forward titration ITC binding studies of the G-quadruplex sequences with single loop mutations.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>( n_{\text{tot}} )</th>
<th>( K_{\text{obs}} ) (M(^{-1}))</th>
<th>( \Delta H_{\text{obs}} ) (kcal/mol)</th>
<th>( -T\Delta S^\circ ) (cal/mol-K)</th>
<th>( \Delta G^\circ ) (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt-hTel</td>
<td>2.0</td>
<td>( K_1 = 8.5 \times 10^8 )</td>
<td>( \Delta H_1 = -43.1 )</td>
<td>( -T\Delta S_1 = 30.1 )</td>
<td>( \Delta G_1 = -12.1 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( K_2 = 6.1 \times 10^7 )</td>
<td>( \Delta H_2 = -6.65 )</td>
<td>( -T\Delta S_2 = -3.95 )</td>
<td>( \Delta G_2 = -10.6 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( K_3 = 6.3 \times 10^8 )</td>
<td>( \Delta H_3 = -30.1 )</td>
<td>( -T\Delta S_3 = 20.8 )</td>
<td>( \Delta G_3 = -9.27 )</td>
</tr>
<tr>
<td>1TTT</td>
<td>2.0</td>
<td>( K_1 = 4.0 \times 10^8 )</td>
<td>( \Delta H_1 = -38.5 )</td>
<td>( -T\Delta S_1 = 26.8 )</td>
<td>( \Delta G_1 = -11.7 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( K_2 = 1.6 \times 10^7 )</td>
<td>( \Delta H_2 = -2.37 )</td>
<td>( -T\Delta S_2 = -7.45 )</td>
<td>( \Delta G_2 = -9.82 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( K_3 = 1.6 \times 10^8 )</td>
<td>( \Delta H_3 = -19.8 )</td>
<td>( -T\Delta S_3 = 11.3 )</td>
<td>( \Delta G_3 = -8.44 )</td>
</tr>
<tr>
<td>2TTT</td>
<td>2.0</td>
<td>( K_1 = 1.4 \times 10^7 )</td>
<td>( \Delta H_1 = -19.7 )</td>
<td>( -T\Delta S_1 = 9.97 )</td>
<td>( \Delta G_1 = -9.72 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( K_2 = 1.3 \times 10^6 )</td>
<td>( \Delta H_2 = -5.15 )</td>
<td>( -T\Delta S_2 = -2.83 )</td>
<td>( \Delta G_2 = -8.33 )</td>
</tr>
<tr>
<td>3TTT</td>
<td>2.0</td>
<td>( K_1 = 2.8 \times 10^8 )</td>
<td>( \Delta H_1 = -39.7 )</td>
<td>( -T\Delta S_1 = 28.2 )</td>
<td>( \Delta G_1 = -11.5 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( K_2 = 2.9 \times 10^7 )</td>
<td>( \Delta H_2 = -2.92 )</td>
<td>( -T\Delta S_2 = -7.25 )</td>
<td>( \Delta G_2 = -10.2 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( K_3 = 1.3 \times 10^8 )</td>
<td>( \Delta H_3 = -20.6 )</td>
<td>( -T\Delta S_3 = 12.2 )</td>
<td>( \Delta G_3 = -8.35 )</td>
</tr>
</tbody>
</table>
Table 3. Thermodynamic parameters derived from the forward titration ITC binding studies of G-quadruplex sequences with multiple loop mutations. The 123TTT mutant resulted in a weak binding isotherm that could not be analyzed.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>$n_{tot}$</th>
<th>$K_{obs}$ (M$^{-1}$)</th>
<th>$\Delta H^\circ_{obs}$ (kcal/mol)</th>
<th>$-T\Delta S^\circ$ (cal/mol·K)</th>
<th>$\Delta G^\circ$ (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12TTT</td>
<td>2.0</td>
<td>$K_1 = 3.7 \times 10^8$</td>
<td>$\Delta H_1 = -37.5$</td>
<td>$T\Delta S_1 = 25.8$</td>
<td>$\Delta G_1 = -11.7$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_2 = 1.6 \times 10^7$</td>
<td>$\Delta H_2 = 1.25$</td>
<td>$-T\Delta S_2 = -11.1$</td>
<td>$\Delta G_2 = -9.83$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_3 = 2.2 \times 10^6$</td>
<td>$\Delta H_3 = -8.36$</td>
<td>$-T\Delta S_3 = -0.27$</td>
<td>$\Delta G_3 = -8.63$</td>
</tr>
<tr>
<td>23TTT</td>
<td>2.0</td>
<td>$K_1 = 4.6 \times 10^8$</td>
<td>$\Delta H_1 = -43.1$</td>
<td>$-T\Delta S_1 = 31.3$</td>
<td>$\Delta G_1 = -11.8$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_2 = 3.3 \times 10^7$</td>
<td>$\Delta H_2 = -3.19$</td>
<td>$-T\Delta S_2 = -7.05$</td>
<td>$\Delta G_2 = -10.2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_3 = 2.1 \times 10^6$</td>
<td>$\Delta H_3 = -12.7$</td>
<td>$-T\Delta S_3 = 4.06$</td>
<td>$\Delta G_3 = -8.63$</td>
</tr>
<tr>
<td>13TTT</td>
<td>2.0</td>
<td>$K_1 = 5.6 \times 10^8$</td>
<td>$\Delta H_1 = -20.6$</td>
<td>$T\Delta S_1 = 8.68$</td>
<td>$\Delta G_1 = -11.9$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_2 = 1.6 \times 10^7$</td>
<td>$\Delta H_2 = 64.2$</td>
<td>$-T\Delta S_2 = -74.1$</td>
<td>$\Delta G_2 = -9.82$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_3 = 7.8 \times 10^6$</td>
<td>$\Delta H_3 = -17.3$</td>
<td>$-T\Delta S_3 = 8.04$</td>
<td>$\Delta G_3 = -9.39$</td>
</tr>
<tr>
<td>123TTT</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>
Table 4. Thermodynamic parameters derived from the reverse titration ITC binding studies.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>$n$</th>
<th>$K_a$ (M$^{-1}$)</th>
<th>$\Delta H_{obs}$ (kcal/mol)</th>
<th>-$T\Delta S^\circ$ (cal/mol$\cdot$K)</th>
<th>$\Delta G^\circ$ (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt-hTel</td>
<td>0.46</td>
<td>$7.4 \times 10^7 \pm 0.9$</td>
<td>$-31.8 \pm 0.8$</td>
<td>24.1</td>
<td>-10.7</td>
</tr>
<tr>
<td>1TTT</td>
<td>0.49</td>
<td>$1.3 \times 10^7 \pm 0.2$</td>
<td>$-30.7 \pm 0.5$</td>
<td>21.0</td>
<td>-9.7</td>
</tr>
<tr>
<td>2TTT</td>
<td>0.52</td>
<td>$5.6 \times 10^6 \pm 0.4$</td>
<td>$-34.6 \pm 0.4$</td>
<td>27.8</td>
<td>-9.2</td>
</tr>
<tr>
<td>3TTT</td>
<td>0.46</td>
<td>$9.4 \times 10^6 \pm 1.2$</td>
<td>$-36.3 \pm 0.7$</td>
<td>26.7</td>
<td>-9.5</td>
</tr>
<tr>
<td>12TTT</td>
<td>0.56</td>
<td>$3.4 \times 10^6 \pm 0.4$</td>
<td>$-22.3 \pm 0.5$</td>
<td>13.5</td>
<td>-8.9</td>
</tr>
<tr>
<td>23TTT</td>
<td>0.46</td>
<td>$3.3 \times 10^6 \pm 0.5$</td>
<td>$-12.2 \pm 0.5$</td>
<td>3.24</td>
<td>-8.8</td>
</tr>
<tr>
<td>13TTT</td>
<td>0.45</td>
<td>$8.8 \times 10^5 \pm 0.8$</td>
<td>$-26.9 \pm 0.4$</td>
<td>17.4</td>
<td>-9.4</td>
</tr>
<tr>
<td>123TTT</td>
<td>1.01</td>
<td>$4.7 \times 10^6 \pm 0.5$</td>
<td>$-15.9 \pm 0.3$</td>
<td>6.9</td>
<td>-9.0</td>
</tr>
</tbody>
</table>
Table 5. Thermodynamic parameters derived from the ITC binding studies of the UP1 binding domains with the wild type G-quadruplex DNA.

