UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Structural and thermodymanic studies of the thrombin-binding aptamer and related quadruplexes

Permalink
https://escholarship.org/uc/item/9r27q226

Author
Tang, Chungfei Esther

Publication Date
2006

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Structural and Thermodynamic Studies of the Thrombin-binding Aptamer
and Related Quadruplexes

by

Chungfei Esther Tang

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry and Chemical Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

[redacted]

Date University Librarian

Degree Conferred:.....

LIBRARY

UCSF LIBRARY

This dissertation is dedication to

My parents – Jokkee Tang 鄭作祺 and Pitching Wong 黃碧清

My sister – Chungshyuan Stephanie Tang

My brothers – Chungshen Johnson Tang & Chungjeng 'Jengz' Tang

And

My beloved son – Yangsheng William Liang

UCSF LIBRARY

ACKNOWLEDGEMENTS

I am first of all indebted to my advisor, Dr. Richard H. Shafer, for his guidance in the last few years. His understanding, encouragement, and patience were the basis for my strength in making it all the way through the graduate program. He has helped me in every aspect of my life with invaluable intellectual discussions and reiterative explanations on many subjects, even going so far as to explain something for the 'tenth' time, helped me to become a better scientist. I have learned many lessons of life simply by knowing him.

I want to thank Dr. Martin Shetlar and Dr. Alan Frankel, who are in my dissertation committee, for their helpful suggestions and insights on my research project. Dr. Shetlar, with his lab right next to ours, is always willing to address my doubts or problems in experimental skills with a welcoming smile. I am also thankful to Dr. Thomas James and Dr. James Cleaver, for the helpful discussions with me when I was preparing to advance to candidacy years ago.

Postdoctoral researchers in the Shafer Lab deserve a formal recognition for I would not be able to finish my experimental work without their mentorship. Dr. Ivan Smirnov gave me the head start when I first joined the lab. I really appreciate his patience and wisdom in helping with troubleshooting and analysis. All metal ion – nucleic acid studies would not have gone as smoothly without Dr. Pinaki Majhi's expertise in calorimetry and his patience in giving me almost a step-by-step recipe and detailed knowledge of the instruments. Dr. Jianying Qi has definitely enriched my life in the lab since she joined us in recent years. I want to thank her for her emotional support as well as all the pep talks about my

UCSF LIBRARY

graduate work, career, and life.

I am especially indebted to our beloved CCB graduate program coordinator. I can never thank her enough for always believing in me (more than I believed in myself!) and refusing me to give up. I really appreciate her time, love, and all the tissue papers I used in the last few months! I am very proud to have a group of wonderful classmates and our precious get-together time. Ngoc-Ha Nguyen, Feixia Chu, Anne Sarver, Alina Dykman, Chandreyee Das, and Nicola Clegg, I wouldn't trade anything else for your friendship.

None of these would be possible without the unconditional love and support of my parents, Jokkee and Pitching, who have given me everything they could and are always there for me during all my ups and downs. I truly appreciate their time and effort in not only raising me, but also their help for taking care of my wonderful son, Yangsheng William Liang, so that I can finish this PhD program. I'd like to thank my sister, Stephanie, who is younger than me but has always been the one 'taking care' of me and who is now the world's greatest aunt to my son. My life would not be complete without my two dearest brothers, Johnson and Jengz, who always believe in me. I am also happy to have my sister-in-law and brother-in-law, Joel Grey and Lingleng Chan, in our family.

In addition to my family, I am lucky to have a wonderful buddy, Joyce Yeung to walk with me, shoulder to shoulder, despite all the outrageous ups and downs in the last few years. I am also grateful to all my friends especially Chienling Chen, Miki and Sean Park, Leeann Liao, Pauline Cheng, Thingfong Kwan, Andreas and Jeanet Ebel, Kichun Lee.

HSE LIBRARY

Structural and Thermodynamic Studies of the Thrombin-binding Aptamer and Related Quadruplexes

Chungfei Esther Tang

Chemistry and Chemical Biology

Approved by:

Richard H Shafer

Chairperson of Committee

ABSTRACT

Nucleic acid quadruplex structures possess several unique properties compared to the canonical nucleic acid structures. These structures are unique in their sequence dependence and metal ion requirements. Evidence for the presence and functional roles of quadruplexes in cellular systems has been increasing rapidly in the past decade. They have also proven to be useful as diagnostic and chemotherapeutic agents. Understanding of the structure and stability of quadruplexes is essential to elucidation of their biological interactions.

Antiparallel quadruplexes contain guanine stretches that are antiparallel to each other, in which half of the nucleosides have the *anti* conformations about the glycosidic bonds and half possess the *syn* conformations. In contrast, the guanine stretches in the parallel topology are parallel to each other with all nucleosides in the *anti* conformation. Although the antiparallel fold is more

NCSF LIBRARY

prevalent, parallel quadruplexes are becoming increasingly common. The influence of individual nucleoside conformation on the overall folding topology was examined in order to understand the forces that determine the quadruplex fold. Selective substitution of rG for dG allowed us to reverse the antiparallel fold of the thrombin binding aptamer (TBA). In addition, this approach also converts a unimolecular quadruplex into a bimolecular one. The reverse substitutions of dG for rG in the all-RNA analogue of TBA confirmed the change in topology from parallel to antiparallel conformation as well as the change from bimolecular to unimolecular. Changing conformational properties of guanine nucleosides allows us to control the quadruplex folding topology.

Metal ion – nucleic acid interactions are crucial in biological system and have significant impact on structures and functions of macromolecules. Quadruplexes are particularly unique in their unusual dependence on specific metal ions for their formation and stabilization through binding via inner sphere coordination. The thermodynamic behavior of quadruplexes stabilized by monovalent and divalent cations was studied by Isothermal Titration Calorimetry (ITC) at various temperatures. Results show that the enthalpy of quadruplex formation is temperature dependent and non-zero heat capacities were determined. These changes in the heat capacity are discussed in terms of changes in exposure of polar and nonpolar groups upon quadruplex formation.

UCSF LIBRARY

TABLE OF CONTENTS

Dedication.....iii

Acknowledgments.....iv

Abstract.....vi

Table of Contents.....viii

List of Tables.....xii

List of Figures.....xiii

CHAPTER 1: INTRODUCTION.....1

1.1 What is quadruplex?.....2

1.1.1 Structural Polymorphism.....2

1.1.2 Metal Ion Specificity.....8

1.2 History.....9

1.3 Biological significance.....10

1.3.1 Possible Biological Roles.....11

A. Telomere and telomerase.....11

B. Bloom's and Werner's Syndrome.....12

C. Oncogenes.....13

D. Triplet Repeats.....15

E. RNA Quadruplexes.....16

F. Quadruplexes as Biosensors.....17

1.4 References.....18

UCSF LIBRARY

CHAPTER 2: Engineering The Quadruplex Fold: Nucleoside Conformation Determines Both Folding Topology And Molecularity In Guanine Quadruplexes 31

2.1 ABSTRACT 32

2.2 INTRODUCTION 33

2.3 MATERIALS AND METHODS 38

2.3.1 Sample Preparation 38

2.3.2 Circular Dichroism 38

2.3.3 UV Measurements 39

2.3.4 Nondenaturing Polyacrylamide Gel Electrophoresis (PAGE) 39

2.3.5 1D Imino Proton NMR 40

2.4 RESULTS 40

2.4.1 Circular Dichroism 40

2.4.2 Polyacrylamide Gel Electrophoresis 52

2.4.3 Thermal Denaturation Studies 54

2.4.4 Imino proton NMR spectra 60

2.5 DISCUSSION 61

2.6 OTHER SEQUENCES 71

2.6.1 CD results for dTETTEL and dTETTELr 72

2.6.2 CD results for dOTEL, rOTEL, and dOTELr 76

2.6.3 CD results of dHTRr1 and dHTRr2 83

2.7 REFERENCES 87

UCSF LIBRARY

CHAPTER 3: HEAT CAPACITY EFFECTS IN THE UNFOLDING OF THE THROMBIN BINDING APTAMER.....94

3.1 ABSTRACT.....95

3.2 INTRODUCTION.....96

3.3 MATERIALS AND METHODS.....97

3.3.1 Sample Preparations.....97

3.3.2 UV Measurements.....97

3.3.3 Circular Dichroism (CD).....97

3.3.4 Isothermal Titration Calorimetry (ITC).....98

3.4 RESULTS.....100

3.4.1 ITC of TBA into 25 mM KCl.....100

3.4.2 ITC of TBA into 140 mM KCl.....110

3.5 DISCUSSION.....117

3.6 CONCLUSIONS.....121

3.7 REFERENCES.....122

UCSF LIBRARY

CHAPTER 4: THERMODYNAMIC CHARACTERIZATION OF THE THROMBIN-BINDING DNA APTAMER 124

4.1 Abstract 125

4.2 introduction 126

4.3 Materials and methods 128

4.3.1 Sample Preparation 128

4.3.2 UV Measurement 128

4.3.3 Circular Dichroism (CD) 128

4.3.4 Isothermal Titration Calorimetry (ITC) 129

4.4 Results 130

4.5 Discussion 142

4.6 Reference 144

CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS 147

UCSF LIBRARY

LIST OF TABLES

CHAPTER 2

Table 1: ε260 values used for concentration determinations. 39

Table 2: Melting temperatures 58

Table 3: Thermodynamic parameters from van't Hoff analysis 60

CHAPTER 3

Table 1: Calculations of ΔC_p.....119

CHAPTER 4

Table 1: Thermodynamic parameters obtained from ITC experiments. 137

Table 2: K for binding of Ba^{2+} and TBA.....139

UCSF LIBRARY

LIST OF FIGURES

CHAPTER 1

Figure 1: Guanine quartet 2

Figure 2: Schematics for various quadruplex structures 4

Figure 3: Comparisons of unimolecular quadruplexes 5

Figure 4: Possible parallel quadruplexes 6

CHAPTER 2

Figure 1: HTR structure 35

Figure 2: Schematic of different quadruplex folding topologies for HTR 36

Figure 3: *anti* and *syn* conformation 36

Figure 4: Schematic of dTBA antiparallel unimolecular quadruplex structure 37

Figure 5: CD spectra for various sequences 42

Figure 6: CD spectra for dTBAr1 45

Figure 7: CD spectra for dTBA 46

Figure 8: CD spectra for rTBA 47

Figure 9: CD spectra for dTBAr1 48

Figure 10: CD spectra for various sequences 51

Figure 11: Non-denaturing polyacrylamide gel 53

Figure 12: T_{max} as a function of strand concentration 56

Figure 13: Quartet imino proton spectra 61

Figure 14: CD spectra for dtettel 73

Figure 15: CD spectra for dtettelr 74

UCSF LIBRARY

Figure 16: CD spectra for dtettelr 75

Figure 17: CD spectra for dOTEL. 77

Figure 18: CD spectra for dOTEL 78

Figure 19: CD spectra for rOTEL 79

Figure 20: CD spectra for rOTEL 80

Figure 21: CD spectra for dOTELr 81

Figure 22: CD spectra for dOTELr 82

Figure 23: CD spectra for dHTRr1 84

Figure 24: CD spectra for dHTRr2 85

CHAPTER 3

Figure 1 UV melting curves for TBA in 25 mM KCl 101

Figure 2: CD spectra of TBA in 25 mM KCl 103

Figure 3: Representative ITC data of TBA in 25 mM KCl 104

Figure 4: ΔH for vs. titration step of TBA in 25 mM KCl 106

Figure 5: ΔH vs. temperature plot of TBA in 25 mM KCl 108

Figure 6: UV melting curves of TBA in 140 mM KCl 111

Figure 7: CD spectra of TBA in 140 mM KCl 112

Figure 8: Representative ITC data of TBA in 140 mM KCl 114

Figure 9: ΔH for vs. titration step of TBA in 140 mM KCl 115

Figure 10: ΔH vs. temperature plot of TBA in 140 mM KCl 116

UCSF LIBRARY

CHAPTER 4

Figure 1: CD spectra of TBA with the addition of barium acetate 131

Figure 2: CD titration of TBA with barium acetate at 10 °C 132

Figure 3: CD titration of TBA with barium acetate at 20 °C 133

Figure 4: CD titration of TBA with barium acetate at 30 °C 134

Figure 5: Representative ITC data of barium acetate titration..... 135

Figure 6: Representative ITC data of lead nitrate titration..... 136

Figure 7: ΔH vs. temperature plot for TBA-barium quadruplex 138

Figure 8: ΔH vs. temperature plot for TBA-lead quadruplex..... 139

Figure 9: van 't Hoff plot..... 140

UCSF LIBRARY

CHAPTER 1: INTRODUCTION

UCSF LIBRARY

1.1 WHAT IS QUADRUPLEX?

It has been known for a long time that tandemly repeated nucleic acid sequences with two or more guanine bases have the propensity to form polymorphic quadruplexes. The basic building blocks that construct these quadruplexes are the guanine quartets stabilized by cyclic Hoogsteen hydrogen bonding as shown in Figure 1 [1].

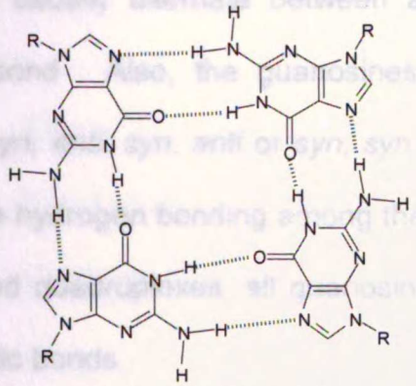

Figure 1: Guanine quartet

1.1.1 Structural Polymorphism

The structural polymorphism of guanine quadruplexes can be classified in several ways [2]. They can be classified in terms of the number of strands that self-associate to form the structure. One strand of oligonucleotides containing at least four runs of guanines bases can fold into an unimolecular quadruplex (Figure 2D&E), while two strands with more than 2 runs of guanine bases can fold and associate into a hairpin dimer (Figure 2B&C). The linear four-stranded

UCSF LIBRARY

(Figure 2A) quadruplex can also form. Quadruplexes can also be grouped into parallel or antiparallel classes by the arrangement and orientation of the strands. These structures can be further distinguished by the *anti* and *syn* conformations of the guanosines about the glycosidic bonds, and by the orientation of the loops to be lateral or diagonal.

In folded quadruplexes, which include the unimolecular / single stranded (Figure 3) and the hairpin dimer / bimolecular structures, the sequential guanosines usually alternate between *anti* and *syn* conformations about the glycosidic bond . Also, the guanosines within each quartet alternate in the *sequence syn, anti, syn, anti* or *syn, syn, anti, anti*. This is important in order to maintain the hydrogen bonding among the bases within the quartets. In the linear, four-stranded quadruplexes, all guanosines are in the *anti* conformations about the glycosidic bonds.

UCSF LIBRARY

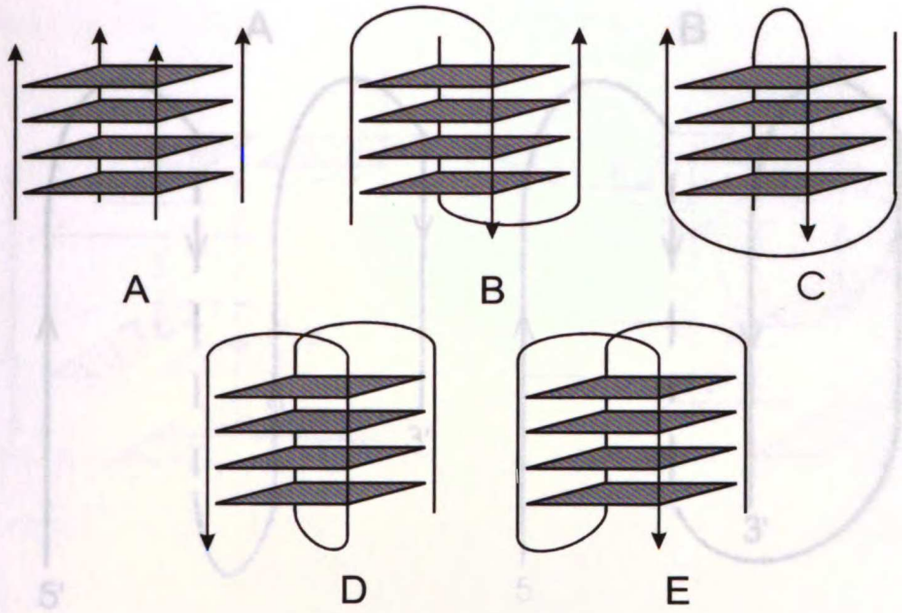

Figure 2: Schematics for various quadruplex structures. A: Linear four-stranded. B: Hairpin dimer. C: Alternative hairpin dimer. D: Unimolecular quadruplex. E: Alternative unimolecular quadruplex.

UCSF LIBRARY

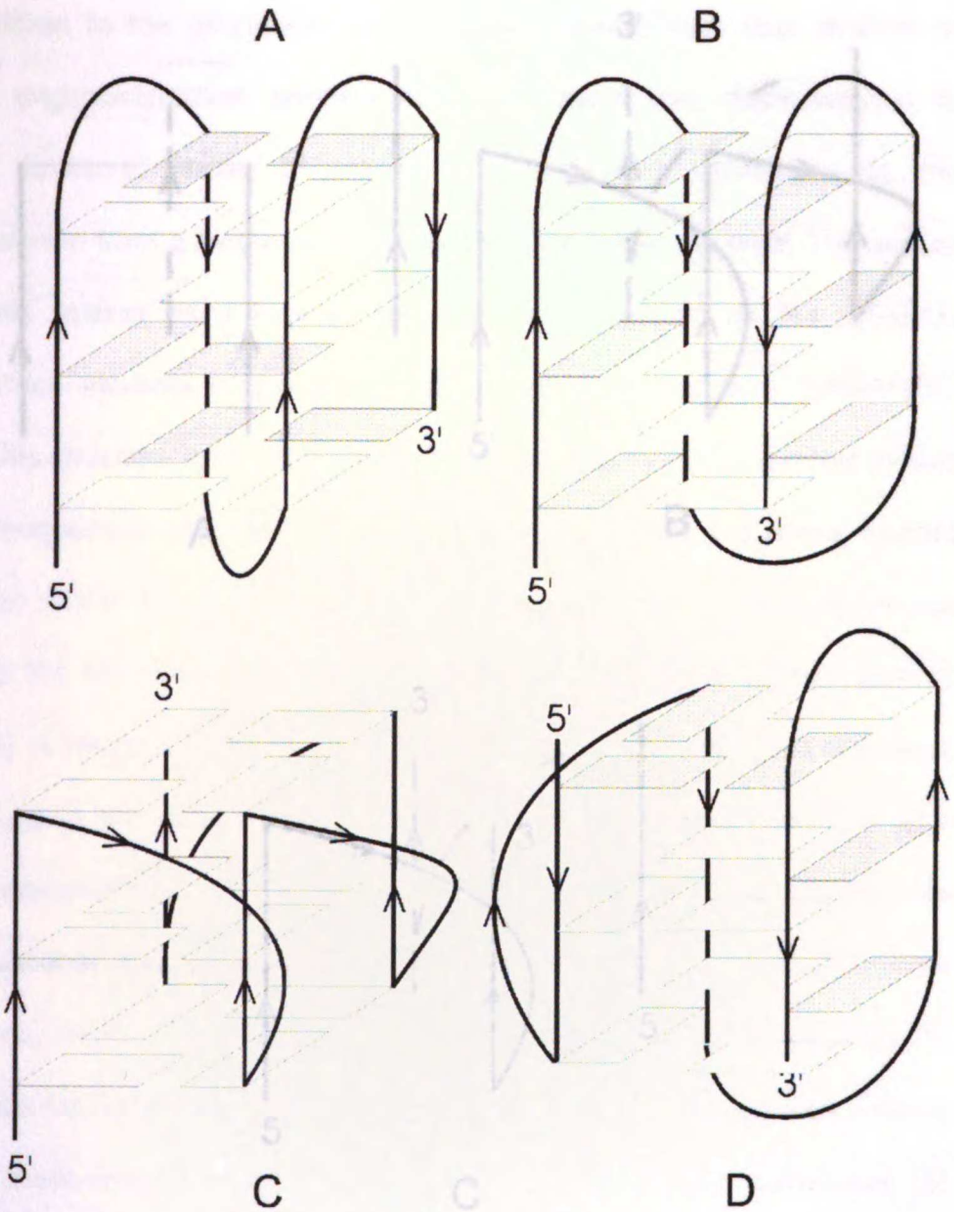

Figure 3: Comparisons of unimolecular quadruplexes. A: Antiparallel folding pattern. B: Alternative antiparallel folding. C: Possible unimolecular parallel folding topology. D: Possible folding containing three parallel guanine stretches and one antiparallel stretch.

UCSF LIBRARY

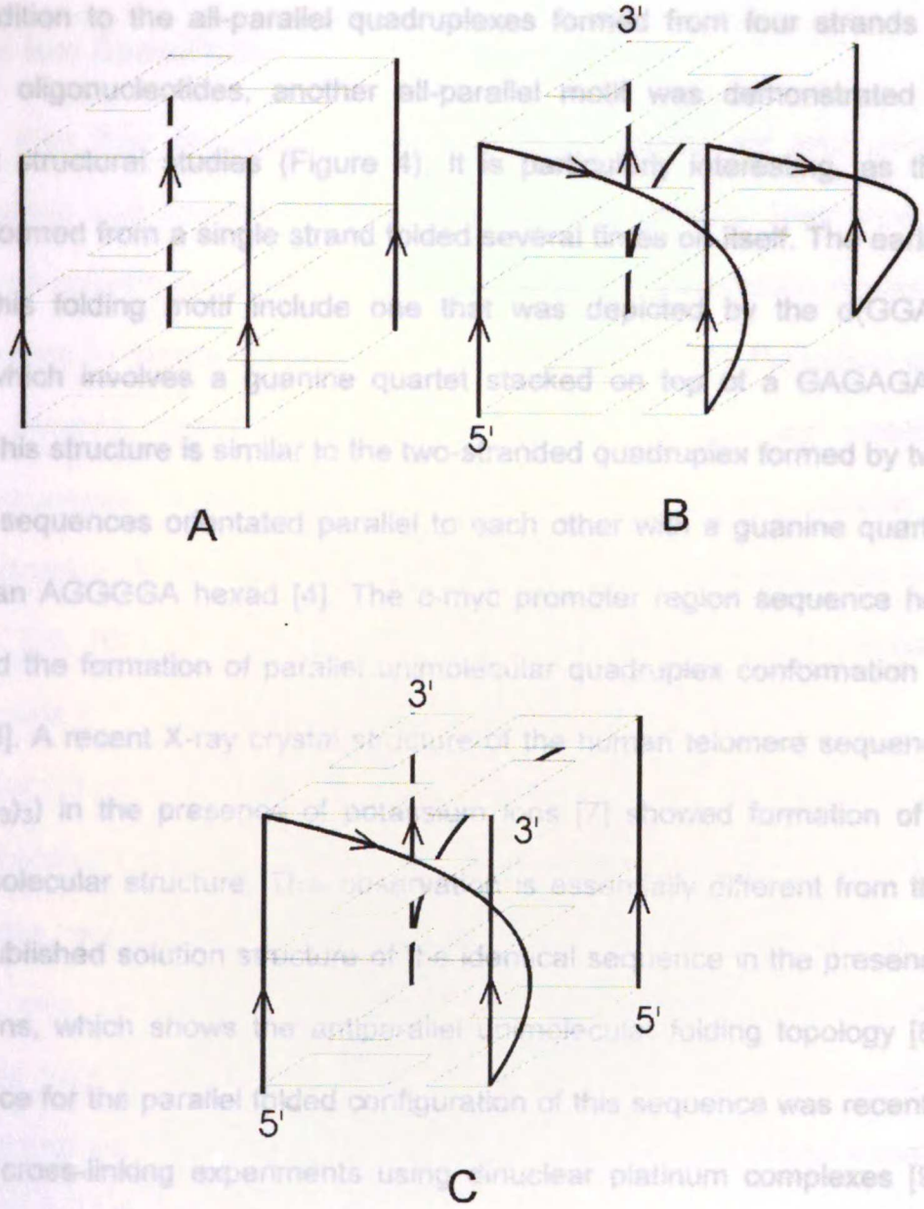

Figure 4: Possible parallel quadruplexes. A: Four stranded; B: Unimolecular; C. Double stranded or bimolecular

UCSF LIBRARY

In addition to the all-parallel quadruplexes formed from four strands of guanine-rich oligonucleotides, another all-parallel motif was demonstrated in more recent structural studies (Figure 4). It is particularly interesting, as this structure is formed from a single strand folded several times on itself. The earlier studies of this folding motif include one that was depicted by the $d(GGA)_4$ sequence, which involves a guanine quartet stacked on top of a GAGAGAG heptad [3]. This structure is similar to the two-stranded quadruplex formed by two $d((GGA)_2G)$ sequences orientated parallel to each other with a guanine quartet stacked on an AGGGGA hexad [4]. The c-myc promoter region sequence has also revealed the formation of parallel unimolecular quadruplex conformation in solution [5, 6]. A recent X-ray crystal structure of the human telomere sequence $d(AG_3(TTAG_3)_3)$ in the presence of potassium ions [7] showed formation of a parallel unimolecular structure. This observation is essentially different from the previously published solution structure of the identical sequence in the presence of sodium ions, which shows the antiparallel unimolecular folding topology [8]. Other evidence for the parallel folded configuration of this sequence was recently provided by cross-linking experiments using dinuclear platinum complexes [9]. The most recent structure of HTR [10] determined by NMR in K+ solution demonstrates a novel unimolecular quadruplex folding topology with hybrid-type mixed parallel / antiparallel guanine strands (Figure 3D).

UCSF LIBRARY

1.1.2 Metal Ion Specificity

Other than their sequence dependence, quadruplexes are also unique among all nucleic acid structures in their metal ion requirements. Their unusual dependence on specific metal ions for their formation and stabilization results in significant binding by means of inner sphere coordination [11-13]. Potassium and sodium are common metals ions known to effectively induce and stabilize quadruplex structures. In additions, there are many other cations capable of inducing quadruplex structures [14]. A few of these cations are monovalent ions such as NH_4^+ [15, 16] and Tl^+ [17]. Divalent cations known to induce quadruplex structures include Sr^{2+} [18], Ba^{2+} [19], and Pb^{2+} [12].

Sen and Gilbert investigated the roles of metal ions in the formation of G4-DNA (four-stranded structures) with the five alkali metal ions [20]. These alkali metal ions have very different efficacies in inducing the formation of quadruplex structures. It was demonstrated that K^+, in millimolar concentrations, is much more effective than other alkali metal ions in promoting quadruplex structure formation and that the K^+-bound quadruplexes also possess the greatest thermal stability. Li^+ on the other hand, was ineffective in the formation and stabilization of quadruplex structures. In general, the trend of the alkali metal cation-induced stabilization of G4-DNA is

$K^+ > Rb^+ > Na^+ > Cs^+ > Li^+$.

One of the possible explanations of this order of efficacies is that the central cavity of the guanine quartet displays crown ether-like size-selective binding of cations [21].

UCSF LIBRARY

The intracellular K^+ concentration is about 140 mM and Na^+ concentration is in the lower range of 5-15 mM. These phenomena are reversed in the extracellular environment [22]. Under both conditions, the biological system is a medium conducive to quadruplex structure formation.

Another phenomena termed the Na^+ - K^+ switch, which was first introduced by Sen and Gilbert [20], regarding the variation of different predominant structures of quadruplexes in the presence of both Na^+ and K^+ [23, 24]. The ability of these cations to stabilize linear vs. folded quadruplexes differs significantly for oligonucleotides containing two or more runs of guanines. Therefore, when both of these alkali cations are present, it is hard to predict the ratio of linear and folded quadruplexes that exists at equilibrium. This is further complicated by differences in the kinetics of formation of these forms. Linear / four-stranded quadruplexes can have very slow formation kinetics [25], whereas unimolecular quadruplexes fold rather rapidly [26, 27]. Hence, it may take very long times to reaching equilibrium, or there may even possibly be kinetic blocks, to reaching equilibrium between these types of quadruplexes.

1.2 HISTORY

The first demonstration of the ability of guanine to form a guanine quartet was made more than forty years ago using deoxyguanosine 3'-phosphate [28]. This phenomenon was also studied with 5'- and 3'-GMP as well as polyinosine and polyguanine [14, 28-32]. GMP was reported to possesses the unique ability to undergo spontaneous formation of this ordered quartet structure [33, 34].

UCSF LIBRARY

1.3 BIOLOGICAL SIGNIFICANCE

Evidence for the presence of quadruplex structures in biological systems is accumulating at an increasing rate. One of the most convincing pieces of evidence for the presence of quadruplex structures in vivo is the demonstration of antibody recognition of the structures in intact nuclei [35]. High-affinity single-chain antibody fragments (scFv) with capabilities of discerning parallel and antiparallel quadruplexes were generated to study the presence of quadruplexes in the macronucleus and micronucleus of Stylonychia lemnae. Experimental evidence showed that the telomeres of Stylonychia macronuclei adopt a quadruplex structure in vivo, suggesting the possibility of a role for this structure in telomere function. More recently, Hurley and co-workers presented direct evidence for quadruplex structures in the c-myc promoter region of human cells. There are many identified proteins that specifically recognize and bind qudruplexes. For example, the products of the Bloom's and Werner's syndrome genes, BLM and WRN, are DNA helicases that unwind a variety of different forms of guanine-quadruplex DNA [36]. With the cellular presence of quadruplexes having been demonstrated, significant biological roles of quadruplexes are highly likely. There are abundant quadruplex-forming motifs widely dispersed in biologically significant regions in the eukaryotic genome. There are many reports of proposing roles of quadruplex structures in replication, transcription, and recombination. Studies have demonstrated their roles in telomere function as well as their association with genetic based disorders in triplet repeat sequences.

UCSF LIBRARY

1.3.1 Possible Biological Roles

A. Telomere and telomerase

The tandem repeat G-rich telomeric sequences, which readily form quadruplex structures, have been extensively studied over the years. The major hypothesis concerning the significance of telomerase activity is the requirement for telomerase in cells that divide indefinitely, making telomerase inhibition a primary chemotherapy target [37-42]. Another significant hypothesis is that loss of telomerase activity is a cause of cellular senescence [43-45] and hence being able to prolong telomerase activity will allow cells to live longer.

During DNA replication, linear chromosomes are shortened and telomerase prevents the shortening process. It is found that tumor cells divide indefinitely and that 85-90% of all tumor cells have high levels of telomerase activity [46-48]. Therefore, an abundance of research has been focusing on inhibiting telomerase for cancer therapy [49-51].

There is a strand bias in the various telomeric DNAs in which the strand running 5' to 3' from the center toward the end of the DNA molecule has more guanine residues than its complement. Furthermore, the guanines in this strand have tandemly repeated short, simple sequence motifs. This guanine-rich nature and the extension of the 3' strand are characteristics of telomeres and may be general features of telomere structures [52]. These telomeric sequences were found to form novel guanine-quadruplex structures and it was suggested that these structures might promote various aspects of telomere functions, including regulation of telomerase or stabilization of chromosomes [53, 54].

UCSF LIBRARY

Telomerase is a DNA polymerase, which is also a reverse transcriptase because of its mechanism of action involves the copying an RNA template into DNA. It is a ribonucleoportein containing its own RNA template as an integral part of the enzyme [55, 56]. Telomerase has been shown to bind preferentially to single-stranded telomeric oligonucleotides. The formation of quadruplexes was shown to influence the extent of telomere elongation by telomerase, and hence might act as a negative regulator of telomerase activity [57]. Quadruplex structures are therefore become promising targets for therapeutic interventions [41, 42, 58-60]. Small molecules such as Porphyrin T4 [61-63], which binds specifically to quadruplex structures, can push the equilibrium toward this folded form, providing a credible approach to telomerase inhibition.

B. Bloom's and Werner's Syndrome

Bloom's syndrome is a rare genetic disorder characterized by hypermutable cells; this leads to many clinical pathologies such as skin lesions and increased occurrences of bacterial infections due to immunodeficiency [64]. Complications arising from this disorder include the formidable chronic lung disease, diabetes, leukemia, and many other types of cancers [64-66]. Bloom's syndrome is caused by defects in the BLM gene, which encodes a DNA helicase. This ATP dependent helicase was reported to be significantly more active at unwinding linear DNA quadruplexes than duplexes [67]. This preference suggested that linear DNA quadruplex may be a natural substrate of BLM in vivo [67].

UCSF LIBRARY

As in the genotype of Bloom's syndrome, Werner's syndrome is also caused by defects in an ATP dependent helicase encoded by the WRN gene [68]. This genetic disorder is associated with genetic instability and early onset of aging including gray hair, muscle atrophy, osteoporosis, cataracts, diabetes mellitus, hyperlipidemia, malignancy, and senile dementia [65, 66, 68]. WRN was shown to unwind certain quadruplexes in the presence of hydrolysable ATP and Mg^{2+} [69]. WRN efficiently unwinds hairpin dimer quadruplexes formed by $d(CGG)_7$ with the presence of a 3'- or 5'-single-strand tail. However, WRN is unable to unwind the corresponding blunt-ended quadruplex, as well as hairpin dimer quadruplexes of a telomeric sequence or four-stranded linear quadruplexes of an IgG switch region sequence. This selective ability of WRN to unwind specific quadruplexes may suggest a specific role of this helicase in DNA metabolism [69].

BLM and WRN are classified as RecQ helicases, which catalyze the ATP dependent unwinding of quadruplex DNA. Hence, this unwinding function of these helicases may be critical during replication and transcription.

C. Oncogenes

Quadruplex structures are also of potential biological significance as regulatory elements as demonstrated by their formation in the G-rich sequence of the c-myc oncogene control region [70]. c-myc has emerged as an attractive anti-cancer therapeutic target over the past few years because its expression has been linked to cellular proliferation as well as inhibition of differentiation associated with several human and animal malignancies, including breast, colon,

UCSF LIBRARY

cervix, and lung cancer [71-73]. There is also evidence indicating that brief inactivation of c-myc results in apoptosis of tumors and osteogenic sarcoma [74]. The nuclease hypersensitive element (NHE) III_1 of the c-myc promoter, a 27-base-pair guanine-rich sequence situated in the upstream from the P1 promoter, has been shown to form intramolecular quadruplex structures under physiological conditions [70, 75]. NHE III_1 controls 85-90% of the transcriptional activation of the c-myc oncogene, the stabilization of the quadruplex structures formed in this sequence by the cationic porphyrin TMPyP4 was shown to suppress further c-myc transcriptional activation [75]. Not only do the c-myc promoter regions contain the quadruplex-forming sequence, but other similar sequences are also found in the promoter regions of various other oncogenes [70]. These observations provide compelling evidence that quadruplexes may play a role in gene regulation as a transcriptional repressor element.

The human B-cell CLL/lymphoma2 (bcl-2) gene is a proto-oncogene found to be overexpressed in a wide range of human cancers including B-cell and T-cell lymphomas, breast, cervical, lung, prostate, and colorectal cancers [76-81]. Unlike other proto-oncogenes, the bcl-2 gene product does not promote cell proliferation but instead its oncogenic potential is achieved by reducing the rate of cell apoptosis. Hence, overexpression of bcl-2 was shown to interfere with the therapeutic action of current cell treatment regimes, which involve inducing apoptosis of tumor cells [82].

Bcl-2 has also emerged as an attractive target for neuroprotective therapy, which acts by preventing apoptotic neuronal cell death in neurodegenerative

UCSF LIBRARY

disorders of aging, including Alzheimer's disease, Parkinson's disease, and stroke. Overexpression of bcl-2 in neuronal cells has been shown to prevent cell death induced by diverse apoptotic stimuli [83, 84].

The P1 promoter regions upstream of the translation start site is a guanine-rich transcriptional promoter region containing multiple transcriptional start sites [85]. This guanine-rich strand can form intramolecular quadruplexes in solution containing K^+ [85]. These results indicate that the quadruplexes identified in the promoter region of the bcl-2 gene are likely to play a role similar to those in the c-myc promoter, in that their formation could serve to modulate gene transcription [86].

D. Triplet Repeats

Triplet repeats have been associated with a variety of human, genetic based disorders [87-89]. The molecular basis of the triplet repeat diseases is expansion, of the number of repeats of a triplet sequence, which can occur in a single generation. Triplet repeat diseases exhibit non-Mendelian inheritance, with genetic characteristic known as anticipation, a term referring to the tendency towards earlier onset and increasing severity with the degree of expansion [90].

The most extensively studied case is the fragile X syndrome, a result of the expansion of the d(CGG) repeat within the human fragile X mental retardation (FMR1) gene [91] and the loss of fragile X mental retardation protein (FMRP) function [92]. Studies show that while d(CGG) is repeated about 5-50 times in normal individuals, it is repeated about 40-230 times in carriers and from about 230 to more than 2000 times in diseased individuals [93]. Although multiple

UCSF LIBRARY

forms of DNA structure have been proposed for these d(CGG) repeats [89], there is substantial evidence that d(CGG) repeats can form quadruplex structures under physiological conditions [91, 94].

E. RNA Quadruplexes

There are several examples of guanine-rich sequences appearing in mRNA that have the capability of forming quadruplexes. RNA quadruplexes may be involved in fragile X syndrome as several reports describe the high affinity of the FMRP protein for quadruplex motifs in mRNA [92, 95-98]. Targeting various mRNA sequences may be a significant function of FMRP, acting as a translational repressor. There may also be a protein-RNA quadruplex interaction involved in transporting the bound mRNA to certain cellular sites. Other studies have also shown that the DNA region containing tandem GGC repeats may form unimolecular quadruplexes that might involved in the mechanism underlying the expansion process [99, 100].

Sequence analysis of the 3'-untranslated region of the insulin growth factor (IGF) II mRNA suggests the presence of guanine quadruplex structures in the guanine-rich region downstream of the cleavage site for endonucleolytic cleavage [101]. It was proposed that the presence of quadruplexes flanking the site may serve to keep the intervening sequence containing the cleavage site available for interactions involved in the posttranscriptional processing events at that site [101].

Studies have also described a translational recoding event in a herpes simplex virus mutant that results in a low level of expression of full-length

UCSF LIBRARY

thymidine kinase, despite the presence of an extra base in the corresponding mutant gene. There is evidence showing that the underlying mechanism of this downstream frameshift is based on the ability to form guanine quadruplex structures [102, 103].

RNA quadruplexes are also involved in the preferential binding of the fd gene 5 protein to its linear form [104, 105]. The resulting RNA-protein complex may be involved in translational repression. Other studies also demonstrated a more than 100 fold greater binding affinity of this gene 5 protein for quadruplex DNA than for the single-stranded DNA [106].

Although RNA aptamers, identified by in vitro selection with quadruplex-folding motif, are relatively rare, it has been shown that the essential recognition element for the prion protein was an RNA quadruplex [107].

F. Quadruplexes as Biosensors

Thrombin binding aptamer (TBA), identified by in vitro selection techniques in an effort to generate tight-binding ligands for the protin thrombin, is the most extensively studied quadruplex-forming sequence [27, 108]. TBA has been exploited in other ways, using its protein recognition cabilities as diagnostic tools in microarray format and in an immunobead assay [109, 110]. In addition, aptamer beacons, based on TBA, have been described as a sensitive tool for protein recognition [111].

UCSF LIBRARY

1.4 REFERENCES

1. Williamson, J.R., *G-quartet structures in telomeric DNA.* Annu Rev Biophys Biomol Struct, 1994. **23**: p. 703-30.
2. Keniry, M.A., *Quadruplex structures in nucleic acids.* Biopolymers, 2000. **56**(3): p. 123-46.
3. Matsugami, A., et al., *An intramolecular quadruplex of (GGA)(4) triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction.* J Mol Biol, 2001. **313**(2): p. 255-69.
4. Kettani, A., et al., *A dimeric DNA interface stabilized by stacked A.(G.G.G.G).A hexads and coordinated monovalent cations.* J Mol Biol, 2000. **297**(3): p. 627-44.
5. Seenisamy, J., et al., *The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4.* J Am Chem Soc, 2004. **126**(28): p. 8702-9.
6. Phan, A.T., Y.S. Modi, and D.J. Patel, *Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter.* J Am Chem Soc, 2004. **126**(28): p. 8710-6.
7. Parkinson, G.N., M.P. Lee, and S. Neidle, *Crystal structure of parallel quadruplexes from human telomeric DNA.* Nature, 2002. **417**(6891): p. 876-80.
8. Wang, Y. and D.J. Patel, *Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex.* Structure, 1993. **1**(4): p. 263-82.

UCSF LIBRARY

9. Ourliac-Garnier, I., et al., *Cross-links of quadruplex structures from human telomeric DNA by dinuclear platinum complexes show the flexibility of both structures.* Biochemistry, 2005. **44**(31): p. 10620-34.

10. Ambrus, A., et al., *Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution.* Nucleic Acids Res, 2006. **34**(9): p. 2723-35.

11. Phillips, K., et al., *The crystal structure of a parallel-stranded guanine tetraplex at 0.95 A resolution.* J Mol Biol, 1997. **273**(1): p. 171-82.

12. Smirnov, I. and R.H. Shafer, *Lead is unusually effective in sequence-specific folding of DNA.* J Mol Biol, 2000. **296**(1): p. 1-5.

13. Smirnov, I.V., et al., *Pb EXAFS studies on DNA quadruplexes: identification of metal ion binding site.* Biochemistry, 2002. **41**(40): p. 12133-9.

14. Guschlbauer, W., J.F. Chantot, and D. Thiele, *Four-stranded nucleic acid structures 25 years later: from guanosine gels to telomer DNA.* J Biomol Struct Dyn, 1990. **8**(3): p. 491-511.

15. Nagesh, N. and D. Chatterji, *Ammonium ion at low concentration stabilizes the G-quadruplex formation by telomeric sequence.* J Biochem Biophys Methods, 1995. **30**(1): p. 1-8.

16. Hud, N.V., Schultze, P., Feigon, J., *Ammonium ion as an NMR probe for monovalent cation coordination sites of DNA quadruplexes.* J Am Chem Soc, 1998. **120**(25): p. 6403.

UCSF LIBRARY

17. Basu, S.S., A. A.; Cocco, M.; Strobel, S. A., *Direct detection of monovalent metal ion binding to a DNA G-quartet by ^{205}Tl NMR.* J Am Chem Soc, 2000. **122**(13): p. 3240-3241.

18. Chen, F.M., *Sr2+ facilitates intermolecular G-quadruplex formation of telomeric sequences.* Biochemistry, 1992. **31**(15): p. 3769-76.

19. Venczel, E.A. and D. Sen, *Parallel and antiparallel G-DNA structures from a complex telomeric sequence.* Biochemistry, 1993. **32**(24): p. 6220-8.

20. Sen, D. and W. Gilbert, *A sodium-potassium switch in the formation of four-stranded G4-DNA.* Nature, 1990. **344**(6265): p. 410-4.

21. Frensdorff, H.K., *Stability constants of cyclic polyether complexes with univalent cations.* J Am Chem Soc, 1971. **93**(3): p. 600-606.

22. Alberts, B., *Molecular biology of the cell.* 3rd ed. 1994, New York: Garland.

23. Hardin, C.C., et al., *Monovalent cation induced structural transitions in telomeric DNAs: G-DNA folding intermediates.* Biochemistry, 1991. **30**(18): p. 4460-72.

24. Miura, T. and G.J. Thomas, Jr., *Structural polymorphism of telomere DNA: interquadruplex and duplex-quadruplex conversions probed by Raman spectroscopy.* Biochemistry, 1994. **33**(25): p. 7848-56.

25. Wyatt, J.R., P.W. Davis, and S.M. Freier, *Kinetics of G-quartet-mediated tetramer formation.* Biochemistry, 1996. **35**(24): p. 8002-8.

26. Mergny, J.L., A.T. Phan, and L. Lacroix, *Following G-quartet formation by UV-spectroscopy.* FEBS Lett, 1998. **435**(1): p. 74-8.

UCSF LIBRARY

27. Smirnov, I. and R.H. Shafer, *Effect of loop sequence and size on DNA aptamer stability.* Biochemistry, 2000. **39**(6): p. 1462-8.

28. Gellert, M., M.N. Lipsett, and D.R. Davies, *Helix formation by guanylic acid.* Proc. Natl. Acad. Sci. USA, 1962. **48**: p. 2013-2018.

29. Tougard, P., J.F. Chantot, and W. Guschlbauer, *Nucleoside conformations. X. An X-ray fiber diffraction study of the gels of guanine nucleosides.* Biochim Biophys Acta, 1973. **308**(7): p. 9-16.

30. Sasisekharan, V., S. Zimmerman, and D.R. Davies, *The structure of helical 5'-guanosine monophosphate.* J Mol Biol, 1975. **92**(2): p. 171-9.

31. Arnott, S. and E. Selsing, *Structures for the polynucleotide complexes poly(dA) with poly (dT) and poly(dT) with poly(dA) with poly (dT).* J Mol Biol, 1974. **88**(2): p. 509-21.

32. Zimmerman, S.B., G.H. Cohen, and D.R. Davies, *X-ray fiber diffraction and model-building study of polyguanylic acid and polyinosinic acid.* J Mol Biol, 1975. **92**(2): p. 181-92.

33. Pinnavaia, T.J., H.T. Miles, and E.D. Becker, *Letter: Self-assembled 5'-guanosine monophosphate. Nuclear magnetic resonance evidence for a regular, ordered structure and slow chemical exchange.* J Am Chem Soc, 1975. **97**(24): p. 7198-200.

34. Chantot, J. and W. Guschlbauer, *Physicochemical properties of nucleosides 3. Gel formation by 8- bromoguanosine.* FEBS Lett, 1969. **4**(3): p. 173-176.

UCSF LIBRARY

35. Schaffitzel, C., et al., *In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei.* Proc Natl Acad Sci U S A, 2001. **98**(15): p. 8572-7.

36. Mohaghegh, P., et al., *The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases.* Nucleic Acids Res, 2001. **29**(13): p. 2843-9.

37. Hurley, L.H., et al., *G-quadruplexes as targets for drug design.* Pharmacol Ther, 2000. **85**(3): p. 141-58.

38. Jenkins, T.C., *Targeting multi-stranded DNA structures.* Curr Med Chem, 2000. **7**(1): p. 99-115.

39. Buys, C.H., *Telomeres, telomerase, and cancer.* N Engl J Med, 2000. **342**(17): p. 1282-3.

40. de Lange, T. and T. Jacks, *For better or worse? Telomerase inhibition and cancer.* Cell, 1999. **98**(3): p. 273-5.

41. Neidle, S. and L.R. Kelland, *Telomerase as an anti-cancer target: current status and future prospects.* Anticancer Drug Des, 1999. **14**(4): p. 341-7.

42. Mergny, J.L., et al., *The development of telomerase inhibitors: the G-quartet approach.* Anticancer Drug Des, 1999. **14**(4): p. 327-39.

43. Epel, E.S., et al., *Cell aging in relation to stress arousal and cardiovascular disease risk factors.* Psychoneuroendocrinology, 2006. **31**(3): p. 277-87.

Serrano, A.L. and V. Andres, *Telomeres and cardiovascular disease: does size matter?* Circ Res, 2004. **94**(5): p. 575-84.

UCSF LIBRARY

45. Blackburn, E.H., *Telomere states and cell fates.* Nature, 2000. **408**(6808): p. 53-6.

46. Wu, A., M. Ichihashi, and M. Ueda, *Correlation of the expression of human telomerase subunits with telomerase activity in normal skin and skin tumors.* Cancer, 1999. **86**(10): p. 2038-44.

47. Fujioka, T., et al., *Telomerase activity in human renal cell carcinoma.* Int J Urol, 2000. **7**(1): p. 16-21.

48. Shimada, M., et al., *The role of telomerase activity in hepatocellular carcinoma.* Am J Gastroenterol, 2000. **95**(3): p. 748-52.

49. Raymond, E., et al., *Agents that target telomerase and telomeres.* Curr Opin Biotechnol, 1996. **7**(6): p. 583-91.

50. Hahn, W.C., et al., *Inhibition of telomerase limits the growth of human cancer cells.* Nat Med, 1999. **5**(10): p. 1164-70.

51. Zumstein, L.A. and V. Lundblad, *Telomeres: has cancer's Achilles' heel been exposed?* Nat Med, 1999. **5**(10): p. 1129-30.

52. Henderson, E.R. and E.H. Blackburn, *An overhanging 3' terminus is a conserved feature of telomeres.* Mol Cell Biol, 1989. **9**(1): p. 345-8.

53. Henderson, E., et al., *Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs.* Cell, 1987. **51**(6): p. 899-908.

54. Blackburn, E.H., *Structure and function of telomeres.* Nature, 1991. **350**(6319): p. 569-73.

UCSF LIBRARY

55. Blackburn, E.H., *The telomere and telomerase: nucleic acid-protein complexes acting in a telomere homeostasis system. A review.* Biochemistry (Mosc), 1997. **62**(11): p. 1196-201.

56. Blackburn, E.H. and C.W. Greider, eds. *Telomeres*. Cold Springs Harbor Monograph. 1995, Cold Spring Harbor Laboratory Press: Plainview, New York.

57. Zahler, A.M., et al., *Inhibition of telomerase by G-quartet DNA structures.* Nature, 1991. **350**(6320): p. 718-20.

58. Autexier, C., *Telomerase as a possible target for anticancer therapy.* Chem Biol, 1999. **6**(11): p. R299-303.

59. Han, H. and L.H. Hurley, *G-quadruplex DNA: a potential target for anti-cancer drug design.* Trends Pharmacol Sci, 2000. **21**(4): p. 136-42.