<table>
<thead>
<tr>
<th>Binding Domain</th>
<th>( n_{\text{tot}} )</th>
<th>( K_{\text{obs}} ) (M(^{-1}))</th>
<th>( \Delta H_{\text{obs}} ) (kcal/mol)</th>
<th>( -T\Delta S^\circ ) (cal/mol·K)</th>
<th>( \Delta G^\circ ) (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BD1</td>
<td>2.0</td>
<td>( K_1 = 1.9 \times 10^8 )</td>
<td>( \Delta H_1 = -39.2 )</td>
<td>( -T\Delta S_1 = 27.9 )</td>
<td>( \Delta G_1 = -11.3 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( K_2 = 5.7 \times 10^7 )</td>
<td>( \Delta H_2 = -3.12 )</td>
<td>( -T\Delta S_2 = -6.11 )</td>
<td>( \Delta G_2 = -9.22 )</td>
</tr>
<tr>
<td>BD2</td>
<td>1.0</td>
<td>( K_1 = 4.6 \times 10^6 )</td>
<td>( \Delta H_1 = -43.1 )</td>
<td>( -T\Delta S_1 = 31.3 )</td>
<td>( \Delta G_1 = -11.8 )</td>
</tr>
</tbody>
</table>
Figure 1. The sequence and structural representation of human telomeric G-quadruplex DNA
Figure 2. Circular dichroism studies of the unfolding of wild type G-quadruplex DNA and sequences with single loop mutants (panel A) and multiple loop mutants (panel B).
Figure 3. ITC binding isotherms for the forward titrations of the single loop mutations of the telomeric G-quadruplex sequence.
Figure 4. ITC binding isotherms for the forward titrations of the multiple loop mutations of the telomeric G-quadruplex sequence
Figure 5. ITC binding isotherms for the reverse titrations of the single loop mutations of the telomeric G-quadruplex sequence
Figure 6. ITC binding isotherms for the reverse titrations of the multiple loop mutations of the telomeric G-quadruplex sequence
Figure 7. Circular dichroism studies of the unfolding of wild type G-quadruplex DNA after the interaction with the individual nucleic acid binding domains of UP1.
Figure 8. ITC binding isotherms for the forward titrations of the wild type telomeric G-quadruplex with binding domain 1 (panel A) and binding domain 2 (panel B) of UP1.
Figure 9. Cartoon representing a three event model for the binding and unfolding of telomeric G-quadruplex DNA by UP1. The forward titration binding isotherms are provided below each event representing the components of the observed enthalpy.
CHAPTER VII
CONCLUSIONS