60. Kerwin, S.M., *G-Quadruplex DNA as a target for drug design.* Curr Pharm Des, 2000. **6**(4): p. 441-78.

61. Fedoroff, O.Y., et al., *NMR-Based model of a telomerase-inhibiting compound bound to G-quadruplex DNA.* Biochemistry, 1998. **37**(36): p. 12367-74.

62. Arthanari, H., et al., *Fluorescent dyes specific for quadruplex DNA.* Nucleic Acids Res, 1998. **26**(16): p. 3724-8.

63. Anantha, N.V., M. Azam, and R.D. Sheardy, *Porphyrin binding to quadrupled T4G4.* Biochemistry, 1998. **37**(9): p. 2709-14.

64. German, J., *Bloom's syndrome.* Dermatologic Clinics, 1995. **13**: p. 7-18.

65. Chakraverty, R.K. and I.D. Hickson, *Defending genome integrity during DNA replication: a proposed role for RecQ family helicases.* Bioessays, 1999. **21**(4): p. 286-94.

66. Auerbach, A.D. and P.C. Verlander, *Disorders of DNA replication and repair.* Curr Opin Pediatr, 1997. **9**(6): p. 600-16.

67. Sun, H., et al., *The Bloom's syndrome helicase unwinds G4 DNA.* Journal of Biological Chemistry, 1998. **273**(42): p. 27587-92.

68. Goto, M., *Werner's syndrome: from clinics to genetics.* Clin Exp Rheumatol, 2000. **18**(6): p. 760-6.

69. Fry, M. and L.A. Loeb, *Human werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence d(CGG)n.* J Biol Chem, 1999. **274**(18): p. 12797-802.

70. Simonsson, T., P. Pecinka, and M. Kubista, *DNA tetraplex formation in the control region of c-myc.* Nucleic Acids Res, 1998. **26**(5): p. 1167-72.

71. Facchini, L.M. and L.Z. Penn, *The molecular role of Myc in growth and transformation: recent discoveries lead to new insights.* Faseb J, 1998. **12**(9): p. 633-51.

72. Marcu, K.B., S.A. Bossone, and A.J. Patel, *myc function and regulation.* Annu Rev Biochem, 1992. **61**: p. 809-60.

73. Spencer, C.A. and M. Groudine, *Control of c-myc regulation in normal and neoplastic cells.* Adv Cancer Res, 1991. **56**: p. 1-48.

74. Jain, M., et al., *Sustained loss of a neoplastic phenotype by brief inactivation of MYC.* Science, 2002. **297**(5578): p. 102-4.

75. Siddiqui-Jain, A., et al., *Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription.* Proc Natl Acad Sci U S A, 2002. **99**(18): p. 11593-8.

76. Akagi, T., E. Kondo, and T. Yoshino, *Expression of Bcl-2 protein and Bcl-2 mRNA in normal and neoplastic lymphoid tissues.* Leuk Lymphoma, 1994. **13**(1-2): p. 81-7.

77. Joensuu, H., L. Pylkkanen, and S. Toikkanen, *Bcl-2 protein expression and long-term survival in breast cancer.* Am J Pathol, 1994. **145**(5): p. 1191-8.

78. Tjalma, W., et al., *Expression of bcl-2 in invasive and in situ carcinoma of the uterine cervix.* Am J Obstet Gynecol, 1998. **178**(1 Pt 1): p. 113-7.

79. Pezzella, F., et al., *bcl-2 protein in non-small-cell lung carcinoma.* N Engl J Med, 1993. **329**(10): p. 690-4.

80. McDonnell, T.J., et al., *Expression of the protooncogene bcl-2 in the prostate and its association with emergence of androgen-independent prostate cancer.* Cancer Res, 1992. **52**(24): p. 6940-4.

81. Baretton, G.B., et al., *Apoptosis and immunohistochemical bcl-2 expression in colorectal adenomas and carcinomas. Aspects of carcinogenesis and prognostic significance.* Cancer, 1996. **77**(2): p. 255-64.

82. Reed, J.C., et al., *Regulation of chemoresistance by the bcl-2 oncoprotein in non-Hodgkin's lymphoma and lymphocytic leukemia cell lines.* Ann Oncol, 1994. **5 Suppl 1**: p. 61-5.

83. Middleton, G., G. Nunez, and A.M. Davies, *Bax promotes neuronal survival and antagonises the survival effects of neurotrophic factors.* Development, 1996. **122**(2): p. 695-701.

84. Allsopp, T.E., et al., *The proto-oncogene bcl-2 can selectively rescue neurotrophic factor-dependent neurons from apoptosis.* Cell, 1993. **73**(2): p. 295-307.

85. Dai, J., et al., *An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution.* J Am Chem Soc, 2006. **128**(4): p. 1096-8.

86. Dexheimer, T.S., D. Sun, and L.H. Hurley, *Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the bcl-2 P1 promoter.* J Am Chem Soc, 2006. **128**(16): p. 5404-15.

87. Wells, R.D., *Molecular basis of genetic instability of triplet repeats.* J Biol Chem, 1996. **271**(6): p. 2875-8.

88. Wells, R.D., S.T. Warren, and M. Sarmiento, *Genetic instabilities and hereditary neurological diseases.* 1998, San Diego, Calif.: Academic Press. xxiii, 829 , [10] of plates.

89. Sinden, R.R., *Biological implications of the DNA structures associated with disease- causing triplet repeats [In Process Citation].* Am J Hum Genet, 1999. **64**(2): p. 346-53.

90. Arthanari, H. and P.H. Bolton, *Functional and dysfunctional roles of quadruplex DNA in cells.* Chem Biol, 2001. **8**(3): p. 221-30.

91. Fry, M. and L.A. Loeb, *The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure.* Proc Natl Acad Sci U S A, 1994. **91**(11): p. 4950-4.

92. Brown, V., et al., *Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome.* Cell, 2001. **107**(4): p. 477-87.

93. Deissler, H., A. Behn-Krappa, and W. Doerfler, *Purification of nuclear proteins from human HeLa cells that bind specifically to the unstable tandem repeat (CGG)n in the human FMR1 gene.* J Biol Chem, 1996. **271**(8): p. 4327-34.

94. Kettani, A., R.A. Kumar, and D.J. Patel, *Solution structure of a DNA quadruplex containing the fragile X syndrome triplet repeat.* J Mol Biol, 1995. **254**(4): p. 638-56.

95. Moine, H. and J.L. Mandel, *Biomedicine. Do G quartets orchestrate fragile X pathology?* Science, 2001. **294**(5551): p. 2487-8.

96. Schaeffer, C., et al., *The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif.* Embo J, 2001. **20**(17): p. 4803-13.

97. Zhang, Y.Q., et al., *Drosophila fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function.* Cell, 2001. **107**(5): p. 591-603.

98. Darnell, J.C., et al., *Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function.* Cell, 2001. **107**(4): p. 489-99.

99. Usdin, K. and K.J. Woodford, *CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro.* Nucleic Acids Res, 1995. **23**(20): p. 4202-9.

100. Usdin, K. and E. Grabczyk, *DNA repeat expansions and human disease.* Cellular and Molecular Life Sciences, 2000. **57**(6): p. 914-31.

101. Christiansen, J., M. Kofod, and F.C. Nielsen, *A guanosine quadruplex and two stable hairpins flank a major cleavage site in insulin-like growth factor II mRNA.* Nucleic Acids Res, 1994. **22**(25): p. 5709-16.

102. Hwang, C.B., et al., *A net +1 frameshift permits synthesis of thymidine kinase from a drug-resistant herpes simplex virus mutant.* Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**(12): p. 5461-5.

103. Horsburgh, B.C., et al., *Translational recoding induced by G-rich mRNA sequences that form unusual structures.* Cell, 1996. **86**(6): p. 949-59.

104. Oliver, A.W., et al., *Preferential binding of fd gene 5 protein to tetraplex nucleic acid structures.* J Mol Biol, 2000. **301**(3): p. 575-84.

105. Oliver, A.W. and G.G. Kneale, *Structural characterization of DNA and RNA sequences recognized by the gene 5 protein of bacteriophage fd.* Biochem J, 1999. **339**(Pt 3): p. 525-31.

106. Wen, J.D., C.W. Gray, and D.M. Gray, *SELEX selection of high-affinity oligonucleotides for bacteriophage Ff gene 5 protein.* Biochemistry, 2001. **40**(31): p. 9300-10.

107. Weiss, S., et al., *RNA aptamers specifically interact with the prion protein PrP.* J Virol, 1997. **71**(11): p. 8790-7.

108. Shafer, R.H. and I. Smirnov, *Biological aspects of DNA/RNA quadruplexes.* Biopolymers, 2001. **56**(3): p. 209-27.

109. Rye, P.D. and K. Nustad, *Immunomagnetic DNA aptamer assay.* Biotechniques, 2001. **30**(2): p. 290-2, 294-5.

110. Lee, M. and D.R. Walt, *A fiber-optic microarray biosensor using aptamers as receptors.* Anal Biochem, 2000. **282**(1): p. 142-6.

111. Hamaguchi, N., A. Ellington, and M. Stanton, *Aptamer beacons for the direct detection of proteins.* Anal Biochem, 2001. **294**(2): p. 126-31.

CHAPTER 2: ENGINEERING THE QUADRUPLEX FOLD: NUCLEOSIDE CONFORMATION DETERMINES BOTH FOLDING TOPOLOGY AND MOLECULARITY IN GUANINE QUADRUPLEXES

2.1 ABSTRACT

The distribution of *syn* and *anti* glycosidic bond configurations differs in the various quadruplex folding topologies. Because RNA prefers the glycosidic bond to be in the *anti* conformation, specific positioning of ribonucleosides into a G-quadruplex-forming DNA sequence provides the possibility of controlling the folding topologies. Substitutions of RNA into the Thrombin Binding Aptamer sequence d(GGTTGGTGTGGTTGG) allows exploration of these possibilities using CD experiments. This method allows the studies of the physical, biochemical, and biological properties of these various folding topologies.

2.2 INTRODUCTION

In addition to the all-parallel quadruplexes formed from four strands of guanine-rich oligonucleotides, another all-parallel motif was demonstrated in more recent structural studies. It is particularly interesting as this structure is formed from a single strand folded several times on itself. The earlier studies of this folding motif include one that was depicted by the $d(GGA)_4$ sequence, which involves a guanine quartet stacked on top of a GAGAGAG heptad [1]. This structure is similar to the two-stranded quadruplex formed by two $d((GGA)_2G)$ sequences orientated parallel to each other with a guanine quartet stacked on an AGGGGA hexad [2]. Study of the *c-myc* promoter region sequence has also revealed the formation of parallel unimolecular quadruplex conformation in solution [3, 4]. A recent X-ray crystal structure of the human telomere sequence (HTR) $d(AG_3(TTAG_3)_3)$ in the presence of potassium ions [5] showed formation of a parallel unimolecular structure (Figure 1). This observation is essentially different from the previously published solution structure of the identical sequence in the presence of sodium ions, which depicts the antiparallel unimolecular folding topology (Figure 2) [6]. Other evidence for the parallel folded configuration of this sequence was recently shown by cross-linking experiments using dinuclear platinum complexes [7].

In light of these observations regarding the various folding topologies for unimolecular quadruplexes, it is of interest to assess the factors that determine whether a given sequence will fold in an antiparallel or parallel fashion. We approach this question here by examining the effect of *syn/anti* (Figure 3)

preferences of individual guanine nucleosides on the overall fold of Thrombin Binding Aptamer (**dTBA**) (Figure 4), using site-specific substitution of deoxyribonucleosides, dG's, with ribonucleosides, rG's. It is well known that in oligo- and polynucleotides, dG has a significantly greater propensity for the *syn* conformation than does rG [8]. In the case of polynucleotides, for example, it is more difficult to induce the Z-form, left-handed conformation in poly(rG-rC)·poly(rG-rC) than in poly(dG-dC)·poly(dG-dC) [9], which can be attributed, at least in part, to the higher energy of the *syn* conformation in RNA, as compared to DNA, in the context of double helical structures.

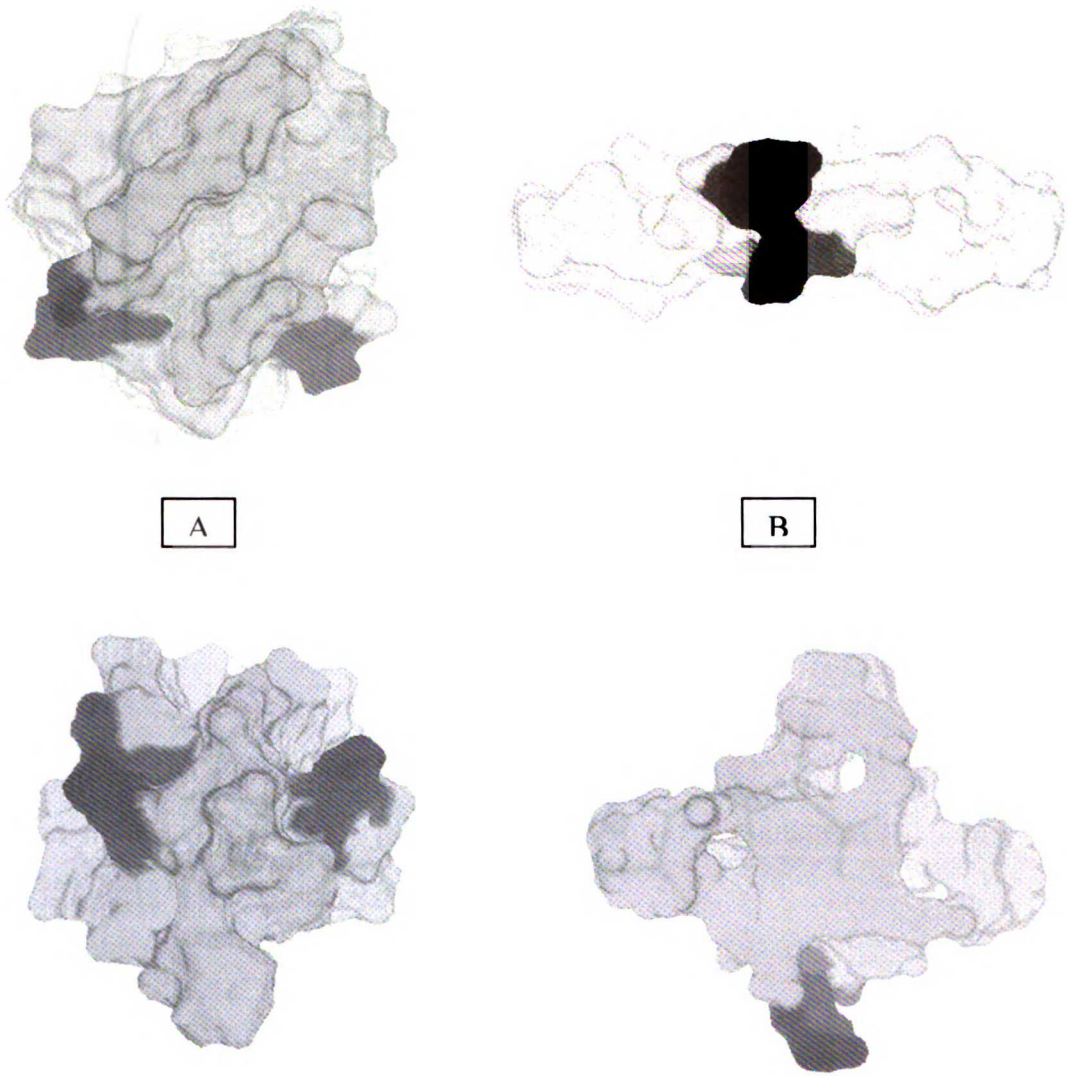

Figure 1: HTR structure viewed perpendicular to helix axis (top row) or viewed down the helix axis (bottom row); the shaded surfaces are the 5' and 3' terminals. A: NMR structure (antiparallel unimolecular quadruplex) B: X-ray structure (parallel unimolecular quadruplex)

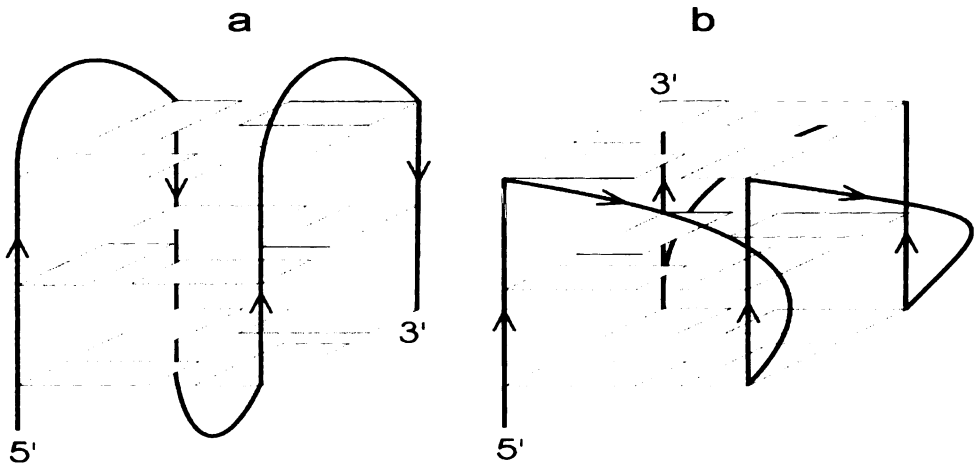

Figure 2: Schematic of different quadruplex folding topologies for HTR. a: NMR structure. b. X-ray structure [5].

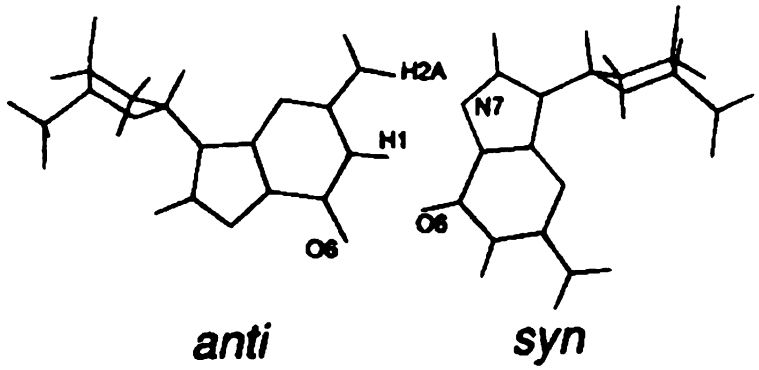

Figure 3: *anti* and *syn* conformation about glycosidic bond of guanosine

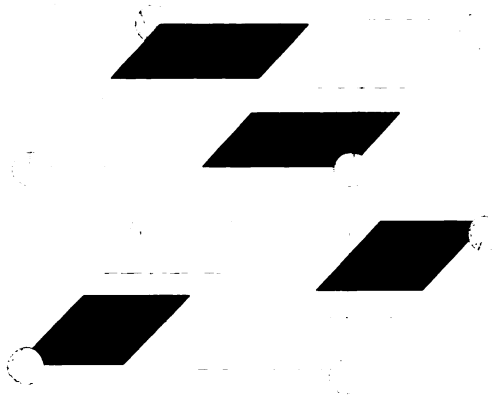

Figure 4: Schematic of **dTBA** antiparallel unimolecular quadruplex structure. The *syn* conformation guanosines are shaded gray.

2.3 MATERIALS AND METHODS

2.3.1 Sample Preparation

All oligonucleotides were purchased from IDT (Coralville, IA) with ion-exchange HPLC purification. The sequences were modified based on the 15-mer sequence of the DNA thrombin binding aptamer, **dTBA**, as follows:

dTBA: **5'-GGTTGGTGTGGTTGG-3'**

dTBAr1: **5'-gGTTgGTGTgGTTgG-3'**

dTBAr2: **5'-GgTTGgTGTGgTTGg-3'**

rTBA: **5'-gguuggugugguugg-3'**

rTBAd1: **5'-GguuGguguGguuGg-3'**

rTBAd2: **5'-gGuugGugugGuugG-3'**

Upper cases are used for deoxyribonucleosides and lower cases for ribonucleosides. The underlined positions represent those nucleosides with *syn* conformation about the glycosidic bond in the unimolecular quadruplex formed by the **dTBA** [10, 11].

Samples were annealed by heating up to 90 °C for 5 minutes, then cooling slowly to 15 °C in 10 mM Li-cacodylate buffer, pH 6.8, and with the addition of 25 mM or 140 mM KCl as indicated.

2.3.2 Circular Dichroism

CD experiments were carried out on an Aviv 215 Circular Dichroism spectrophotometer equipped with a peltier temperature controller. All CD spectral

scans were accumulated over the wavelength range 220-320 nm, using 1 cm path-length cells from Starna Cells Inc, at 15 °C, unless otherwise noted. The scan of buffer alone was subtracted from each sample.

2.3.3 UV Measurements

Concentrations were determined spectrophotometrically, based on ε_{260} values (Table 1) estimated by nearest neighbor methods [12], using 1 cm path-length cells from Starna Cells Inc. UV melting curves were measured in a Cary 100 spectrophotometer equipped with a peltier heating device and run at a heating rate of 0.2 °C/min. Thermodynamic analysis was carried out with an all-or-none model for quadruplex unfolding, assuming temperature-independent enthalpy and entropy changes.

Sequence	**L(mol strand-cm)$^{-1}$**
dTBA	147,920
dTBAr1	150,020
dTBAr2	150,100
rTBA	156,380
rTBAd1	154,280
rTBAd2	154,160

Table 1: ε260 values used for concentration determinations.

2.3.4 Nondenaturing Polyacrylamide Gel Electrophoresis (PAGE)

Samples of 80 μM concentration were heated in 10 mM Tris acetate, pH 6.8, and 25 mM KCl for 1 min at 90 °C, then cooled down slowly to room temperature. Samples were then run on a 10 cm × 10 cm, 20% polyacrylamide (19:1 acrylamide:bisacrylamide ratio) nondenaturing gel, made up in 40 mM Tris-acetate buffer containing 25 mM KCl (also used for the running buffer) for 2.5 hours at 240V in a cold room at 4 °C. Gels were stained with stains-all and destained in 70% ethanol.

2.3.5 1D Imino Proton NMR

NMR spectra were obtained on a Varian 600 MHz NMR instrument with a watergate pulse [13] for solvent suppression. Samples (500 μM strand) were prepared in 90 % H_2O/10 % ^2H_2O in the presence of 25 mM KCl and 10 mM perdeuterated Tris-acetate, pH 6.8.

2.4 RESULTS

2.4.1 Circular Dichroism

Circular dichroism (CD) spectra can be used to assess the nature of the quadruplex fold, based on observations that antiparallel quadruplexes present a long-wavelength positive maximum near 293 nm in the presence of sodium or potassium, while parallel quadruplexes exhibit a CD maximum near 265 nm [14]. This difference has been attributed to the requirement for half of the deoxyguanosine nucleosides to possess the *syn* conformation about the

glycosidic bond in order to fully realize the hydrogen bonding pattern of the guanine quartet in antiparallel conformations [15]. Although the use of CD to distinguish parallel from antiparallel folding topologies has proved to be reliable in general, some have suggested that it may not always lead to the correct result [16]. Nonetheless, the vast majority of published reports are consistent with these observations; for example, there is no instance of a known parallel quadruplex with a single, positive peak near 293 nm. Thus the very distinct spectra described below can be confidently used to distinguish parallel from antiparallel quadruplex structures.

In Figure 5A, we present CD spectra for the thrombin binding aptamer, **dTBA**. The spectrum in the presence of buffer alone shows several low-amplitude bands indicative of a relatively unstructured state, and similar to what has been reported earlier in the absence of added metal ions [17]. Addition of 25 mM KCl results in a dramatic overall change in the CD spectrum, with a large positive maximum near 293 nm, characteristic of antiparallel quadruplex structures, in this case an antiparallel, unimolecular folded quadruplex. The high resolution NMR structure of this sequence stabilized by K+ reveals the overall fold to be that shown in Figure 4 and also establishes that half of the deoxyguanosines are in the *syn* conformation [10, 11]. Similar observations have been reported for all other quadruplexes involving antiparallel folds, including both unimolecular and bimolecular constructs [18, 19].