The research conducted in this dissertation provides an in-depth investigation of the fundamental biophysical properties of human telomeric G-quadruplex DNA and characterizes novel binding interactions of a small molecule and a protein with the G-quadruplex structure. The recent discovery of G-quadruplex forming sequences in the telomeric regions of the chromosome has resulted in an intense interest in the fundamental nature of the G-quadruplex structure and stability as well as for the identification of ligands and proteins that have specificities for the DNA structure. The recognition and binding of proteins to G-quadruplex DNA structures lends credence to the idea that these structures may play a regulatory role in biological processes such as replication and gene transcription.

Collaborative efforts with Dr. Jinbiao Ma and Lei Ding, UAB Department of Biochemistry and Molecular Genetics, led us to our investigations on the interaction of the protein UP1 with the telomeric G-quadruplex DNA. The UP1 protein is an abundant nuclear protein and closely associated with telomerase activity leading to exciting biological implications. Previous reports implicated that UP1 could bind G-quadruplex forming sequences and destabilize the structure. However, very little information was available with regard to the binding energetics and/or the mechanism of G-quadruplex
unfolding. Our research has provided considerable new insight into the recognition, binding, and unfolding of G-quadruplex DNA by UP1. We report an in-depth thermodynamic analysis of the interaction of UP1 with both the Na\(^+\) and K\(^+\) conformations of the telomeric G-quadruplex DNA that resulted the emergence of an energetic argument for the UP1 protein’s ability to unfold G-quadruplex structure. In order to overcome the innate structural stability of the folded G-quadruplex DNA a considerable enthalpic energy barrier must be exceeded. We found that UP1 binds the G-quadruplex structure and couples the energetics of the binding interaction to unfold the G-quadruplex structure and maintain an unstructured conformation in the bound state.