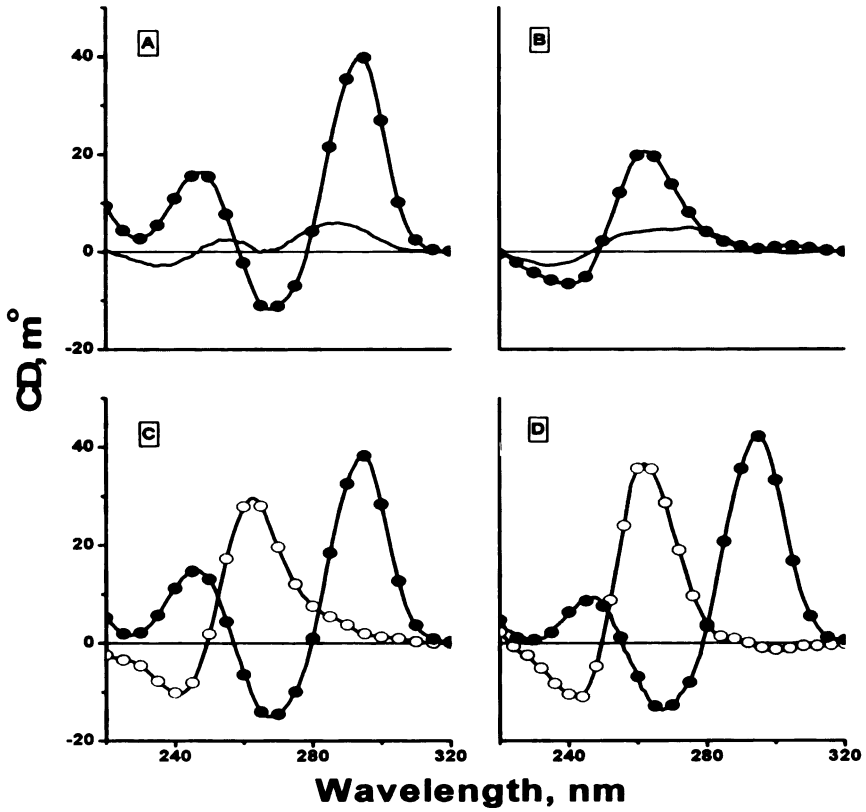

Figure 5: CD spectra for various sequences at 7μM concentration in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, unless otherwise specified. A: **dTBA** in buffer alone (——) or with 25 mM KCl (•); B: **rTBA** in buffer alone (——) or with 25 mM KCl (•); C: **dTBAr1**, in buffer with 25 mM KCl, 5 °C, measured at 60 μM in a 0.1 cm cell; (○), **dTBAr2** in buffer with 25 mM KCl (•); D: **rTBAd1** in buffer with 25 mM KCl (•), **rTBAd2** in buffer with 25 m M KCl (○).

When the same experiment is carried out with the all-RNA analog, **rTBA**, the CD spectrum in the absence of added metal ion resembles that of **dTBA**, consisting of very low amplitude features characteristic of a minimally structured state (Figure 5B). The addition of KCl (25 mM) results again in a change in the CD spectrum, but to a significantly different pattern. Most importantly, the long-wavelength maximum is shifted to 265 nm, characteristic of parallel quadruplex structures. This spectrum, consisting of a single positive peak, is clearly distinct from that of **dTBA** and indicates a major change in overall structure for **rTBA**. In particular, a parallel quadruplex conformation implies that all nucleosides are in the *anti* conformation, demonstrating that replacement of all deoxyribonucleosides with their corresponding ribonucleosides results in a refolding of the antiparallel structure to a parallel structure.

We next made more selective substitutions of ribo- for deoxyribonucleosides. Our rationale was based on the well known observation that ribonucleosides strongly prefer the *anti* conformation about the glycosidic bond to the *syn* conformation, as mentioned above. Thus selective substitution of the dG nucleosides in **dTBA** that possess the *syn* conformation about the glycosidic bond may significantly destabilize the antiparallel fold, possibly leading to a parallel fold.

The NMR solution structures for **dTBA** reveal that dG nucleosides in positions 1, 5, 10, and 14 all possess the *syn* conformation [10, 11]. Thus we substituted just those positions with rG nucleosides, leading to the oligonucleotide **dTBAr1**. For this sequence, addition of KCl to buffer results in a

positive CD peak at 265 nm, similar to that of the all-RNA sequence **rTBA**, as illustrated in Figure 5C. This result is fully consistent with our predictions, based on the observation that ribonucleosides strongly prefer the *anti* over the *syn* conformation when incorporated in oligo- or polynucleotides. This spectrum was obtained at relatively high strand concentration (and low temperature); CD spectra obtained under standard conditions showed little difference from those observed in buffer alone (Figure 6).

Additional efforts were made to stabilize the quadruplex formed by the chimeric **dTBAr1**. It has been shown that addition of 30% methanol (MeOH) to the buffer stabilizes quadruplex structure formation (Smirnov and Shafer, Biopolymers, in press). This was indeed shown in the CD spectra in Figure 6 of the chimeric **dTBAr1** in buffer solution containing of 30% MeOH. The addition of MeOH results in an overall change in the CD spectrum, with a positive maximum peak at 265 nm. In order to confirm that addition of MeOH only stabilized the quadruplex structure, which results in the increase in the amplitude of the corresponding CD peaks, we did a similar addition to **dTBA** (Figure 7) **and rTBA** (Figure 8).

Results of native polyacrylamide gel electrophoresis (shown below) suggest the possibility that parallel structures may consist of more than one strand. Hence, we investigated structure formation of **dTBAr1** in higher concentration conditions, which results in more distinct overall CD changes, as shown in Figure 5C and Figure 9. This concentration dependence indicative of the possibility of multi-stranded parallel structures is investigated further below.

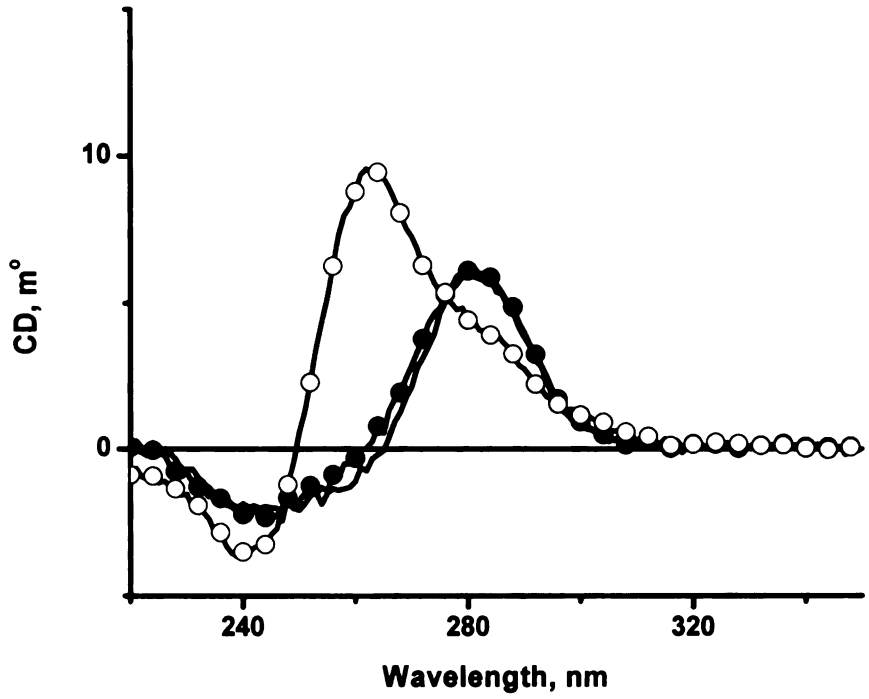

Figure 6: CD spectra for **dTBAr1** at 7 μM in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **dTBAr1** in buffer alone (—), with 25 mM KCl (●), or with 25 mM KCl and 30% MeOH(○).

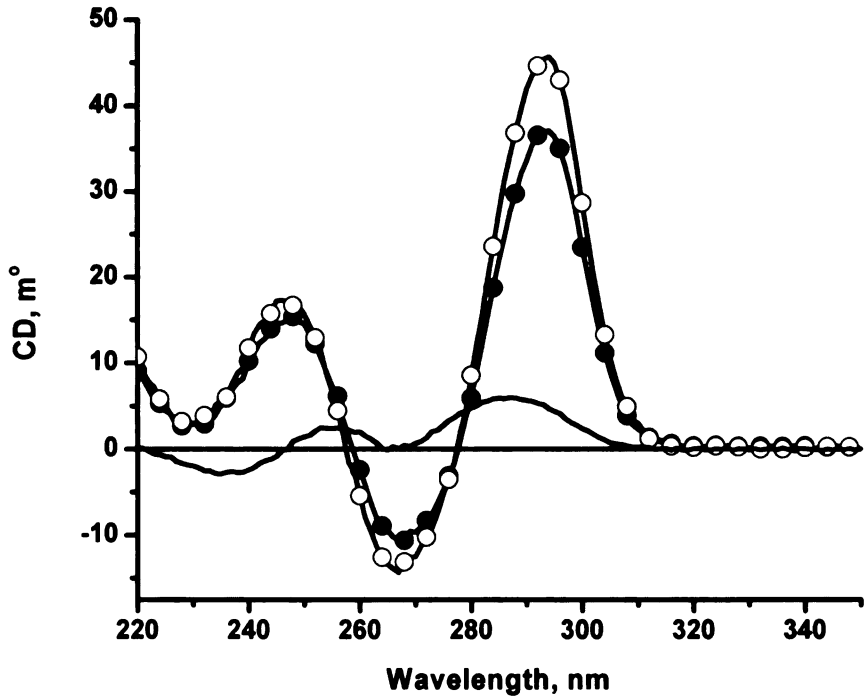

Figure 7: CD spectra for **dTBA** at 7 μM in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **dTBA** in buffer alone (—), with 25 mM KCl (•), or with 25 mM KCl and 30% MeOH (○).

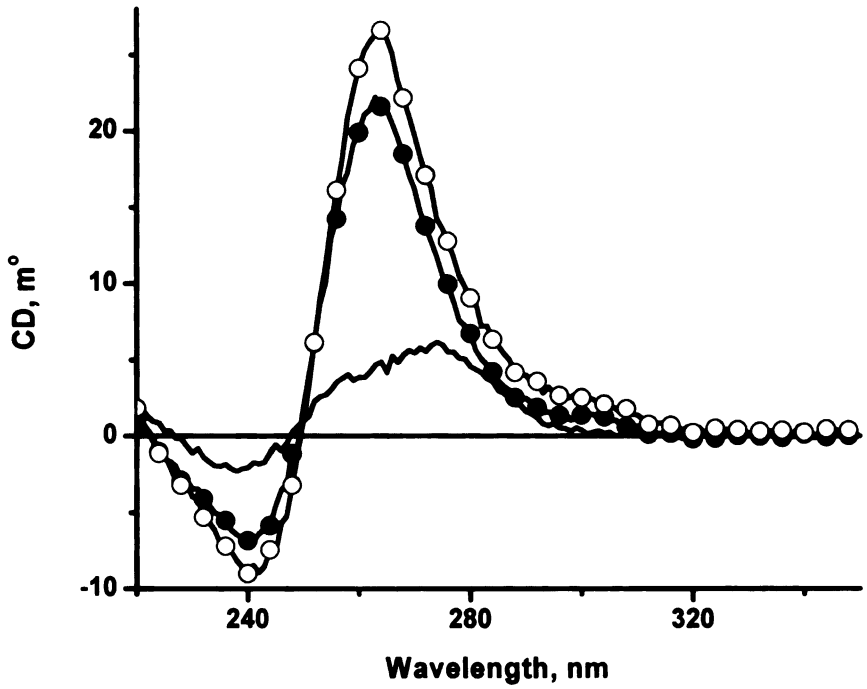

Figure 8: CD spectra for **rTBA** at 7 μM in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **rTBA** in buffer alone (—), with 25 mM KCl (•), or with 25 mM KCl and 30% MeOH (○).

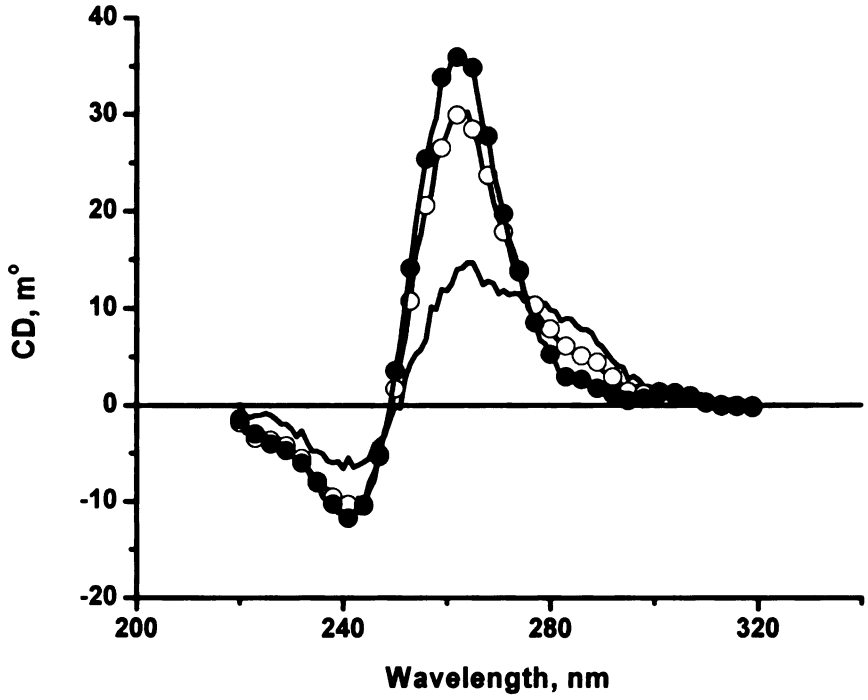

Figure 9: CD spectra for **dTBAr1** at 60 μM in 10 mM Li-cacodylate at 15 °C, or 5 °C in 0.1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **dTBAr1** with 25 mM KCl at 15 °C (—), with 25 mM KCl at 5 °C (○),and with 140 mM KCl at 5°C (●).

As an additional test of our hypothesis that the *syn-anti* behavior of individual nucleosides plays a major role in affecting the overall quadruplex fold, we examined a control sequence, **dTBAr2**, in which the other four dG residues (those which possess the *anti* conformation in the antiparallel quadruplex formed by **dTBA**) have been replaced by rG. Here we expected little or no difference compared to **dTBA** and indeed this turns out to be the case, as illustrated in Figure 5C, where it can be seen that **dTBAr2** exhibits a positive CD peak at 293 nm, consistent with an antiparallel fold.

Having converted the antiparallel fold of **dTBA** to a parallel fold by selective substitution of rG for dG, we turned to the converse, i.e. converting the parallel fold of **rTBA** to an antiparallel fold. Thus we investigated the properties of **rTBAd1**, in which the rG nucleosides at positions 1, 5, 10 and 14 are substituted by dG. Here we have selectively replaced syn-phobic nucleosides with syn-philic ones and now this chimeric sequence folds into an antiparallel quadruplex, as demonstrated by its CD spectrum shown in Figure 5D. Again, replacing the other rG nucleosides with dG, to yield **rTBAd2**, the relevant control sequence, leads to a parallel quadruplex fold, similar to its parent sequence **rTBA**, also shown in Figure 5D.

We have repeated these CD spectra in the presence of 100 mM NaCl instead of 25 mM KCl (Figure 10). For **dTBA**, **rTBA, dTBAr2**, and **rTBAd1**, the spectra observed in NaCl are similar to those observed in KCl, except with smaller amplitudes. In the case of **dTBAr1** and **rTBAd2**, addition of NaCl led to little or no change compared to buffer alone, indicating no quadruplex formation,

most likely due to the extremely weak strand association for these sequences under these conditions.

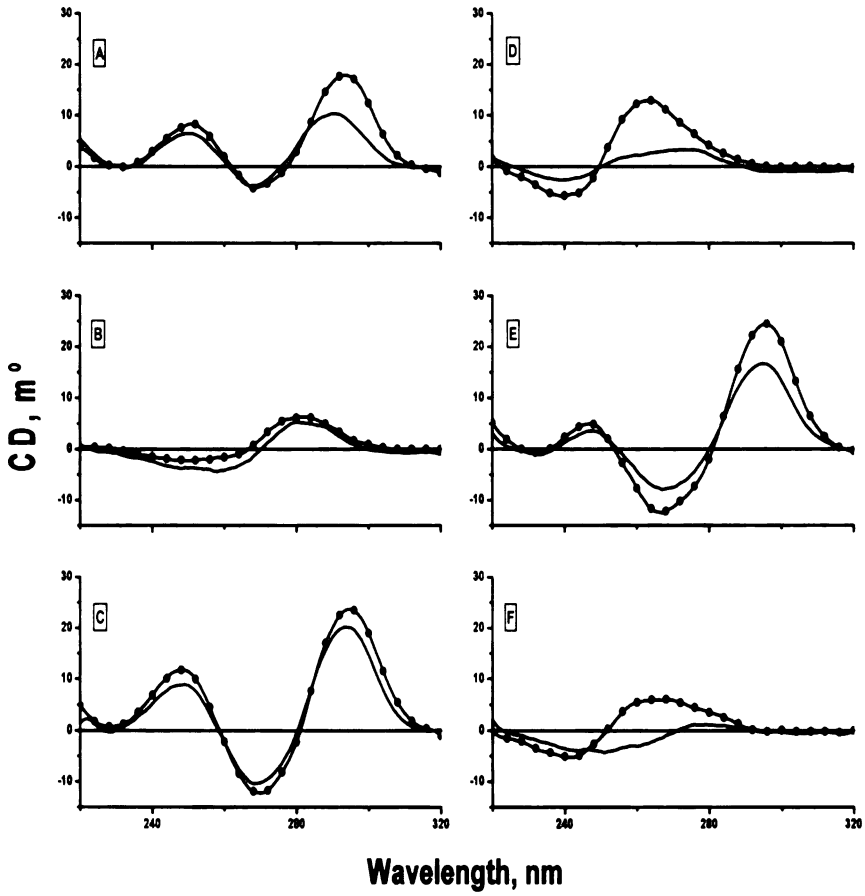

Figure 10: CD spectra for various sequences at 7 μM concentration in 10 mM Li-cacodylate, in buffer alone (——) or with 100 mM NaCl (•), at 15 °C in 1 cm path cells. A: **dTBA** ; B: **dTBAr1**; C: **dTBAr2**; D: **rTBA**; E: **rTBA**1; F: **rTBA**2.

2.4.2 Polyacrylamide Gel Electrophoresis

In order to further examine the possibility that some of the species considered here consist of more than one strand, we carried out native polyacrylamide gel electrophoresis. In Figure 11, it is apparent that the sequences which exhibit positive CD peaks at 293 nm (**dTBA**, **dTBAr2**, and **rTBAd1**) exhibit higher electrophoretic mobility in comparison to those sequences possessing a CD peak at 265 nm (**rTBA**, **rTBAd2** and **dTBAr1**). The position of the slower-moving bands, near the gel position for the 20- and 27-mer components of the ladder, suggests that these structures are composed of more than one strand, and are most likely two-stranded rather than four-stranded species; the latter would be expected to migrate even more slowly in the gel. There is evidence for some minor species in some of the lanes, which could represent small amounts of impurities or a small population of unfolded or differently folded oligonucleotide. In this regard, we note that these samples were run at relatively high concentration in order to facilitate the formation of any two-stranded structures. Nonetheless, the positions of the major species are clear, and indicate a unimolecular fold for the antiparallel structures and a bimolecular fold for the parallel structures.

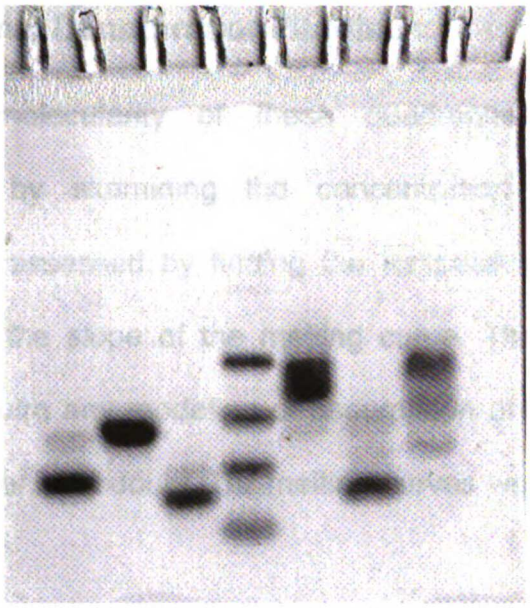

Figure 11: Non-denaturing polyacrylamide gel run at 4 °C. Lanes 1 – 7: **dTBAr2**, **dTBAr1**, **dTBA**, DNA ladder (dT10, dT15, dT20, dT27), **rTBA**, **rTBAd1**, **rTBAd2**, respectively.

2.4.3 Thermal Denaturation Studies

The molecularity of these quadruplex structures can be further investigated by examining the concentration dependence of the melting temperature, assessed by finding the temperature, T_{max}, corresponding to the maximum in the slope of the melting curve. This measure of thermal stability does not require any modeling or assumption of molecularity and thus is useful when the latter is in doubt. UV melting curves were measured for each of these sequences.

The effect of strand concentration on the melting temperature for **dTBA**, **rTBA** and **rTBAd1** is summarized in Figure 12, where it can be seen that the T_{max} of the all-RNA sequence, **rTBA**, exhibits significant concentration dependence while that of the other two sequences appears independent of strand concentration. Among these three sequences, only **rTBA** exhibited a substantial hysteresis and a biphasic character in its melting profile. In such cases, the T_{max} was determined for the first transition in the heating direction. Similar behavior has been reported for this sequence by others [20]. While this suggests a more complicated set of conformations for this sequence and possible kinetic issues, rendering thermodynamic analysis difficult, the strand concentration dependence is nonetheless apparent. The two antiparallel structures, **dTBA** and **rTBAd1**, do not exhibit a concentration-dependent T_{max} and hence behave as unimolecular species, consistent with their high gel mobilities. In contrast, both gel behavior and concentration dependence of

melting temperature support a two-stranded species for the parallel quadruplexes in this study.

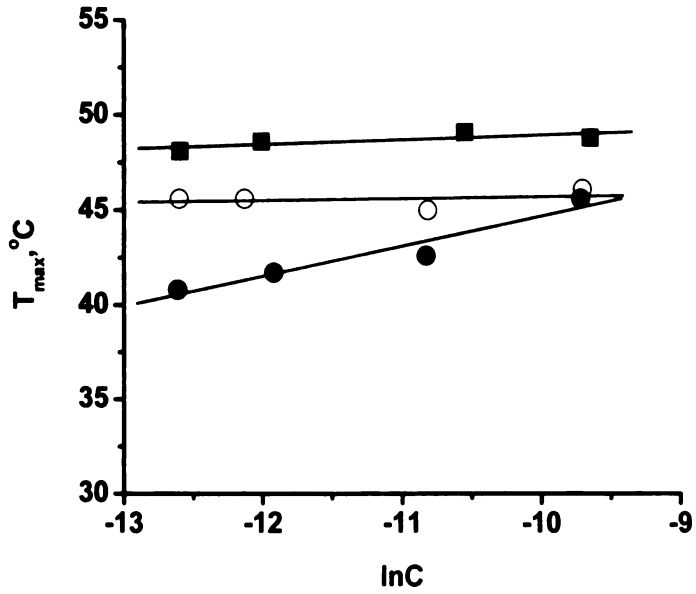

Figure 12: Thermal denaturation temperature, T_{max}, as a function of strand concentration. Melting curves were determined in 10 mM Li-cacodylate buffer, pH 6.8, containing 25 mM KCl. Straight lines represent least-squares fit to the data points.

Results for T_{max} values are presented in Table 2. The all-DNA sequence, **dTBA**, melts at 46 °C, in excellent agreement with earlier work of sSmirnov and Shafer on this sequence under similar conditions [21]. As this sequence folds into a unimolecular quadruplex, its melting temperature is independent of strand concentration. In the case of the parallel structures, however, specifying a melting temperature is more ambiguous, as one must also specify strand concentration. Nonetheless, at relatively high strand concentration, **dTBAr1** exhibits the lowest T_{max}, and hence has the lowest thermal stability overall. For other parallel quadruplex structures, **rTBA**, at high strand concentration, exhibits a melting temperature similar to that of **dTBA** but a melting temperature about 5 °C lower than **dTBA** measured at low strand concentration, while **rTBAd2** melts at a fairly low temperature, 29 °C, at low strand concentration. These results are consistent with the need to measure the CD spectrum of **dTBAr1** at low temperature and high strand concentration. Similarly, they are consistent with the lack of quadruplex formation by **dTBAr1** and **rTBAd2** in NaCl solution, as these two sequences form the least stable quadruplexes in KCl.

Sequence	T_{max}, °C
dTBA	46
dTBAr1 (80 μM)	22
dTBAr2	52
rTBA	41
rTBAd1	49
rTBAd2	29

Table 2: Melting temperatures determined by UV thermal denaturation. Samples run in 10 mM Li-cacodylate buffer, pH 6.8 containing 25 mM KCl and at 7 μM strand concentration unless otherwise specified.

In the case of the remaining, unimolecular antiparallel quadruplexes, **dTBAr2** and **rTBAd1**, both exhibit higher T_{max} values than **dTBA**. This leads to the general observation that all antiparallel quadruplexes have higher melting temperatures than all of the parallel quadruplexes, indicating that, at least for this base sequence, the parallel folding topology is less thermally stable than the antiparallel topology. Furthermore, among the three antiparallel structures, the all-DNA sequence is the least stable, while among the three parallel structures; the all-RNA sequence is the most stable.

Results from thermodynamic analysis of the UV melting curves for those sequences forming unimolecular antiparallel quadruplexes are presented in

Table 3. Similar analysis for parallel-forming species was not carried out due to significant hysteresis in their heating and cooling curves. In terms of thermal stability, the all-DNA sequence is the least stable while the mainly DNA sequence, **dTBAr2**, is the most stable. The intermediate thermal stability of the mainly RNA sequence, **rTBAd1**, suggests that RNA can essentially fold in a manner similar to that of its DNA analog when the barrier to taking on the *syn* conformation is removed. Both RNA-containing sequences in Table 3 appear to have greater enthalpic contributions to the stability of their folded structures. This may be due to stacking differences and/or other specific effects. The slightly lower thermal stability of **rTBAd1** compared to **dTBAr2** derives from destabilizing entropic effects, which more than compensate the slight enthalpic preference for the former. Since differences in these terms are small, it is difficult to assess their real significance. What does appear unambiguous is that both RNA-containing sequences are more stable than the all-DNA sequence.

Sequence	T_m (°C, ± 0.5)	ΔH° (kcal/mol, ±2)	ΔS° (cal/(mol-deg), ±4)	ΔG°(37 °C) (kcal/mol, ±0.2)
dTBA	45.9	40.3	127	0.93
dTBAr2	51.8	45.3	139	2.21
rTBAd1	49.4	46.2	144	1.72

Table 3: Thermodynamic parameters from van't Hoff analysis of the thermal denaturation of sequences forming unimolecular antiparallel quadruplexes in 25 mM KCl.

2.4.4 Imino proton NMR spectra

One of the signature properties of structures based on the guanine quartet involves the presence of imino proton resonances in the ~10.5 – 12.5 ppm range [18, 19]. Thus we measured the imino proton spectrum for both **dTBA** and **rTBA** in order to confirm the presence of a quadruplex conformation in these sequences. The imino proton spectrum of **dTBA** has been well characterized [10, 11] and we include it here simply for reference. Results for both **dTBA** and **rTBA** are presented in Figure 13. The spectrum for **dTBA** resembles that already reported, again indicative of guanine imino protons engaged in quartet hydrogen bonding. Similarly, the spectrum for **rTBA** also contains resonances characteristic of guanine quartets.

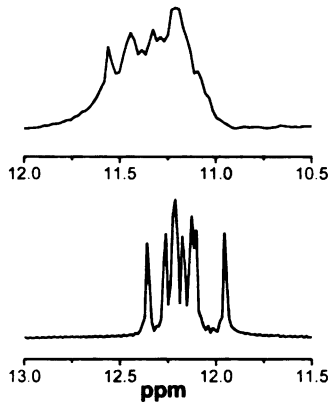

Figure 13: Quartet imino proton spectra for **rTBA** (top) and **dTBA** (bottom), 500 μM strand, in 10 mM perdeuterated Tris-acetate buffer, pH 6.8, and 25 mM KCl, 5 °C, .

The spectrum of each structure covers a similar range of frequencies, with that for **dTBA** slightly smaller than that for **rTBA**. Additionally, the resolution is substantially better in the spectrum for **dTBA**. The apparently broader lines in the **rTBA** spectrum may be due to the fact that this structure consists of more than one strand and its spectrum may possess more than eight resonances. For our purposes, however, the significance of these spectra is that they confirm the presence of a guanine quartet based structure for **rTBA**, the all-RNA version of this sequence.

2.5 DISCUSSION

Guanine quadruplex stability is affected by multiple forces. Of primary significance is the ability to form complete guanine quartets. In general, the more quartets in a structure, the more stable it will be, all other things held constant. In the case of folded quadruplexes, we can anticipate a role for the loop sequences

in modulating the stability of these structures. The importance of loop length and sequence in determining the thermal stability of **dTBA** has been shown previously [21]. More recently, the effect of loop length on quadruplex stability has been demonstrated in other sequence contexts [16]. Evidence suggesting the possibility of parallel unimolecular quadruplexes appeared over a decade ago. For example, in an early study on the human telomere repeat sequence [22], CD spectra, in the presence of K^+ ions, showed peaks at both 265 nm and 293 nm, suggesting the presence of both parallel and antiparallel quadruplex structures. Several years later, a report appeared describing both DNA and RNA forms of the *Tetrahymena* repeat sequence, $d(T_2G_4)_4$, including data suggesting that the RNA oligonucleotide was both unimolecular and parallel [23].

More recently, several structural studies demonstrating a DNA parallel quadruplex conformation have appeared. Patel and co-workers described an unusual four-stranded DNA structure, in the presence of 150 mM Na^+, in which two strands of d(GGAGGAG) formed a quartet and an AGGGA hexad, with all GG segments parallel to each other; two such structures complexed with each other via stacking of their respective hexads [2]. Thus within each two-stranded hexad-quartet structure, all GG segments are parallel to each other, but those of one such structure are antiparallel to those of the other.

Katahira and co-workers have reported the NMR solution structure of the 12-mer $d(GGA)_4$ in the presence of 10 mM NaCl and 100 mM KCl [24]. This structure involves a single strand folding in a parallel fashion to form one guanine quartet stacked over a GAGAGAG septet. As in the AGGGGA hexad structure

mentioned above, two such structures stack via their GAGAGAG heptads, giving rise to a two-stranded structure. Shortly following that report, an X-ray crystal structure for the human telomere sequence, $d(AG_3(T_2AG_3)_3)$, appeared [25]. This structure, obtained in the presence of K^+ ions (and a quadruplex-binding ligand), exhibited a parallel folding topology, in striking contrast to the antiparallel, folded structure reported about a decade earlier for the same sequence by NMR in solution, in the presence of Na^+ ions [6]. More recently, evidence for a two-stranded, parallel quadruplex formed by a two-fold repeat of the human telomere has been described, in which the two strands each contribute half of the guanines in each quartet [26]. G-rich sequences from the control region of the *c-myc* promoter have also been shown to form a unimolecular parallel quadruplex [3, 4, 27].

RNA quadruplexes have been studied less extensively than DNA quadruplexes. One example is a study demonstrating different quadruplex structures formed by $d(G_2AG_2T_4G_2AG_2A)$ and its RNA analog, $r(G_2AG_2U_4G_2AG_2A)$; the DNA sequence forms a typical bimolecular antiparallel, diagonal loop type structure while the RNA sequence forms a dimer of two parallel unimolecular quadruplexes [28, 29]. This result echoes the results mentioned earlier, in which CD and gel electrophoresis data indicated, under certain conditions, an antiparallel, unimolecular structure for $d(T_2G_4)_4$ in contrast to a parallel, unimolecular structure for $r(U_2G_4)_4$ [23]. Some physical studies have been carried out on an RNA quadruplex-containing aptamer targeted to the Fragile-X mental retardation protein, although its folding topology is not apparent

[30, 31].

At this time, then, there is ample documentation of both parallel and antiparallel folded quadruplexes, yet relatively little is known regarding the forces which determine how a specific sequence folds. One study has shown that the normally antiparallel folded quadruplex formed by $d(G_4(T_4G_4)_3)$ can be induced to fold to a parallel quadruplex in the presence of K^+ when certain single nucleosides are replaced with a locked nucleic acid (LNA) analog [32]. No clear evidence for change from antiparallel to parallel was reported, however, in the presence of Na^+. This approach also sought to exploit the preference of the LNA analog for the *anti* conformation, yet in two of the three cases where conversion from antiparallel to parallel was essentially complete, the substitutions occurred at sites that normally possess the *anti* conformation. Thus the rationale for those results is not clear. In the results presented here, the role of the substituted nucleosides is unambiguously demonstrated via the control sequences and the results are seen for both K^+ and Na^+.

A more recent report has appeared describing the effect of replacing all quartet deoxyguanosines in **dTBA** with the LNA analog, which led to a structure with a parallel-like CD spectrum and substantially lower thermal stability than **dTBA** [33]. No information concerning the molecularity of this structure was provided, although the observed hysteresis in CD melting profiles for this modified sequence indicated very slow kinetics of quadruplex formation, consistent with the possibility of a multi-stranded species. In this regard, it is of interest to note that some two-stranded quadruplexes have been shown to

exhibit very slow kinetics of formation/dissociation [26] and thus the observation of hysteresis in melting curves is consistent with either two- or four-stranded structures. The specific role of the *syn – anti* properties of the LNA analog was not considered.

The potential role of loops in directing a parallel or antiparallel fold has been investigated by Neidle and co-workers [16]. In that study, it was concluded that loop length may play a major role in determining the overall fold. Specifically, the study finds that the parallel structure is the only stable form when all loops consist of a single base; in most other cases, both parallel and antiparallel folds are possible. This conclusion is supported by previously published structures of parallel quadruplexes, both DNA and RNA, which possess two or three single-base loops [24, 27, 29, 34].

The influence of loops on the CD spectrum of quadruplexes has also been considered. For example, in a study on the CD spectra of a series of telomere sequences from vertebrates, Bolton and co-workers have suggested that in spectra possessing positive peaks near 260 and 290 nm, that the longer wavelength peak may arise from the three-base (TTA) loops while the other peak arises from the quartet core [35]. Similarly, others have shown that the CD spectrum of the four-fold repeat of *Tetrahymena*, $d(T_2G_4)_4$, in the presence of NaCl possesses positive peaks at both 260 and 290 nm [23, 36]; one explanation is that the unusual fold of this sequence, involving two lateral loops and one double-reversal, or propeller, loop [37] is both part parallel and part antiparallel, and thus may have a CD spectrum with both peaks. However, the alternative

explanation attributing the two positive peaks to two different species cannot be ruled out, especially in light of gel electrophoresis results suggesting the presence of two different species under these conditions [23].

Our CD results provide some additional insight into this question, as parallel and antiparallel species display CD spectra possessing either the positive peak at 265 nm or at 293 nm, but not both. The loops, likely composed of two or three bases, must have different folds yet do not appear to make a separate contribution to the CD signal. Of course, it is difficult to generalize from a single example, and thus the general CD properties of parallel and antiparallel quadruplexes remain to be elaborated.

The main focus, however, of this work is to investigate the effect of the guanine nucleoside on the nature of the quadruplex fold. Specifically, we have demonstrated that dG residues at certain positions can be replaced with rG residues (a difference of a single oxygen atom on the sugar ring) resulting in a major change in the overall folding topology. Starting with all-DNA (antiparallel conformation), substitution of specific dG residues can reverse the sense of the fold. Similarly, the parallel fold of the all-RNA sequence can be reversed by specific substitution of rG with dG. We rationalize these changes by noting that RNA nucleosides exhibit a strong preference for the *anti* conformation about the glycosidic bond, related to the fact that ribonucleosides exhibit the C3'-endo (N) type of sugar pucker. This preference for the *anti* conformation in RNA is not absolute, as there are examples of RNA structures containing *syn* ribonucleosides, such as Z RNA [9], hairpins [38], G-A mismatches in duplexes

[39] and ribozymes [40]. Nonetheless, within the context of the sequences studied here, the inherent preference of ribonucleosides for the *anti* conformation appears to be sufficiently strong to force quadruplex refolding.

In addition to reversing the quadruplex fold from antiparallel to parallel, specific replacement of dG with rG also changes the molecularity of the structure from one to two strands. At this time, we can only speculate in a limited manner on the arrangement of the two strands in the parallel quadruplex formed by **dTBAr1** (and by **rTBA** and **rTBAd2** as well). As illustrated in Figure 4 in Chapter 1, several different arrangements are possible for parallel quadruplexes: (i) a single unimolecular species – Figure 4B, (ii) a dimer of unimolecular quadruplexes and (iii) a single quadruplex composed of two strands - Figure 4C. Our results indicate that the latter two forms are possible arrangements for the parallel structures investigated here; additional studies will be necessary to distinguish (ii) from (iii). Examples of type (ii) structures in the literature involve a single quartet stacked on a larger guanine quartet-containing platform composed of purines, such as an AGGGGA hexad [2, 29] or an GAGAGAG heptad [24]. Given the sequence of **TBA**, this may be unlikely to form in the parallel quadruplexes studied here, leaving type (iii) as the more likely strand configuration.

It is well appreciated experimentally that in most DNA and RNA structures, the *anti* conformation about the glycoside bond is more frequently observed than the *syn* conformation, implying that the former is more stable than the latter. Relating these observations to the conformational properties of the nucleic acid

building blocks is considerably more complicated, as the glycosidic torsion angle is coupled to other degrees of freedom, such as sugar pucker. For example, solution studies on purine nucleosides indicate only relatively small differences between *syn* and *anti* conformations [8], yet this similarity in energy is not observed at the oligo- or polynucleotide level. A clearer correlation is seen, however, regarding sugar pucker, with ribonucleosides preferring the C3' pucker and deoxyribonucleosides preferring the C2' pucker. Furthermore, computations have led to quite different energetics between the various conformational states of nucleosides and nucleotides [41-44].

The results described herein demonstrate that the observed preference of ribonucleosides for the *anti* conformation about the glycosidic bond in a large number of structures also holds for the guanine quadruplex. The energetic difference between the *syn* and *anti* conformations of rG, incorporated into oligomeric structures, is sufficiently large to destabilize the antiparallel fold relative to the parallel fold. Similarly, the relative ease with which dG can exhibit the *syn* conformation reverses the situation, resulting in an antiparallel fold for a mainly RNA chimeric sequence. Thus there does not appear to be anything inherent to RNA chemistry or structure preventing antiparallel folding, other than the glycosidic bond conformational properties. The switch in folding is not the only change that occurs, however, as the parallel species examined here consist of more than one strand, while the antiparallel species are unimolecular. This observation raises the question of the relationship between folding topology and molecularity. The **TBA** sequence presents no obvious obstacle to simply forming

a unimolecular, parallel quadruplex. Unimolecular parallel structures have been reported containing one-base, two-base and three-base loops [16, 27]. Apparently there is some driving force for a two-stranded, parallel structure in this case such that the unimolecular species has a higher free energy than the corresponding two-stranded structure, suggesting that the entropic cost of associating two strands is less than the enthalpic cost of forming the unimolecular parallel conformation.

One possible factor influencing the molecularity of parallel quadruplexes is the sequence in the loops. All of the known structures of unimolecular parallel DNA quadruplexes, i.e. those in which a single strand folds into a complete parallel quadruplex, contain adenine in one or more loops and also contain at least one single-base loop [4, 24, 25, 27]. In contrast, the TBA sequence studied here contains no adenine and no single-base loop sequence. This suggests a possible role for loop sequence in determining the folding pattern for quadruplexes and also is consistent with earlier studies indicating that single-base loops are most readily accommodated in unimolecular parallel quadruplexes [16].

While it is somewhat premature to generalize, all RNA quadruplexes investigated to date appear to exhibit a parallel folding topology, while only DNA quadruplexes have been found in both parallel and antiparallel folds. Such differences could have important biological implications. For example, the intramolecular guanine quadruplex in mRNA that binds the fragile-X mental retardation protein [45] is likely a parallel, rather than an antiparallel, structure.

Formation of 8-methylguanine from exposure to certain carcinogens has been reported [46, 47]; this could potentially occur in both DNA and RNA *in vivo*, and impact the nature of quadruplexes formed in G-rich regions of the genome or of RNA, due to the influence of methylation at the 8 position on the syn-anti conformational preferences of the modified nucleoside. The inactivity of the RNA version of a DNA quadruplex aptamer that binds ATP could be due to a change from antiparallel to parallel folding [48]. Finally, it is likely that antiparallel quadruplex structures proposed for G-rich RNA sequences, such as that shown by *in vitro* selection to bind tightly to prions [49], will need revision.

2.6 OTHER SEQUENCES

Similar studies by CD were carried out with different sequences. Selection of one of the sequences was based on the report describing both DNA and RNA forms of the *Tetrahymena* repeat sequence, **dTETTEL** [23]. Once again the sequence was modified based on the DNA form $d(T_2G_4)_4$ with reference to its solution structure formed in Na^+-containing solution [37]. The other sequence studied was based on the *Oxytricha* repeat sequence, **dOTEL**, with the modification based on the solution structure reported in the presence of metal cation [50]. A similar notation was adopted for these sequences, with upper cases used for deoxyribonucleosides, lower cases for ribonucelosides, and underlined positions represent those nucleosides with *syn* conformation about the glycosidic bond in the quadruplex structures formed by the DNA forms. Sample preparations were same as described in section 2.3.1.

dTETTEL: 5'-TTGGGGTTGGGGTTGGGGTTGGGG-3'

dTETTELr: 5'- TTgGGGTTGggGTTGgGGTTgGGG -3'

dOTEL: 5'-GGGGTTTTGGGGTTTTGGGGTTTTGGGG-3'

dOTELr: 5'-gGgGTTTTgGgGTTTTgGgGTTTTgGgG-3'

rOTEL: 5'-gggguuuugggguuuugggguuuuggggg-3'

dHTRr1: 5'-AGgGTTAgGgTTAGgGTTAgGg-3'

dHTRr2: 5'-AgGgTTAGgGTTAgGgTTAGgG-3'

2.6.1 CD results for dTETTEL and dTETTELr

In Figure 14, there is no dramatic overall change in CD spectrum for **dTETTEL** with the addition of 25 mM KCl. There are two major CD peaks at 265 nm and 293 nm indicative of the presence of mixture of parallel and antiparallel quadruplex structures. However, for the chimeric **dTETTELr** in the presence of 25 mM KCl (Figure 15), the major structures are the parallel structures indicated by the major CD maximum peak at 265 nm.

The chimeric **dTETTELr** in the presence of 100 mM NaCl has a different CD pattern. As can be seen in Figure 16, the two peaks with maxima at 265 nm and 293 nm of almost equivalent amplitude are indicative that approximately equal amount of parallel and antiparallel quadruplex structures are formed by the chimeric **dTETTELr** in the presence of 100 mM NaCl. This result does not replicate the results reported by Sugimoto and co-workers [50]. Also, since the modification was based on the solution structure of **dTETTEL** in the presence of Na^+, it is surprising that the chimeric sequence formed more stable parallel structures in the presence of KCl as opposed to NaCl. Further studies could be carried out to investigate the possibility of multi-stranded parallel quadruplex formation as shown in the **dTBA** and corresponding chimeric sequences.

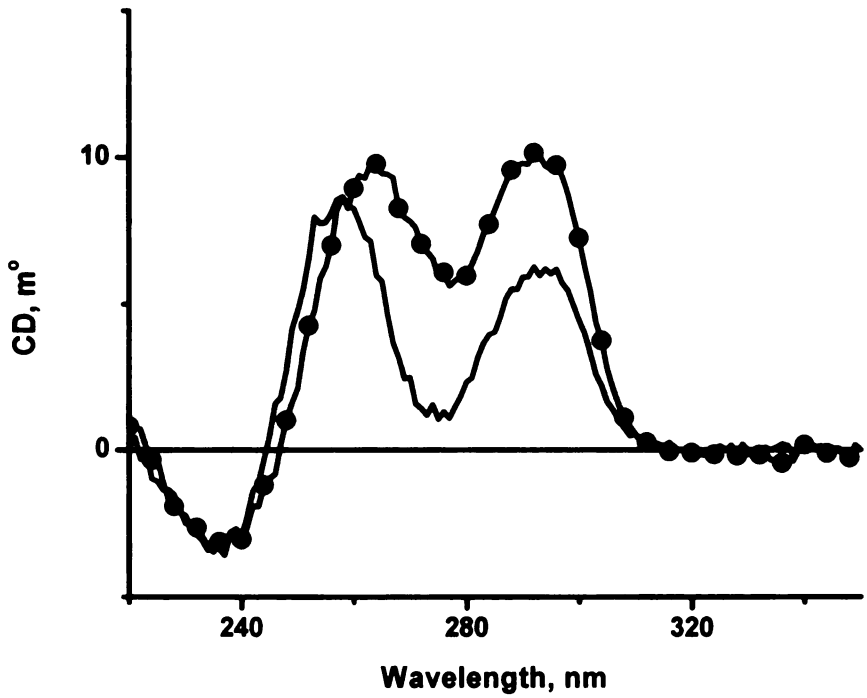

Figure 14: CD spectra for **dtettel** at 1 OD(Concentration) in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. dtettel in buffer alone (—) or with 25 mM KCl (•).

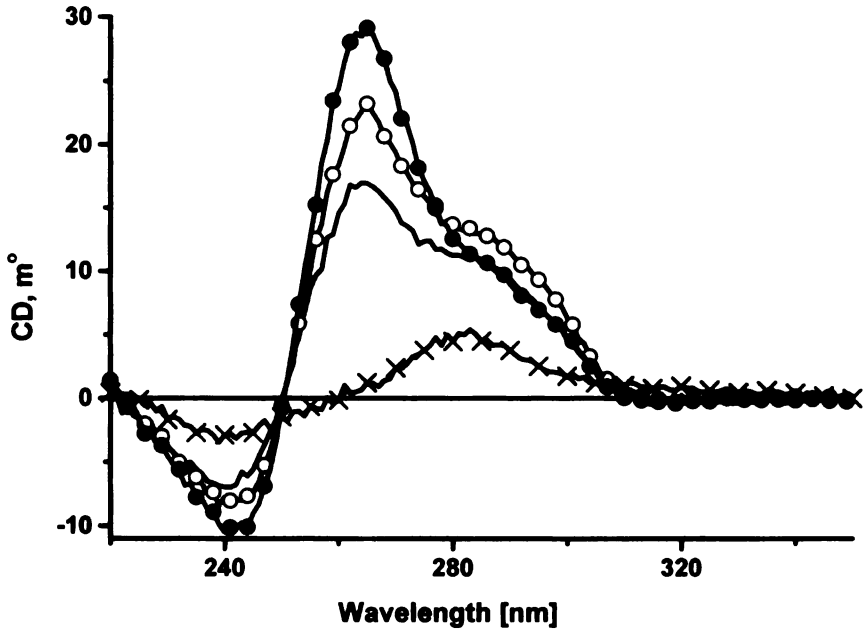

Figure 15: CD spectra for **dtetteIr** at 1 OD in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down, or noted otherwise. **dtetteIr** in buffer alone (—), with 25 mM KCl (○), two days later heated at 90 °C (x) and cooled down to 15 °C (●).

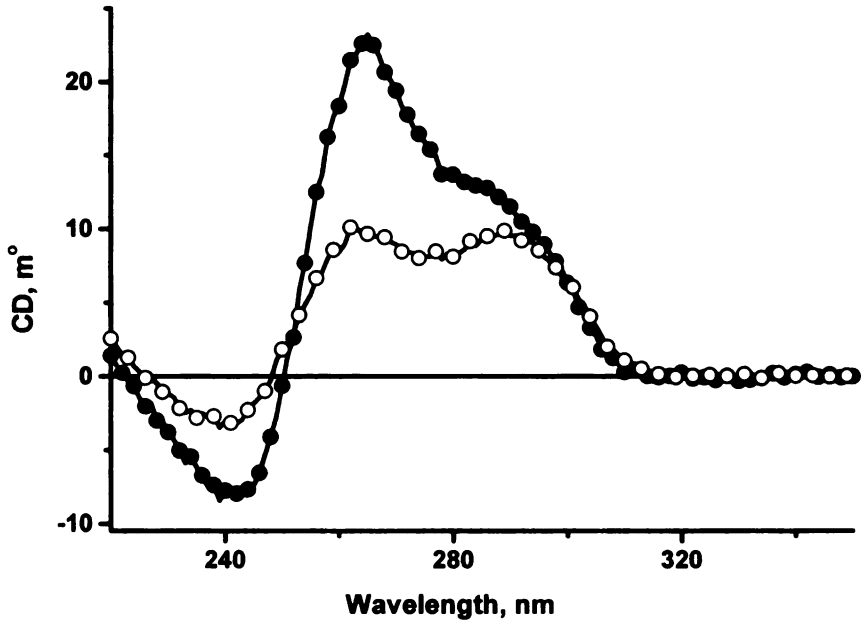

Figure 16: CD spectra for **dtettelr** at 1 OD in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **dtettelr** in buffer with 25 mM KCl (•) or with 100 mM NaCl (○).

2.6.2 CD results for dOTEL, rOTEL, and dOTELr

As was seen in Figure 17, addition of 25 mM KCl to **dotel** results in a CD spectrum of a positive maximum CD peak at 293 nm characteristic of antiparallel quadruplex structures. However, a different CD spectrum is (Figure 18) obtained by the addition of NaCl at both concentrations of 50 mM and 100 mM.

When the same experiments were carried out with the all RNA analog, **rOTEL**, the addition of both 25 mM KCl (Figure 19) and 100 mM NaCl (Figure 20), resulted in significant CD changes were observed. Both spectra showed one large, positive maximum at 265 nm with the characteristic of parallel quadruplex structures.

The chimeric **dOTELr** also revealed a quite significant overall CD change (Figure 21) with the expected positive CD peak at 265 nm appearing with the addition of 25 mM KCl. Once again, no considerable CD changes (Figure 22) in the presence of NaCl at both concentrations.