To further develop our research of the interaction of UP1 with telomeric G-quadruplex DNA we pursued an investigation of the effects of loop sequence mutations on the binding interaction. Prior to examining the influence of the loop sequence on the binding interaction we examined the structure and stability of G-quadruplex DNAs with loop sequence mutations. Systematic purine to pyrimidine mutations (TTA → TTT) of the G-quadruplex loop sequences were designed and the calorimetric investigation revealed that the thermal stability of the G-quadruplex was influenced by the loop sequence mutations. An example is the sequence where all of the adenines are mutated to thymines and the thermal stability was decreased by 15° C. DSC experiments allowed us to directly monitor the change in heat capacity as a function of temperature and derive the enthalpy of unfolding of the G-quadruplex structure. We were able to parse the enthalpic contributions of the base stacking interactions between the adenine and guanine nucleotides and quantitate their contributions to the overall enthalpy of unfolding. As the number of adenine to thymine mutations was increased the apparent enthalpy of
unfolding decreased systematically. An intriguing result of the study was the loss of the biphasic melting profile of the G-quadruplex as the number of loop sequence mutations increased. The biphasic nature of the transition has been noted by several research groups who actively study G-quadruplex DNA structure and stability, but the phenomenon is poorly understood. The DSC melting profile is composed of a low temperature transition and a high temperature transition. The higher temperature transition is commonly used as a measure of the thermal stability of the G-quadruplex.

As the number of adenine to thymine mutations in the loop sequence was increased we observed a complete loss of the low temperature transition suggesting that the transition was due to the loss of the enthalpic contribution of the adenine-guanine base stacking interactions. We concluded that the low temperature transition was due to the adenine and guanine base stacking interaction and the unfolding pathway of the telomeric G-quadruplex is preceded by un-structuring of the loop sequences prior to the denaturation of the G-tetrads.

We combined our efforts from the previous projects and we investigated the influence of loop sequence mutations on the recognition, binding, and unfolding of G-quadruplex structure by UP1. Structural studies of UP1 and short telomeric sequences have suggested that there are key binding interaction between the phenyalanine residues in the nucleic acid binding domains of the protein and the adenine-guanine base stacks of the G-quadruplex structure. Using telomeric G-quadruplex sequences that had systematic mutations of the loop sequence we demonstrated that the loops of the G-quadruplex structure are critical components for recognition and unfolding by UP1. We concluded that loop 2 of the G-quadruplex structure was clearly the recognition site for UP1 binding.
and any mutation that removed the adenine from the second loop sequence was poorly tolerated and drastically reduced the binding energetics and subsequent UP1 unfolding. Our research also noted differences in the binding affinities of the individual nucleic acid binding domains (BD1 and BD2) for G-quadruplex DNA. BD1 has a very high affinity and favorable binding energetic for G-quadruplex DNA. In contrast, BD2 interacts very weakly with the structure. In the case where the binding domains were not physically linked together there was little capability to unfold the structure and suggests a synergy between the DNA binding domains to accomplish the task of efficiently unfolding the G-quadruplex.

In addition to protein interactions with G-quadruplex DNA we also investigated the interaction of the DNA binding ligand, actinomycin D, with the telomeric G-quadruplex DNA. Based on our previous research focused on actinomycin D binding to guanine rich sequences in duplex DNA, we investigated the potential for an interaction with the guanine rich sequences of the human telomere that formed G-quadruplex structures. We were the first research group to report that actinomycin D could bind to the G-quadruplex formed by the telomeric sequence d[AGGG(TTAGGG)₃], and induce a structural modulation upon complex formation. We concluded from our research that actinomycin D bound to both the Na⁺ and K⁺ structural conformations with similar binding affinity and the structures of the ligand-G-quadruplex complexes as characterized by CD spectropolarimetry had distinct structural similarities. The structural characteristics of the bound G-quadruplexes were unique to either of the initial structures and their thermal stability was markedly increased. The ligand induced structural
conversion provided our first insights into the dynamic nature of the G-quadruplex DNA structures.

Our investigations have proven to be beneficial and insightful in understanding the biophysical properties that govern ligand and protein interactions with the telomeric G-quadruplex DNA. The research that is encompassed in this dissertation has provided key insights regarding ligand and protein interactions with G-quadruplex structures, and also leads to new questions and avenues of research to pursue. Energetic and structural characterizations have allowed us to accomplish our research objectives and have made it clear that the G-quadruplex DNA is a dynamic nucleic acid conformation that has only begun to reveal its functional aspects and complex interactions in vivo. Perhaps one of the more exciting prospects for future research is the high probability that there yet to be discovered ligands and proteins that specifically recognize and interact with G-quadruplex DNAs.
REFERENCES