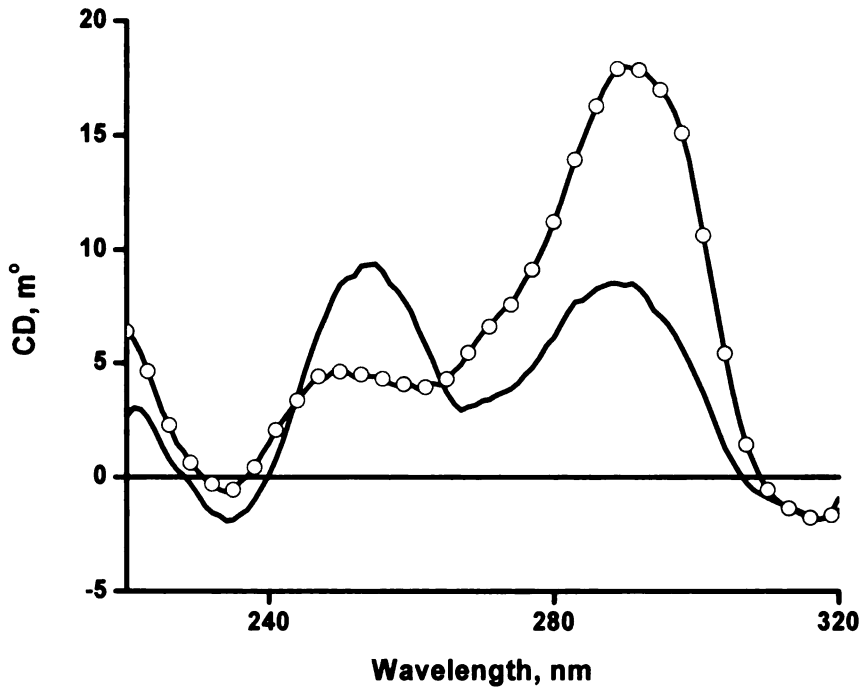

Figure 17: CD spectra for **dOTEL** at 1 OD in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **dOTEL** in buffer alone (—) or with 25 mM KCl (○).

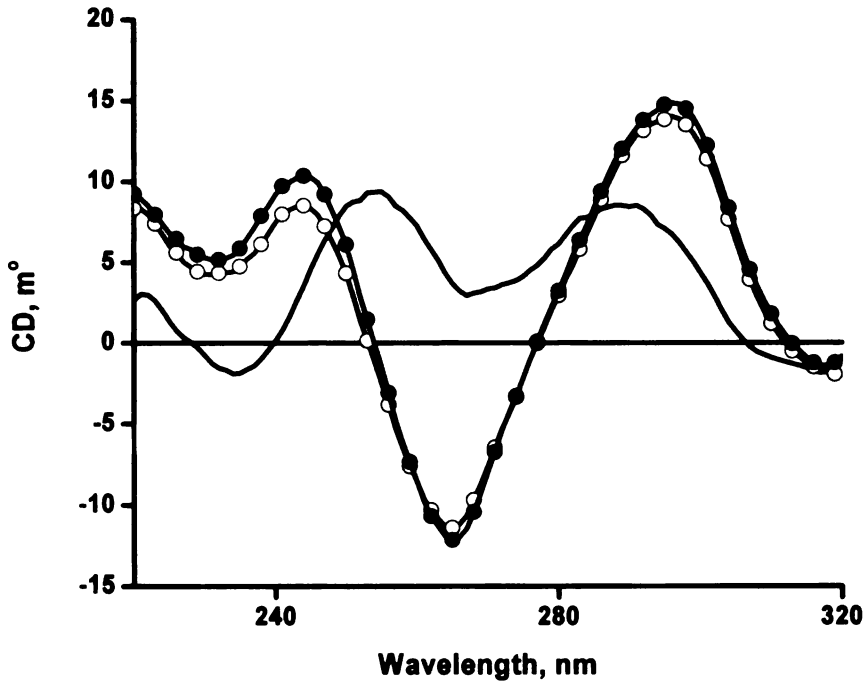

Figure 18: CD spectra for **dOTEL** at 1 OD in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **dotel** in buffer alone (—), with 50 mM NaCl (○),and with 100 mM NaCl (●).

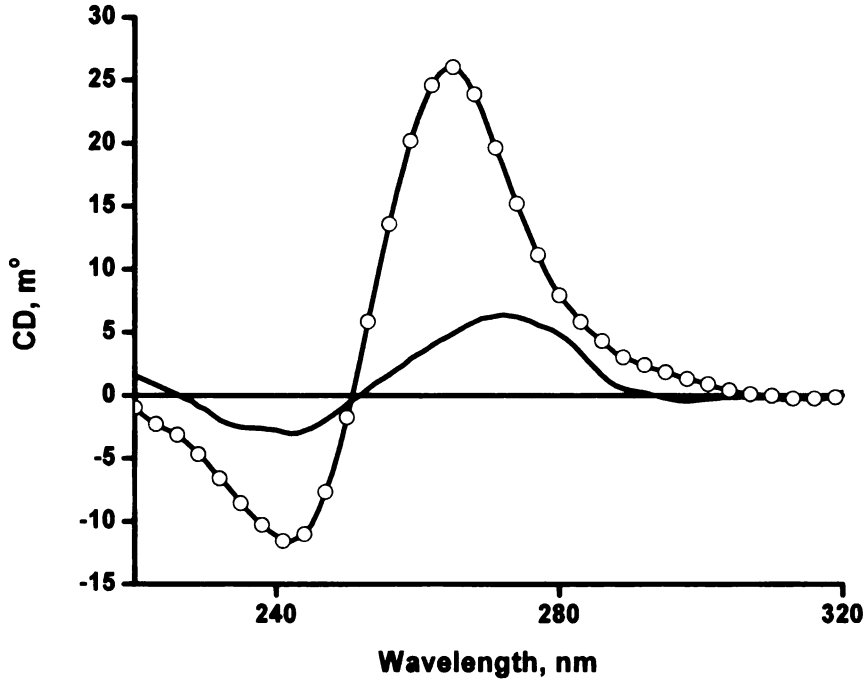

Figure 19: CD spectra for **rOTEL** at 1 OD in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **rotel** in buffer alone (—) or with 25 mM KCl (○).

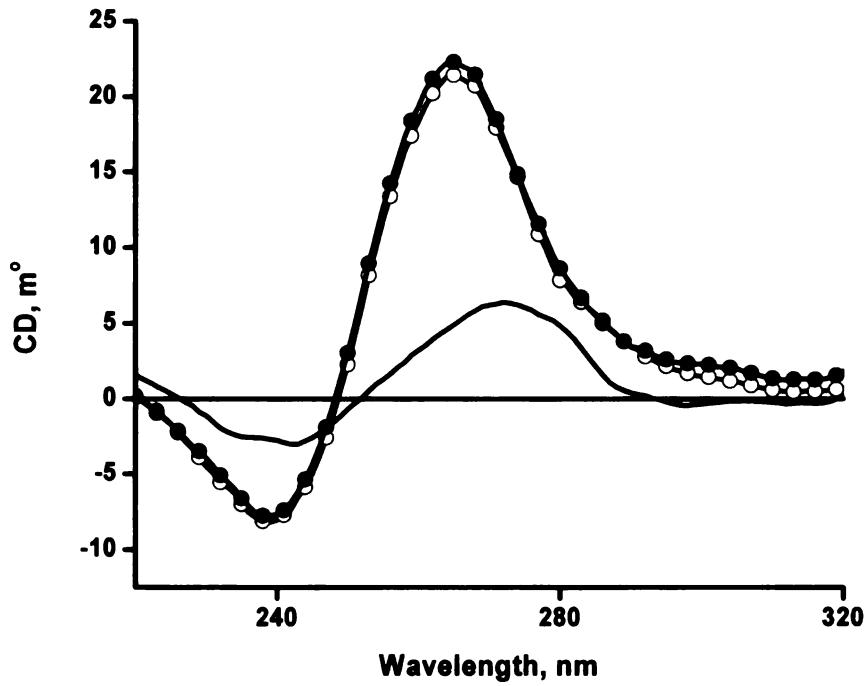

Figure 20: CD spectra for **rOTEL** at 1 OD in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **rOTEL** in buffer alone (—), with 50 mM NaCl (○),and with 100 mM NaCl (●).

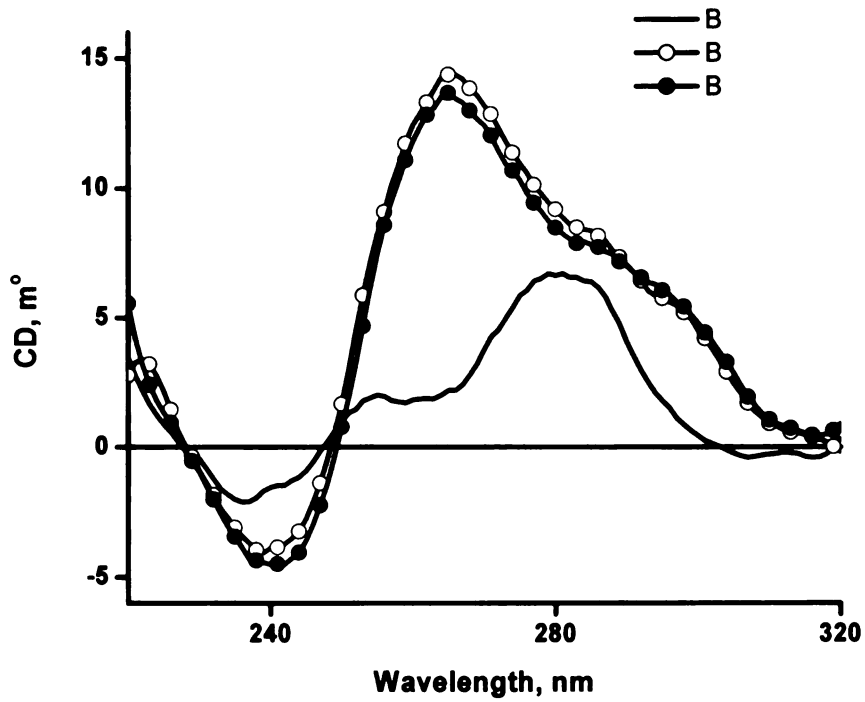

Figure 21: CD spectra for **dOTELr** at 1 OD in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **dOTELr** in buffer alone (—), with 25 mM KCl (○),and with 25 mM KCl at 5°C (●).

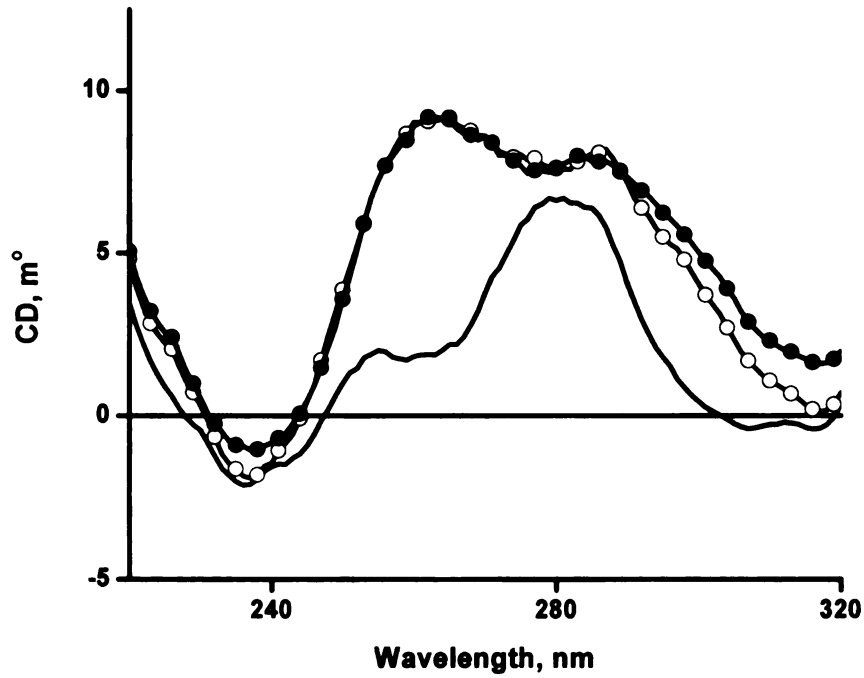

Figure 22: CD spectra for **dOTELr** at 1 OD in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **dOTELr** in buffer alone (—), with 50 mM NaCl (○),and with 100 mM NaCl (●).

2.6.3 (

s a sc

quadr

specr

two da

to dH

struct

prese

chang

this 2

days

signa

2.6.3 CD results of dHTRr1 and dHTRr2

Figure 23 shows the CD spectra of **dHTRr1**. As seen in Figure 23A, there is a small positive maximum CD peak at 293 nm characteristic of antiparallel quadruplex structures as a result of the addition of 100 mM NaCl to **dHTRr1**. The spectrum remains the same after leaving the solution in the same condition for two days.

There is a more significant change in CD with the addition of 25 mM KCl to **dHTRr1** with a positive peak at 265 nm, which is a characteristic of parallel structures, and also a broad shoulder extended to 293 nm. This may indicate the presence of mixed parallel and antiparallel structures. However, the CD spectrum changed peak having a maximum at 293 nm in the presence of 30% MeOH and this 293 nm peak increases when the solution left in the same condition for 2 days.

As for the other chimeric sequence, **dHTRr2**, no significant change in CD signals occurs with the addition of 100 mM NaCl (Figure 24).

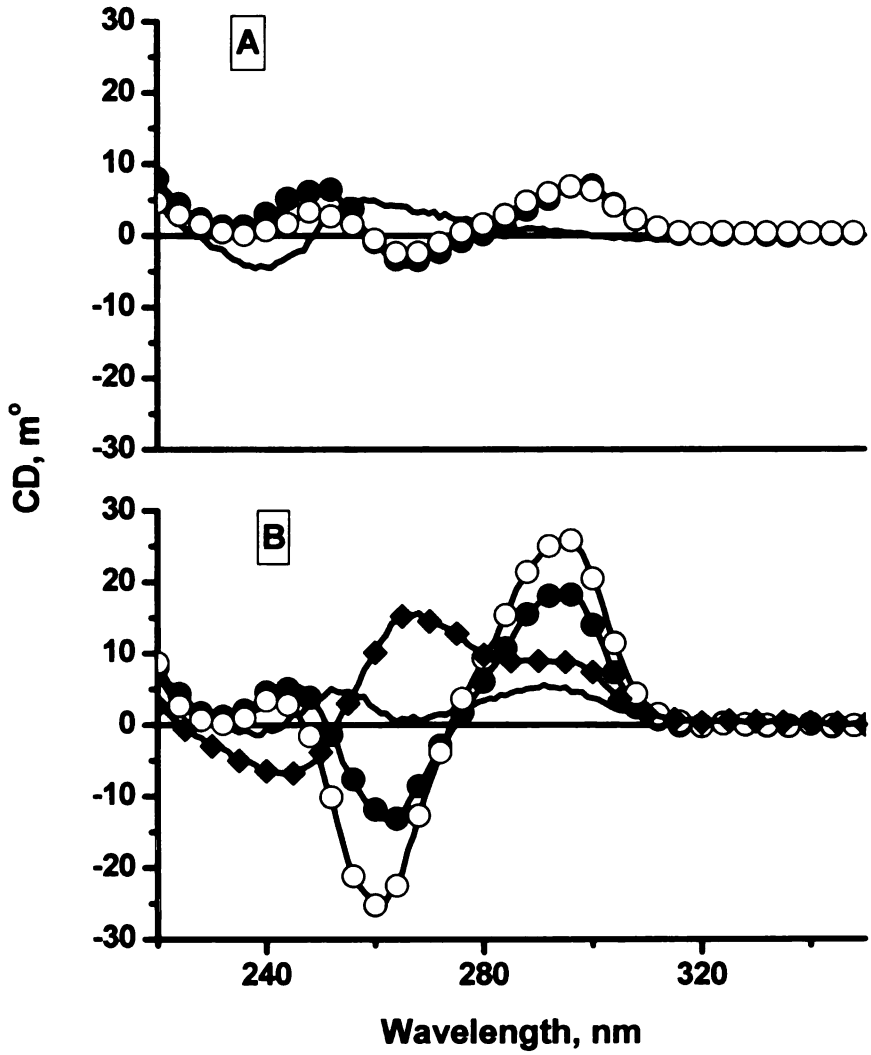

Figure 23: CD spectra for **dHTRr1** at 1 OD in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. A: **dHTRr1** in buffer alone (—), with 100 mM NaCl (•), and 2 days in 100 mM NaCl (○). B: **dHTRr1** in buffer alone (—), with 25 mM KCl (•), with 25 mM KCl and 30% MeOH (○), or 2 days in 25 mM KCl and 30% MeOH (♦)

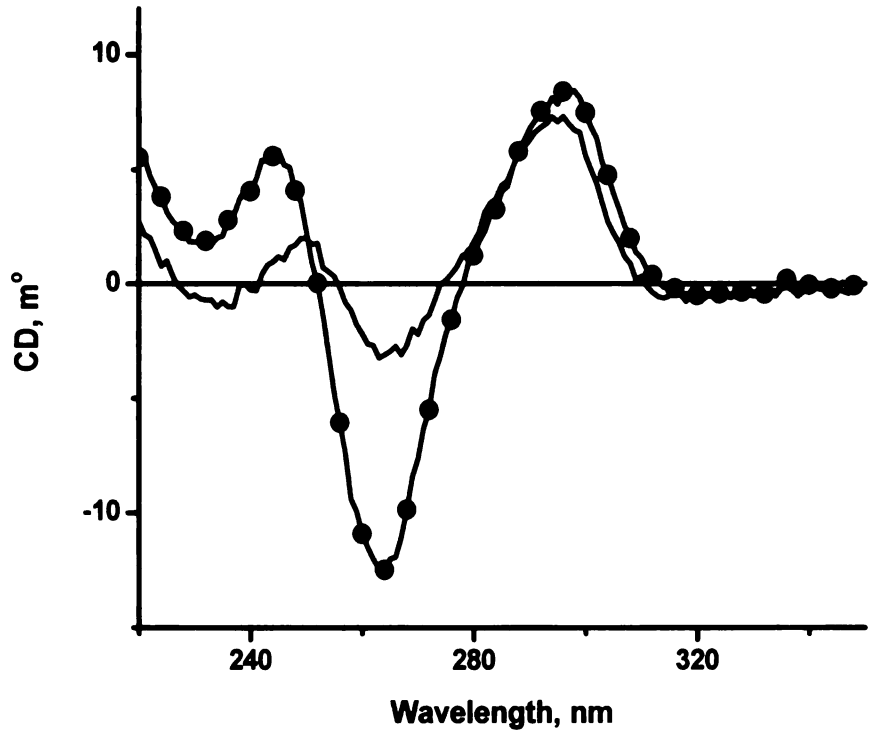

Figure 24: CD spectra for **dHTRr2** at 1 OD in 10 mM Li-cacodylate at 15 °C in 1 cm path cells, after being heated to 90 °C for 5 min and slowly cooled down. **dHTRr2** in buffer alone (—) and with 100 mM NaCl (•).

In summary, we have demonstrated how to engineer the overall folding pattern of guanine quadruplexes by making specific substitutions of ribo- for deoxyriboguanosines, or vice versa. These results clearly establish the role of individual nucleoside conformational preferences on the overall quadruplex fold, and illustrate how a minor modification, addition or deletion of a single oxygen atom on several nucleosides, can have a profound impact on nucleic acid secondary, tertiary, and quaternary structure.

2.7 REFERENCES

1. Matsugami, A., et al., *An intramolecular quadruplex of (GGA)(4) triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction.* J Mol Biol, 2001. **313**(2): p. 255-69.
2. Kettani, A., et al., *A dimeric DNA interface stabilized by stacked A.(G.G.G.G).A hexads and coordinated monovalent cations.* J Mol Biol, 2000. **297**(3): p. 627-44.
3. Seenisamy, J., et al., *The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4.* J Am Chem Soc, 2004. **126**(28): p. 8702-9.
4. Phan, A.T., Y.S. Modi, and D.J. Patel, *Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter.* J Am Chem Soc, 2004. **126**(28): p. 8710-6.
5. Parkinson, G.N., M.P. Lee, and S. Neidle, *Crystal structure of parallel quadruplexes from human telomeric DNA.* Nature, 2002. **417**(6891): p. 876-80.
6. Wang, Y. and D.J. Patel, *Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex.* Structure, 1993. **1**(4): p. 263-82.
7. Ourliac-Garnier, I., et al., *Cross-links of quadruplex structures from human telomeric DNA by dinuclear platinum complexes show the flexibility of both structures.* Biochemistry, 2005. **44**(31): p. 10620-34.
8. Saenger, W., *Principles of Nucleic Acid Structure.* 1984, Berlin: Springer-Verlag. 556.

9. Hall, K., et al., *'Z-RNA'--a left-handed RNA double helix.* Nature, 1984. **311**(5986): p. 584-6.

10. Macaya, R.F., et al., *Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution.* Proc Natl Acad Sci U S A, 1993. **90**(8): p. 3745-9.

11. Wang, K.Y., et al., *A DNA aptamer which binds to and inhibits thrombin exhibits a new structural motif for DNA.* Biochemistry, 1993. **32**(8): p. 1899-904.

12. Gray, D.M., S.H. Hung, and K.H. Johnson, *Absorption and circular dichroism spectroscopy of nucleic acid duplexes and triplexes.* Methods Enzymol, 1995. **246**: p. 19-34.

13. Piotto, M., V. Saudek, and V. Sklenar, *Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions.* J Biomol NMR, 1992. **2**(6): p. 661-5.

14. Hardin, C.C., A.G. Perry, and K. White, *Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids.* Biopolymers, 2001. **56**(3): p. 147-94.

15. Lu, M., Q. Guo, and N.R. Kallenbach, *Thermodynamics of G-tetraplex formation by telomeric DNAs.* Biochemistry, 1993. **32**(2): p. 598-601.

16. Hazel, P., et al., *Loop-length-dependent folding of G-quadruplexes.* J Am Chem Soc, 2004. **126**(50): p. 16405-15.

17. Smirnov, I. and R.H. Shafer, *Lead is unusually effective in sequence-specific folding of DNA.* Journal of Molecular Biology, 2000. **296**(1): p. 1-5.

18. Patel, D.J., et al., *Structures of guanine-rich and cytosine-rich quadruplexes formed in vitro by telomeric, centromeric and triplet repeat disease sequences*, in *Oxford Handbook of Nucleic Acid Structures*, S. Neidle, Editor. 1999, Oxford University Press: Oxford. p. 349-453.

19. Keniry, M.A., *Quadruplex structures in nucleic acids.* Biopolymers, 2001. **56**(3): p. 123-46.

20. Sacca, B., L. Lacroix, and J.L. Mergny, *The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides.* Nucleic Acids Res, 2005. **33**(4): p. 1182-92.

21. Smirnov, I. and R.H. Shafer, *Effect of loop sequence and size on DNA aptamer stability.* Biochemistry, 2000. **39**(6): p. 1462-1468.

22. Balagurumoorthy, P. and S.K. Brahmachari, *Structure and stability of human telomeric sequence.* J Biol Chem, 1994. **269**(34): p. 21858-69.

23. Sugimoto, N., T. Ohmichi, and M. Sasaki, *The stability of DNA and RNA G-Quartets.* Nucleosides Nucleotides Nucl Acids, 1996. **15**: p. 559-567.

24. Matsugami, A., et al., *An intramolecular quadruplex of (GGA)(4) triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction.* J Mol Biol, 2001. **313**(2): p. 255-69.

25. Parkinson, G.N., M.P. Lee, and S. Neidle, *Crystal structure of parallel quadruplexes from human telomeric DNA.* Nature, 2002. **417**(6891): p. 876-80.

26. Phan, A.T. and D.J. Patel, *Two-repeat human telomeric d(TAGGGTTAGGGT) sequence forms interconverting parallel and*

antiparallel G-quadruplexes in solution: distinct topologies, thermodynamic properties, and folding/unfolding kinetics. J Am Chem Soc, 2003. **125**(49): p. 15021-7.

27. Ambrus, A., et al., *Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization.* Biochemistry, 2005. **44**(6): p. 2048-58.

28. Liu, H., et al., *A comparison of the properties and the solution structure for RNA and DNA quadruplexes which contain two GGAGG sequences joined with a tetranucleotide linker.* Nucleosides, Nucleotides, Nucleic Acids, 2002. **21**: p. 785-801.

29. Liu, H., et al., *A dimeric RNA quadruplex architecture comprised of two G:G(:A):G:G(:A) hexads, G:G:G:G tetrads and UUUU loops.* J Mol Biol, 2002. **322**(5): p. 955-70.

30. Darnell, J.C., et al., *Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function.* Cell, 2001. **107**(4): p. 489-99.

31. Ramos, A., D. Hollingworth, and A. Pastore, *G-quartet-dependent recognition between the FMRP RGG box and RNA.* Rna, 2003. **9**(10): p. 1198-207.

32. Dominick, P.K. and M.B. Jarstfer, *A conformationally constrained nucleotide analogue controls the folding topology of a DNA g-quadruplex.* J Am Chem Soc, 2004. **126**(16): p. 5050-1.

33. Randazzo, A., et al., *Structural studies on LNA quadruplexes.* Nucleosides Nucleotides Nucleic Acids, 2005. **24**(5-7): p. 795-800.

34. Phan, A.T., et al., *An interlocked dimeric parallel-stranded DNA quadruplex: a potent inhibitor of HIV-1 integrase.* Proc Natl Acad Sci U S A, 2005. **102**(3): p. 634-9.

35. Rujan, I.N., J.C. Meleney, and P.H. Bolton, *Vertebrate telomere repeat DNAs favor external loop propeller quadruplex structures in the presence of high concentrations of potassium.* Nucleic Acids Res, 2005. **33**(6): p. 2022-31.

36. Rezler, E.M., et al., *Telomestatin and diseleno sapphyrin bind selectively to two different forms of the human telomeric G-quadruplex structure.* J Am Chem Soc, 2005. **127**(26): p. 9439-47.

37. Wang, Y. and D.J. Patel, *Solution structure of the Tetrahymena telomeric repeat d(T2G4)4 G-tetraplex.* Structure, 1994. **2**(12): p. 1141-56.

38. Proctor, D.J., et al., *Restricting the conformational heterogeneity of RNA by specific incorporation of 8-bromoguanosine.* J Am Chem Soc, 2003. **125**(9): p. 2390-1.

39. Pan, B., S.N. Mitra, and M. Sundaralingam, *Crystal structure of an RNA 16-mer duplex R(GCAGAGUUAAAUCUGC)2 with nonadjacent G(syn).A+(anti) mispairs.* Biochemistry, 1999. **38**(9): p. 2826-31.

40. Sigel, R.K., et al., *Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif.* Nat Struct Mol Biol, 2004. **11**(2): p. 187-92.

41. Gorb, L., O. Shishkin, and J. Leszczynski, *Charges of phosphate groups. A role in stabilization of 2'-deoxyribonucleotides. A DFT investigation.* J Biomol Struct Dyn, 2005. **22**(4): p. 441-54.

42. Shishkin, O.V., et al., *Conformational analysis of canonical 2-deoxyribonucleotides. 2. Purine nucleotides.* J Biomol Struct Dyn, 2004. **22**(2): p. 227-44.

43. Foloppe, N., L. Nilsson, and A.D. MacKerell, Jr., *Ab initio conformational analysis of nucleic acid components: intrinsic energetic contributions to nucleic acid structure and dynamics.* Biopolymers, 2001. **61**(1): p. 61-76.

44. Foloppe, N., et al., *Intrinsic conformational energetics associated with the glycosyl torsion in DNA: a quantum mechanical study.* Biophys J, 2002. **82**(3): p. 1554-69.

45. Darnell, J.C., et al., *Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function.* Cell, 2001. **107**(4): p. 489-99.

46. Netto, L.E., et al., *Identification of C8-methylguanine in the hydrolysates of DNA from rats administered 1,2-dimethylhydrazine. Evidence for in vivo DNA alkylation by methyl radicals.* J Biol Chem, 1992. **267**(30): p. 21524-7.

47. Gasparutto, D., et al., *Excision of 8-methylguanine site-specifically incorporated into oligonucleotide substrates by the AlkA protein of Escherichia coli.* DNA Repair (Amst), 2002. **1**(6): p. 437-47.

48. Huizenga, D.E. and J.W. Szostak, *A DNA aptamer that binds adenosine and ATP.* Biochemistry, 1995. **34**(2): p. 656-65.

49. Weiss, S., et al., *RNA aptamers specifically interact with the prion protein PrP.* J Virol, 1997. **71**(11): p. 8790-7.

50. Smith, F.W., P. Schultze, and J. Feigon, *Solution structures of unimolecular quadruplexes formed by oligonucleotides containing Oxytricha telomere repeats.* Structure, 1995. **3**(10): p. 997-1008.

CHAPTER 3: HEAT CAPACITY EFFECTS IN THE UNFOLDING OF THE THROMBIN BINDING APTAMER

3.1 ABSTRACT

This study is designed to investigate the thermodynamic behavior of nucleic acid quadruplexes. The goal in this study is an enhanced understanding of the thermodynamics related to the formation of DNA quadruplexes in the presence of KCl. Experimental techniques used in this study are UV/CD spectroscopy, isothermal titration calorimetry (ITC) and gel electrophoresis. We performed UV, CD, and gel electrophoresis studies to confirm the molecularity of the structures and also to know the temperature range in which the quadruplex conformation is stable. ITC measurements were repeated at different temperatures to determine the temperature dependence of the enthalpy change accompanying quadruplex formation. Quadruplex formation was exothermic at high temperature but with a decreasing magnitude as the temperature decreased. These results are discussed in terms of the changes in heat capacity that accompany quadruplex formation.

3.2 INTRODUCTION

Heat capacity changes (ΔC_p) associated with structural folding transitions are prevalent in protein folding but are usually considered minor or negligible in thermodynamic studies in nucleic acid folding. However, with improved technology for thermodynamic measurements, along with more advanced understanding of nucleic acid folding, there are increasingly more nucleic acid studies addressing the significance of ΔC_p in thermodynamic analysis of nucleic acid structure formation[1, 2]. It is important to include ΔC_p terms in order to improve free energy calculations for nucleic acid structure prediction.

In light of the significant energetic contributions of ΔC_p, structural changes in nucleic acid are associated with changes in ΔC_p along with changes in free energy (ΔG), enthalpy (ΔH) and enthropy (ΔS). In addition, the sign and magnitude of ΔC_p may also suggest molecular details of the folding process.

In this ITC study of the thrombin-binding aptamer (TBA), along with UV melting and circular dichroism (CD) measurements, we show that ΔH is in fact temperature-dependent and the non-zero values of quadruplex formation are also comparable to the calculated ΔC_p, based on the changes in exposure of polar and nonpolar surfaces upon quadruplex unfolding.

3.3 MATERIALS AND METHODS

3.3.1 Sample Preparations

All DNA thrombin binding aptamer oligonucleotides, TBA (5'-GGTTGGTGTGGTTGG-3'), were purchased from IDT (Coralville, IA) with ion-exchange HPLC purification. Samples were annealed by heating up to 90 °C for 5 minutes, then cooling slowly to 15 °C in 10 mM Tris-acetate, pH 6.8, and with the addition of 25 mM or 140 mM KCl as indicated.

3.3.2 UV Measurements

Concentrations were determined spectrophotometrically, based on $\varepsilon 260$ = 147,920 L(mol strand-cm)$^{-1}$, estimated by nearest neighbor methods, using 1 cm path-length cells. UV melting curves were measured in a Cary 100 spectrophotometer equipped with a peltier heating device and run at a heating rate of 0.2 °C/min. Thermodynamic analysis of melting curves was carried out with an all-or-none model for quadruplex unfolding, assuming temperature-independent enthalpy and entropy changes.

3.3.3 Circular Dichroism (CD)

CD experiments were carried out on an Aviv 215 Circular Dichroism spectrophotometer equipped with a peltier temperature controller. All CD spectral scans were accumulated over the wavelength range 220-320 nm at 15 °C, unless otherwise noted. The scan of buffer alone was subtracted from each sample.

3.3.4 Isothermal Titration Calorimetry (ITC)

ITC experiments were carried out on the VP-ITC from Microcal. We have chosen to run the ITC experiments by titrating DNA in the syringe into metal in the cell instead of the more common way of ligands titrated into macromolecules. This is because the purpose of this work is to study the temperature dependence of the enthalpy change for quadruplex unfolding. Hence, we do not need the full energetics of the quadruplex formation, where free energies and binding constants can be measured. Here, we are only measuring the enthalpy change associated with each titration step and by titrating DNA to the excess metal in the cell, we are maintaining the condition of constant ionic strength. 10 μL of about 900 μM TBA in 10 mM Tris-Acetate buffer in the syringe were titrated into the calorimeter cell containing 1.4398 mL of either 25 mM or 140 mM KCl. In each ITC experiment, we take the average of the ΔH measured from five titration steps for experimental data collected at 15, 20, 25, 30, and 35°C.

Two control experiments were carried out to take into account the heat changes associated with processes other than the formation of quadruplexes. In the first control experiment, 10 μL of 900 μM TBA in 10 mM Tris-Acetate buffer in the syringe were titrated into the calorimeter cell containing 10 mM Tris-Acetate buffer. This is the typical control experiment done in ITC to account for dilution of the titrant. The second control experiment involved titrating 10 μL aliquots of 10 mM Tris-Acetate buffer in the syringe into cell containing 1.4398 mL of either 25 mM or 140 mM KCl. This was needed to account for dilution effects of the salt solution in the cell.

Monovalent metal ions such as potassium need to be present in excess amounts, in contrast to divalent metal ions (see chapter 4), in order to induce quadruplex formation. For this reason, titrations of potassium into a DNA solution are not practical.

3.4 RESULTS

3.4.1 ITC of TBA into 25 mM KCl

UV melting experiments were carried out to obtain the appropriate range of temperatures for ITC experiments. The lowest temperature was chosen such that there is no or minimal pre-existing quadruplex structures without the addition of K^+, whereas the highest temperature was chosen such that there is maximum quadruplex structure before it is thermally denatured. This is important because at low temperature, where there are pre-existing structures, there will be heat used to compensate the energy needed to undo the structure prior to formation of the quadruplex. At the high temperature range, where quadruplex structures are not fully formed, the enthalpy change measured with ITC would not be an accurate representation of the enthalpy change associated with the formation of quadruplexes. From the UV melt curves of 7 μM TBA in buffer only and in 25 mM of KCl in 10 mM Tris Acetate buffer (Figure 1), we have chosen a temperature range of 15-35 °C for these ITC studies. This is because at temperatures lower than 15 °C, there is a weak UV signal indicative of some preliminary structures even in the absence of added salt. As for TBA in the presence of 25 mM KCl, the UV absorption signal begins to decrease at 35 °C, which is indicative of the beginning of the melting transition.

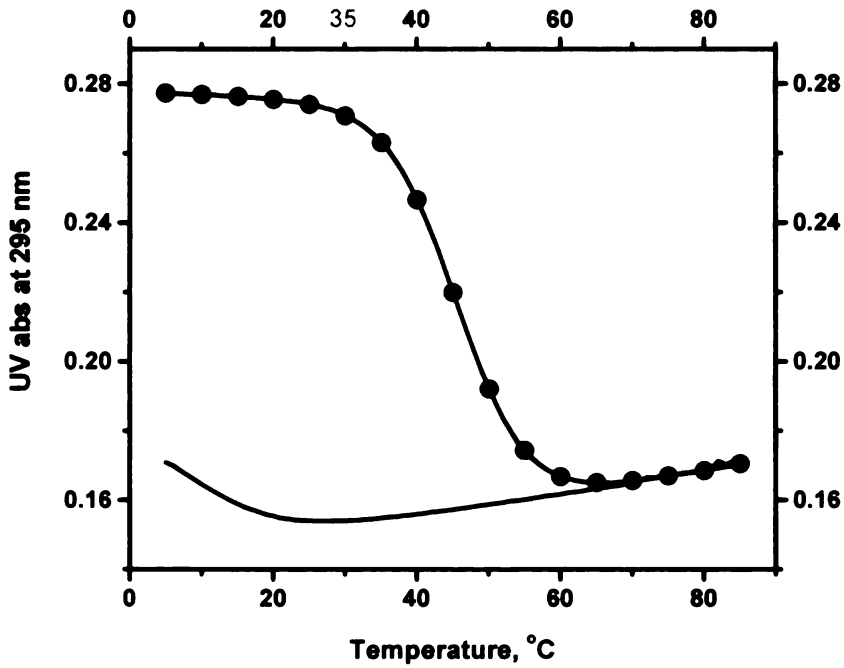

Figure 1: UV melting of 7 μM TBA in in 10 mM Tris-acetate, pH 6.8, only (—) and with 25 mM KCl (•)

Circular dichroism (CD) spectra can be used to assess the formation of quadruplex structures; hence only CD spectra of TBA in buffer at lower temperatures and CD spectra of TBA in 25 mM KCl at higher temperatures were obtained. The spectra (**Figure 2**) confirmed the minimal pre-existing structure at lower temperatures and maximum quadruplex structures at the higher temperatures.

ITC experiments allow us to measure the heat evolved during quadruplex formation as a function of the amount of titrant from the mixing of TBA and potassium solutions. **Figure 3** shows representative ITC data, where multistep addition of TBA in the syringe is made into a 25mM KCl solution in the cell. In Panel A1 and B1, each peak corresponds to the heat released on addition of an aliquot of TBA to KCl, which was recorded as the differential power signal with respect to time during the experiments. Integration of each peak with respect to time and normalization per mole of added TBA yields the apparent enthalpy change, ΔHapp (**Figure 3**: panel A2 and B2).

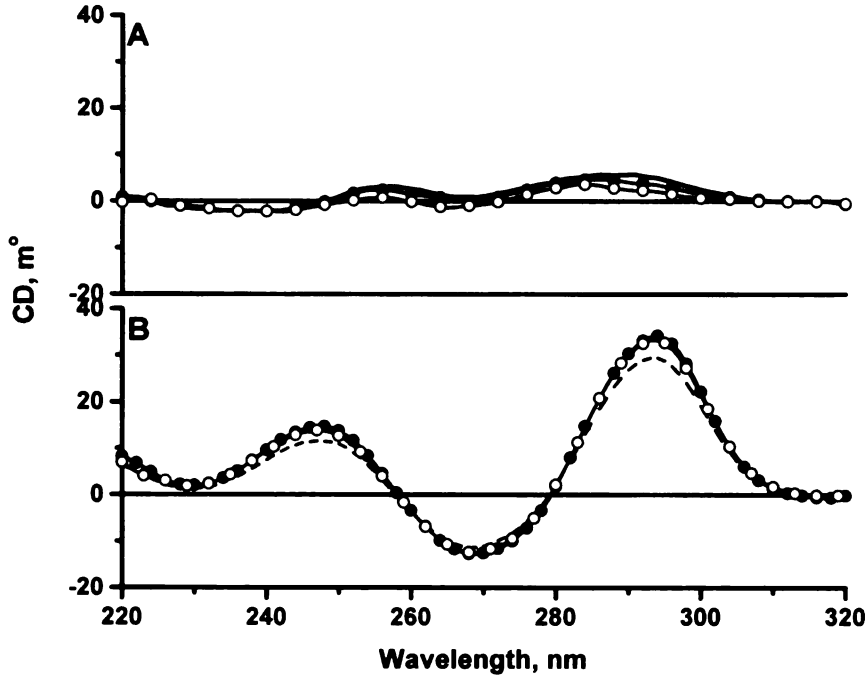

Figure 2: CD spectra of 7 μM TBA in10 mM Tris-acetate, pH 6.8, only (A) and with 25 mM KCl (B)

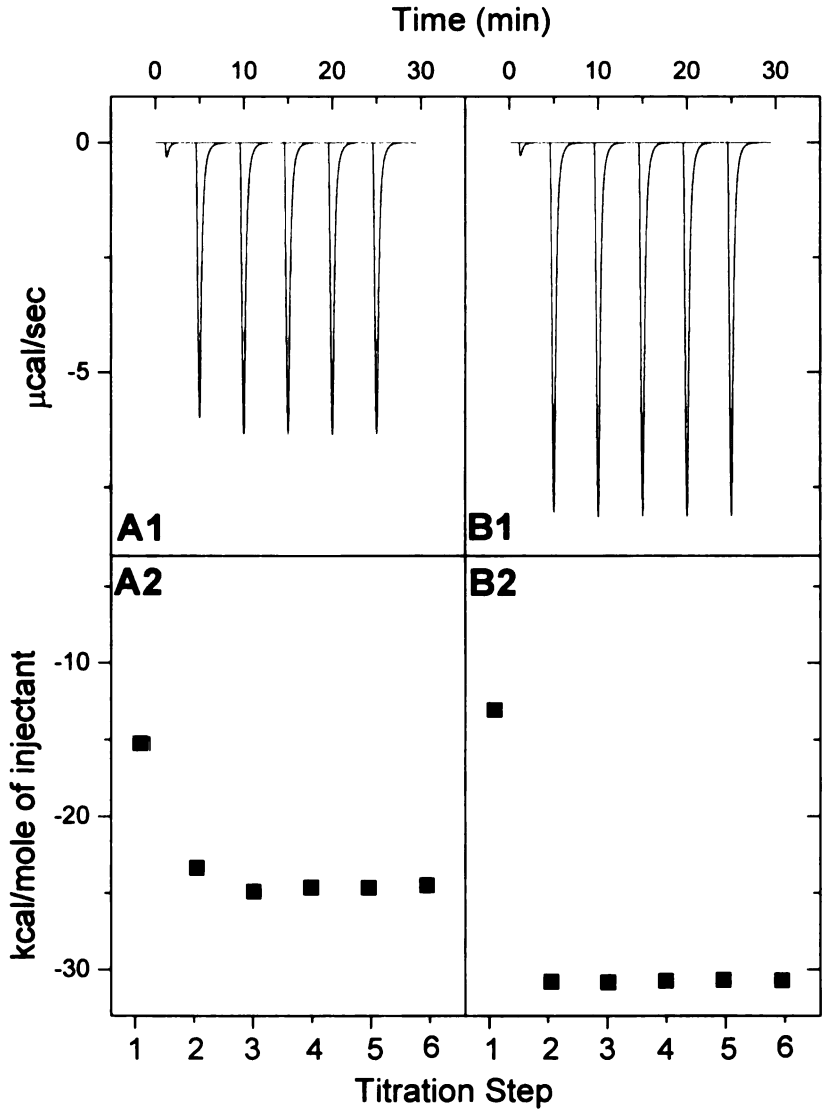

Figure 3: Representative ITC data of TBA in 10 mM Tris-acetate, pH 6.8, titrated into 25mM KCl. A:15°C and B: 30°C

Part of the measured ΔH_{app} is caused by the heat of dilution, typically for the solution in the syringe because each injection corresponds to about 150-fold dilution. However, there is evidence that the small dilution of metal in the cell actually has a significant heat effect as well [3, 4]. Hence, we include a second control experiment for the metal dilution effect, where buffer in the syringe is titrated into KCl. Hence the enthalpy change of dilution of TBA, ΔH_{dilDNA}, is measured by the control experiments of titrating DNA in buffer alone into the cell containing buffer only, while the enthalpy change for dilution of KCl is measured by titrating buffer into the KCl solution. These effects are related to the enthalpy of quadruplex formation, ΔH_{fold}, by the equation:

$$\Delta H_{app} = \Delta H_{fold} + \Delta H_{dilDNA} + \Delta H_{dilKCl}$$

The ΔH_{dilDNA} and ΔH_{dilKCl} of the control experiments were subtracted from the ΔH_{app} of the ITC experiments. **Figure 4** shows the ΔH_{fold} for TBA quadruplex formation vs. titrant addition. The average ΔH_{fold} values at different temperatures were then used to estimate the positive heat capacity change, ΔC_p, for the unfolding of TBA.

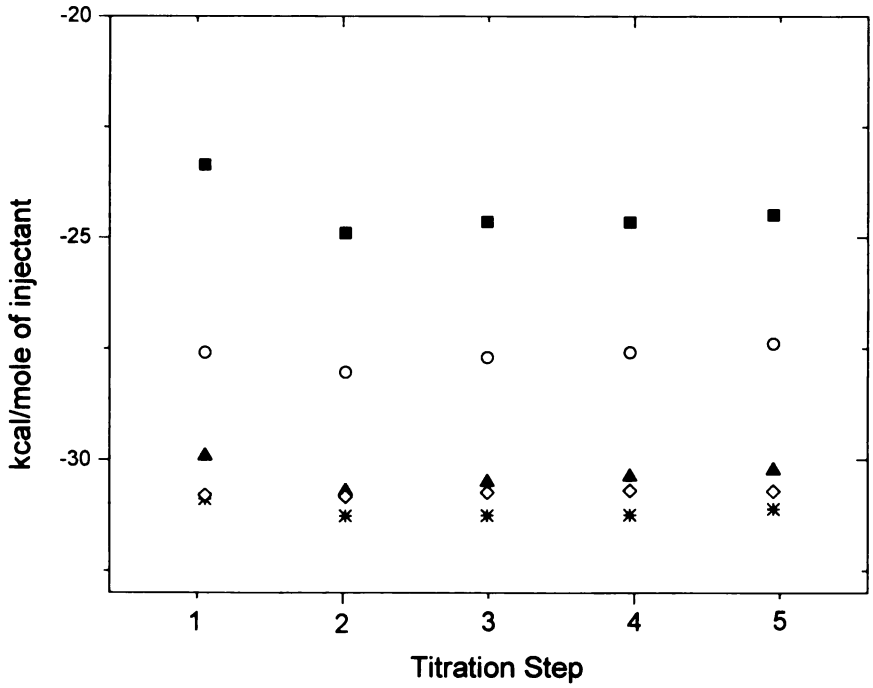

Figure 4: ΔH for TBA quadruplex formation in 25 mM KCl vs. titration step at different temperatures: 15°C (■), 20°C (○), 25°C (▲), 30°C (*), and 35°C (◊)

C_p is defined as amount of heat required to raise the temperature of a substance by one degree at constant pressure. C_p can be a measure of the temperature dependence of the enthalpy or entropy.

$$C_p = \left[\frac{\Delta q}{\Delta T}\right]_p = \left[\frac{\partial H}{\partial T}\right]_p = \left[T\frac{\partial S}{\partial T}\right]_p$$

Hence, changes in C_p associated with the folding or unfolding of a biopolymer can be expressed as:

$$\Delta C_p = \left[\frac{\partial \Delta H}{\partial T}\right]_p = \left[T\frac{\partial \Delta S}{\partial T}\right]_p$$

A common experimental method for determining ΔC_p is by measuring ΔH as a function of temperature as in ITC experiments (**Figure 3**). The ΔH_{fold} at different temperatures in **Figure 4** were averaged and these values were used in plotting ΔH_{fold} as a function of temperature as shown in **Figure 5**.

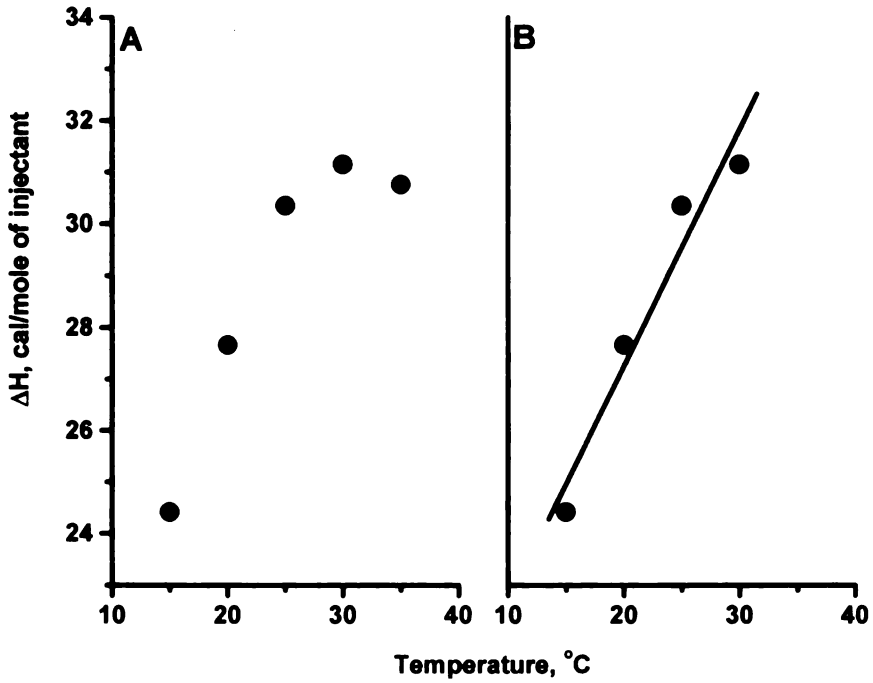

Figure 5: ΔH vs. temperature plot for TBA in 25 mM KCl. A: ΔH for temperature 15-35 °C. B: Linear regression of the ΔH vs. temperature plot, $\Delta Cp = 458$ cal $K^{-1}mol^{-1}$

Since the ΔH values measured from the ITC experiments were for quadruplex folding, the values with opposite signs were plotted against various temperatures in order to obtain the ΔC_p for the unfolding process. ΔC_p was obtained from the slope of the linear regression of the ΔH vs. temperature plot. In **Figure 5**A, the last data point appeared to be an outlier and hence, in order to get a better linear regression, it was omitted. From the UV melting curves in Figure 1, 35°C falls at the very beginning of the transition. Hence, it is possible that at this temperature, not all quadruplexes are formed, which is why there is less heat evolved. There is also a possibility that ΔC_p is not constant over different temperature ranges.

3.4.2 ITC of TBA into 140 mM KCl

Since the range of temperatures for the study with 25 mM KCl was limited to only 15-30 °C, the same study was repeated with higher salt concentration so that a broader range of temperatures can be used in order to get a better linear fit of the enthalpy change.

UV melting experiments were repeated at the same condition, except with a higher concentration of 140 mM KCl. As shown in **Figure 6**, the melting transition is about to start at 40 °C, hence the temperature range 15-40 °C, which is also supported by the CD spectra in **Figure 7**.

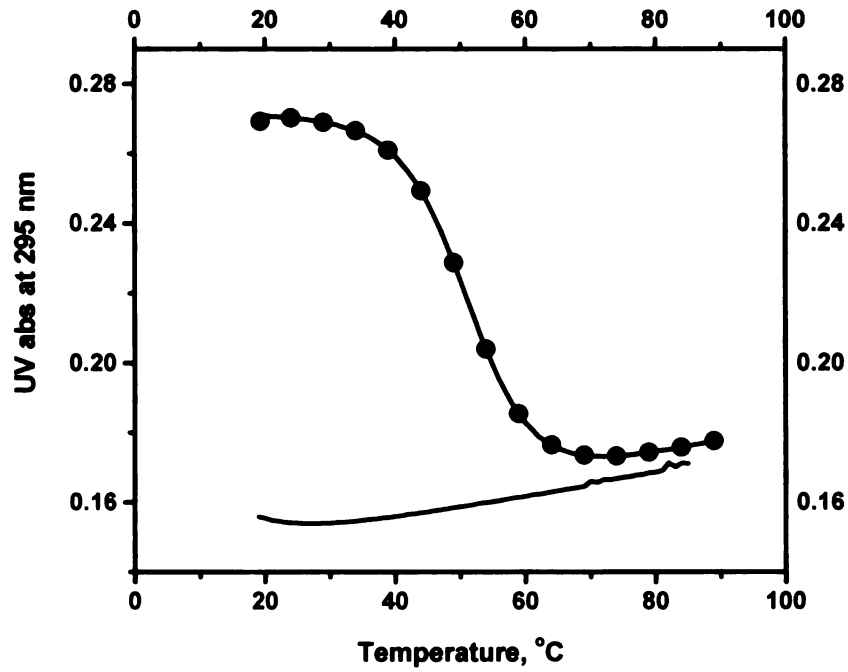

Figure 6: UV melt of 7 μM TBA in in 10 mM Tris-acetate, pH 6.8, only (—) and with 140 mM KCl (•)

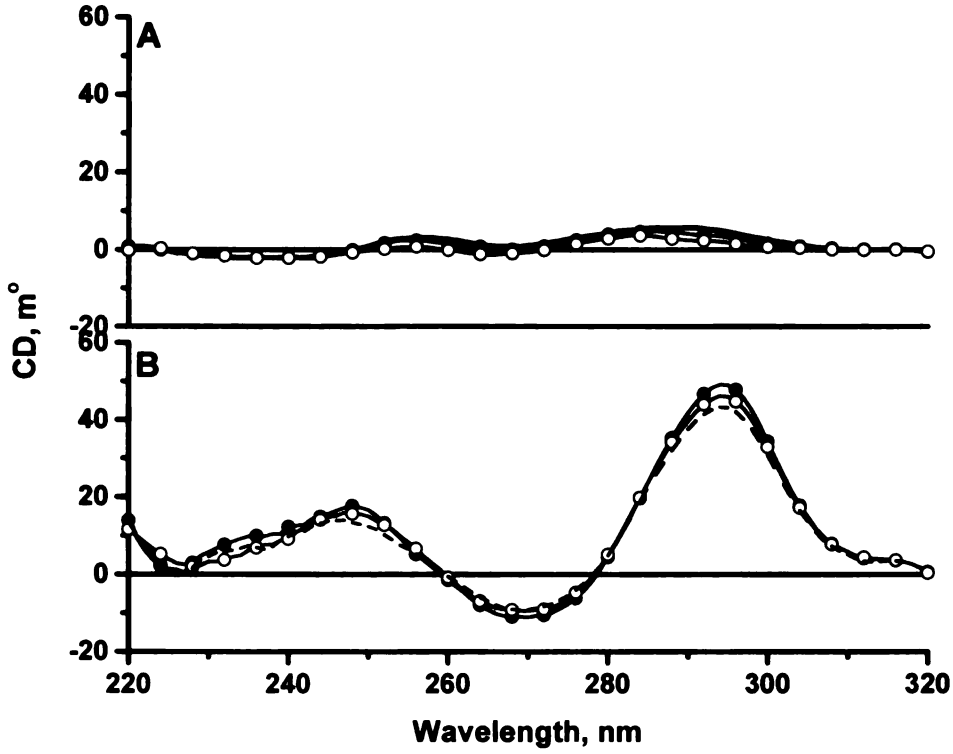

Figure 7: CD spectra of TBA in in 10 mM Tris-acetate, pH 6.8, only (A) and with 140 mM KCl (B) at different temperatures

Figure 8 shows representative ITC data, where multistep additions of TBA in the syringe were made into 140 mM KCl in the cell. Once again, integration of peaks in Panels A1 and B1 with respect to time and normalization per mole of added TBA yields the apparent heat change, ΔH_{app}. The study was carried out the same as before, with similar control experiments except for the higher concentration of KCl at 140 mM. ΔH for TBA quadruplex formation and controls vs. titrant additions obtained from the ITC experiments were also plotted (**Figure 9**) on the same graph and the average net ΔH values were again used for linear fitting (**Figure 10**) in order to get a value of 340 cal $K^{-1}mol^{-1}$.

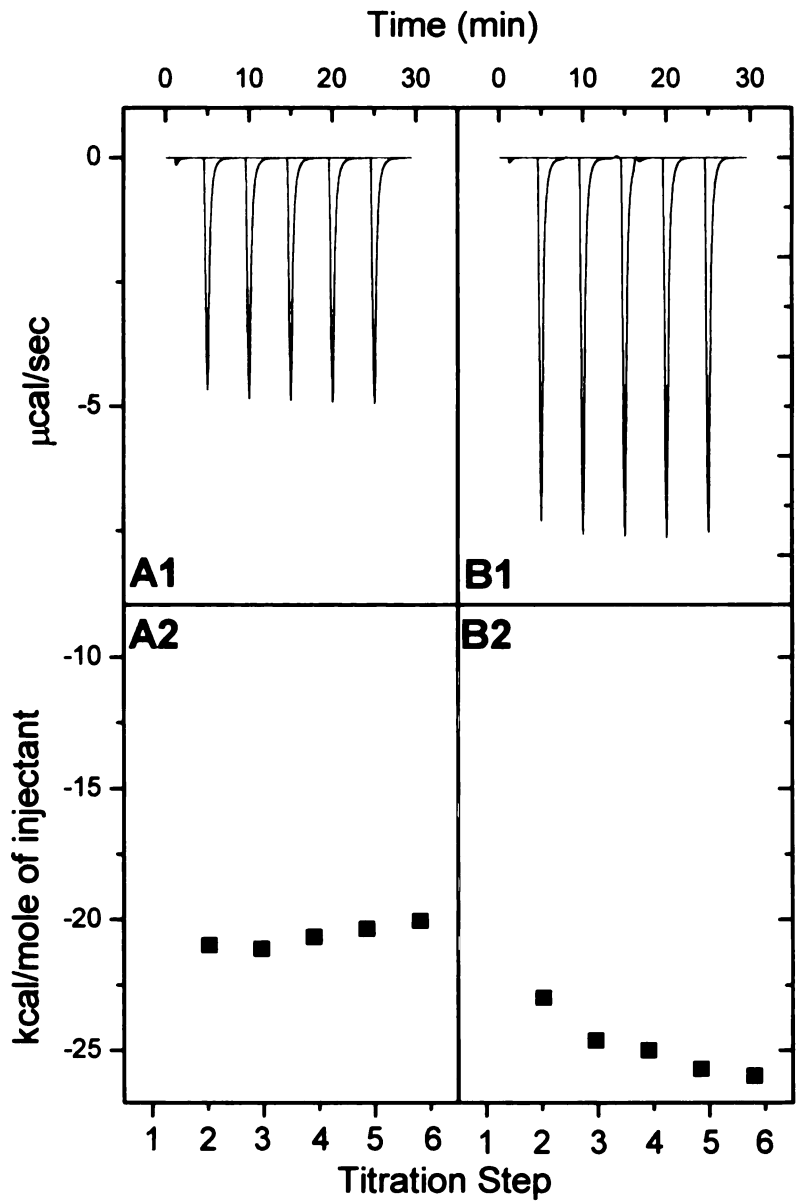

Figure 8: Representative ITC data of TBA in 10 mM tris-acetate, pH 6.8, titrated into 140 mM KCl. A: Titration at 15°C and B: 30°C.

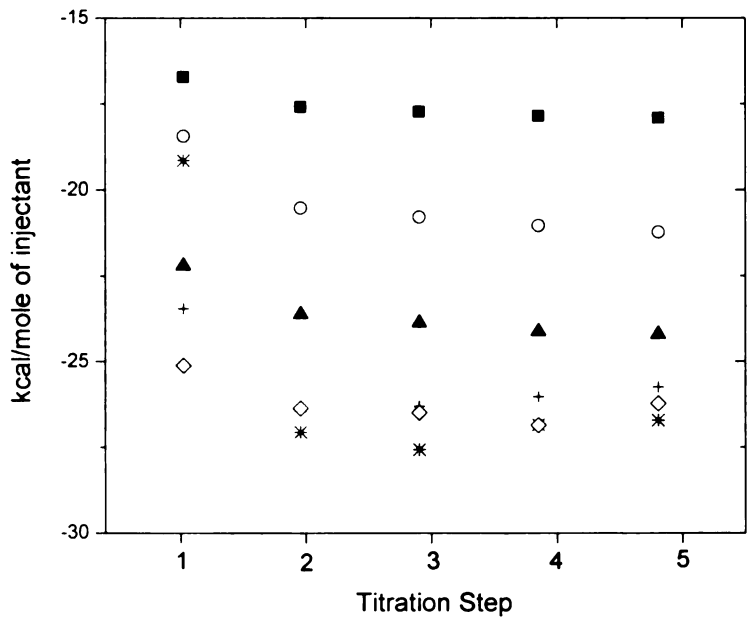

Figure 9: ΔH for TBA quadruplex formation in 140 mM KCl vs. titration step at different temperatures: 15°C (■), 20°C (○), 25°C (▲), 30°C (+), 35°C (◊) and 40°C(*)

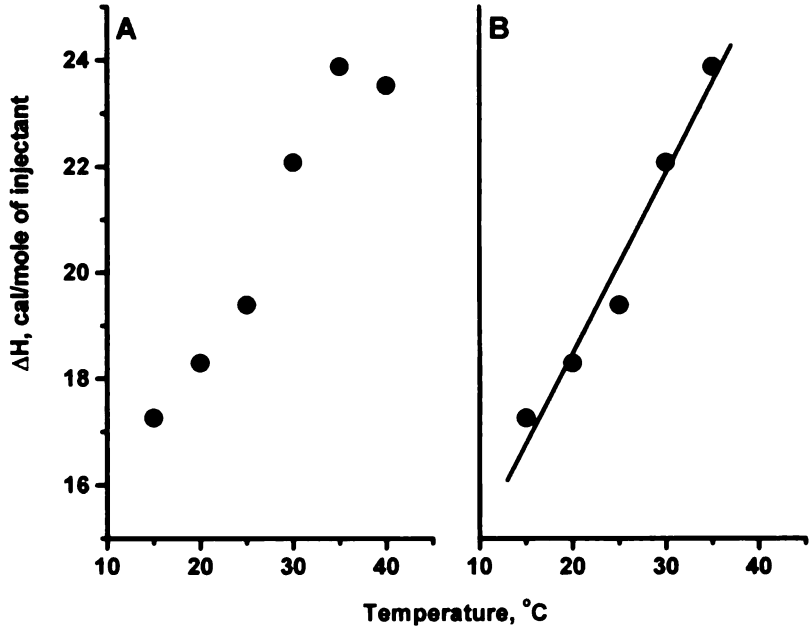

Figure 10: ΔH vs. temperature plot for TBA in 140 mM KCl. A: ΔH for temperature 15-35 °C. B: Linear regression of the ΔH vs. temperature plot, ΔC_p = 340 cal $K^{-1}mol^{-1}$

3.5 DISCUSSION

Reported ΔC_p values for nucleic acid structural transitions ranges widely from zero to values as much as −100 cal K^{-1}base $pair^{-1}$ [5]. This large variation in ΔC_p has significant impact on our ability to predict the structural stability of nucleic acids quantitatively and accurately from sequence. Most of the predictions for nucleic acid stability are base on the assumption that $\Delta C_p = 0$, [6, 7]; but if ΔC_p is on the order of a larger non-zero value, such as the 400 cal $K^{-1}mol^{-1}$ we obtained in our studies of unfolding of the TBA quadruplex, then the prediction of enthalpy change will be 4 kcal $K^{-1}mol^{-1}$ more for every 10 K increase in temperature.

ΔC_p has long been considered in studies of protein folding; which is predominantly originates from solvent effects, including both solute-solvent and solvent-solvent interactions. In addition to solvent effects, the ΔC_p of nucleic acid folding also derives from internal solute effects, such as conformational entropy, changes within coupled states, electrostatic, vibrational modes, and others [6].

As in protein folding, solvent effects are predominantly viewed as the major causes in the ΔC_p of nucleic acid folding as differential solvation of polar and nonpolar surfaces is much easier to probe experimentally than internal solute effects. Hence, ΔC_p of TBA unfolding was computed from the change in polar and nonpolar surfaces as described below.

The prediction of ΔC_p can be calculated based on the empirical relationship between changes in water-accessible surface area and methods for

doing this have been developed by several groups [8, 9]. The resulting expression, used to compute the ΔC_p for the unfolding of TBA, based on data for the ΔC_p of transfer of small hydrocarbons and amides from water to pure liquid phase [9, 10] is:

$$\Delta C_p = 0.32\Delta A_{np} - 0.14\ \Delta A_p \qquad (1)$$

ΔA_p and ΔA_{np} represent changes in surface area of polar and nonpolar water-accessible surfaces respectively, which we obtained from the GETAREA web server (version 1.1) [11], provided by the Sealy Center for Structural Biology at the University of Texas Medical Branch. This service allows users to obtain solvent accessible surface area or solvation energy using their program FANTOM (Fast Newton-Raphson Torsion Angle Minimizer), which calculates low-energy conformations of macromolecules [11] by specifying data files containing Cartesian coordinates in the format of Protein Data Bank (pdb) file.

A_p and A_{np} for the TBA quadruplex structures were obtained from GETAREA with the pdb files from NMR studies of solution structure of TBA by Feigon's and Bolton's groups [12, 13]. Individual nucleotides and nucleosides of TBA were used to compute A_p and A_{np} for the unfolded states. The ΔA_p and ΔA_{np} were used to compute ΔC_p with equation (1). As shown in Table 1, the average ΔC_p is about 316 cal $K^{-1}mol^{-1}$. This value is in close agreement with the experimental ΔC_p values, which are 458 cal $K^{-1}mol^{-1}$ 25 mM KCl and 340 cal $K^{-1}mol^{-1}$ for 140 mM KCl. Since individual nucleosides and nucleotides were used to calculate the water-accessible polar and nonpolar surface area, the

contribution of solvent effects calculated from equation (1) is the maximum possible.

	A_{p_fold}	A_{np_fold}	A_{p_unfold}	A_{np_unfold}	ΔA_p	ΔA_{np}	ΔC_p
bolton	1183.4	1510.9	3690.6	3364.4	2507.2	1853.4	242.09
feigon1	1467.2	1377.6	3690.6	3364.4	2223.4	1986.7	324.48
feigon2	1490.9	1366.7	3690.6	3364.4	2199.8	1997.6	331.28
feigon3	1461.5	1376.1	3690.6	3364.4	2229.1	1988.3	324.17
feigon4	1484	1364.8	3690.6	3364.4	2206.6	1999.6	330.95
feigon5	1474.1	1427.5	3690.6	3364.4	2216.5	1936.8	309.47
feigon6	1454.7	1360.6	3690.6	3364.4	2235.9	2003.8	328.19
feigon7	1505.8	1390.4	3690.6	3364.4	2184.8	1974	325.81
feigon8	1439.8	1384.2	3690.6	3364.4	2250.8	1980.1	318.52
feigon9	1410.8	1384.2	3690.6	3364.4	2279.8	1980.1	314.46
feigon10	1445.1	1370.9	3690.6	3364.4	2245.5	1993.5	323.54
feigon11	1441.7	1396.2	3690.6	3364.4	2248.9	1968.1	314.95
feigon12	1457.7	1396.5	3690.6	3364.4	2232.9	1967.9	317.11
dGMP	277.46	206.13					
dTMP	219.09	244.61					
dG	156.39	247.65					

Table 1: Calculations of ΔC_p from changes in water-accessible polar and nonpolar surface area.

Due to the highly polyanionic nature of nucleic acids, the contributions to ΔC_p for nucleic acid folding may be different from those of other biopolymers. In addition to solvent effects, the presence of pre-existing structures that must be undone has been observed to play an important role for ΔC_p for nucleic acid folding [2, 5].

At lower temperatures, more heat is required to undo the pre-existing structures; hence this reduces the net heat generated from the formation of the quadruplex structures. However, at higher temperatures, the pre-existing

structures are less stable, requiring less heat to undo structures prior to quadruplex formation. Hence the net ΔH is greater in magnitute at higher temperature.

3.6 CONCLUSIONS

The enthalpy of quadruplex folding/unfolding is temperature dependent and we can determine the corresponding heat capacity change (ΔC_p). As in the case of protein, unfolding of DNA results in positive ΔC_p. Values of ΔC_p can be calculated theoretically using established methods, which only accounts for the contribution from changes in water-accessable polar and nonpolar surface areas. In the case of TBA, there is a good agreement between the measured and calculated ΔC_p values. Therefore, the contributions of pre-existing structures to ΔC_p in the unfolding of TBA quadruplexes are minor.

3.7 REFERENCES

1. Mikulecky, P.J. and A.L. Feig, Heat capacity changes associated with DNA duplex formation: salt- and sequence-dependent effects. Biochemistry, 2006. 45(2): p. 604-16.
2. Mikulecky, P.J. and A.L. Feig, Heat capacity changes associated with nucleic acid folding. Biopolymers, 2006. 82(1): p. 38-58.
3. Young, T.F. and O.G. Vogel, The Relative Heat Contents of the Constituents of Aqueous Sodium Chloride Solutions. Thermochemistry of Sodium Chloride Solutions, 1932. 54: p. 3030-40.
4. Young, T.F. and J.S. Machin, Heat Content and Heat Capacity of Aqueous Sodium Chloride Solutions. 1936. 2254-60.
5. Holbrook, J.A., et al., Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: interpretation in terms of coupled processes of formation and association of single-stranded helices. Biochemistry, 1999. 38(26): p. 8409-22.
6. Sturtevant, J.M., Heat capacity and entropy changes in processes involving proteins. Proc Natl Acad Sci U S A, 1977. 74(6): p. 2236-40.
7. Cate, J.H., et al., Crystal structure of a group I ribozyme domain: principles of RNA packing. Science, 1996. 273(5282): p. 1678-85.
8. Privalov, P.L. and G.I. Makhatadze, Contribution of hydration and non-covalent interactions to the heat capacity effect on protein unfolding. J Mol Biol, 1992. 224(3): p. 715-23.

9. Spolar, R.S. and M.T. Record, Jr., Coupling of local folding to site-specific binding of proteins to DNA. Science, 1994. 263(5148): p. 777-84.

10. Myers, J.K., C.N. Pace, and J.M. Scholtz, Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. Protein Sci, 1995. 4(10): p. 2138-48.

11. Fraczkiewicz, R. and W. Braun, Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. J. Comp. Chem, 1998. 19(3): p. 319.

12. Macaya, R.F., et al., Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. Proc Natl Acad Sci U S A, 1993. 90(8): p. 3745-9.

13. Wang, K.Y., et al., A DNA aptamer which binds to and inhibits thrombin exhibits a new structural motif for DNA. Biochemistry, 1993. 32(8): p. 1899-904.

CHAPTER 4: THERMODYNAMIC CHARACTERIZATION OF THE THROMBIN-BINDING DNA APTAMER

4.1 ABSTRACT

DNA quadruplex structures based on the guanine quartet are typically stabilized by monovalent cations such as K^+, Na^+, or NH_3^+. Certain divalent cations can also induce quadruplex formation, such as Ba^{2+}, Pb^{2+}, and Sr^{2+}. Here we studied the binding of these divalent cations with unusually high affinity to the thrombin binding aptamer, d(GGTTGGTGTGGTTGG), inducing a unimolecular folded structure.

At micromolar concentrations the binding is stoichiometric in folding the aptamer. The binding constants and other thermodynamic parameters can be determined by Isothermal Titration Calorimetry and Circular Dichroism titration. Once again, nonzero ΔC_p values are obtained in these measurements.

4.2 INTRODUCTION

It has become increasingly evident that the quadruplex-forming sequences are able to adopt various types of conformations and structures [1-5]. Hence, energetic information associated with different conformations is crucial in understanding the thermodynamic and kinetic stability of different quadruplex structures under various solution conditions. From such information, one can develop important interrelationships between the structures and potential biological functions.

The common factors affecting quadruplex stability include variation in strand, loop, cation, and pH. In this chapter, the main focus is to investigate specific metal cation - DNA interactions, particularly the divalent cations. This specificity of metal-binding in quadruplex formation is unique, as opposed to the looser counterion condensation binding behavior of most monovalent metals with nucleic acids in general. Though monovalent cations, such as the commonly used K^+, are specific in quadruplex formation, divalent cations have more apparent specific binding interactions, with quadruplex-forming nucleic acid sequences due to their greater affinity for binding [6-12].

Several studies have measured the thermodynamic stability of quadruplex structures in the presence of fixed excess cation concentrations using thermal denaturation [8, 13]. However, under these conditions, the secondary and non-specific binding of the excess cations can complicate the interpretation of the results.

The focus of this study is to explore the thermodynamic property of specific binding of divalent cations in quadruplex formation. The objective is to investigate this specific binding in by directly measuring their DNA binding affinity directly. Several previous thermodynamic studies on the thrombin binding aptamer (TBA) with differential scanning calorimetry have resulted in some difference in interpretation of the measured values [8, 13].

In this study, stoichiometric titration of divalent metals into TBA, in which the folding process was monitored by circular dichroism (CD), allowed us to measure the binding affinity by the analysis of these titration curves. The binding constants are then compared to the thermodynamic parameters obtained from isothermal calorimetry (ITC).

4.3 MATERIALS AND METHODS

4.3.1 Sample Preparation

All DNA thrombin binding aptamer oligonucleotides, TBA (5'-GGTTGGTGTGGTTGG-3'), were purchased from IDT (Coralville, IA) with ion-exchange HPLC purification. Samples were annealed by heating up to 90 °C for 5 minutes, then cooling slowly to 15 °C in 10 mM Tris-acetate, pH 6.8.

Divalent metals (Ba^{2+}, Pb^{2+}, Sr^{2+}) were prepared with 10 mM Tris-acetate, pH 6.8, which is the same buffer used in TBA solution.

4.3.2 UV Measurement

Concentrations were determined spectrophotometrically, based on ε_{260} = 147,920 L(mol strand-cm)$^{-1}$, estimated by nearest neighbor methods [14], using 1 cm path-length cells.

4.3.3 Circular Dichroism (CD)

CD experiments were carried out on an Aviv 215 Circular Dichroism spectrophotometer equipped with a peltier temperature controller. All CD spectral scans were accumulated over the wavelength range 220-360 nm at 10 °C, 20 °C, and 30 °C. The scan of buffer alone was subtracted from each sample. In CD titration experiments, metal ions were titrated into TBA in small increments. The quadruplex formation was monitored by following the CD signal at 301 nm for the binding of Ba^{2+}. The titration data was then fit to a single binding site model to yield the binding constant.

4.3.4 Isothermal Titration Calorimetry (ITC)

ITC experiments were carried out on the VP-ITC from Microcal. Divalent metal in the syringe was titrated into the calorimeter cell containing 1.44 mL of about 7 μM TBA in 10 mM tris acetate buffer. Experiments were done at various temperatures namely 15, 20, 25, and 30°C. In the first set of control experiments, divalent metal in the syringe was titrated into the calorimeter cell containing 10 mM tris acetate buffer. The other control experiments were done by titrating 10 mM tris acetate buffer into the calorimeter cell containing TBA in buffer. Two control experiments at corresponding temperatures were carried out to take into account the heat changes associated with processes other than the formation of quadruplexes.

4.4 RESULTS

Circular dichroism (CD) spectra can be used to assess the nature of the quadruplex fold, based on observations that antiparallel quadruplexes typically present a long-wavelength positive maximum near 293 nm for TBA in the presence of Na^+ or K^+ [15, 16]. The effect of adding barium acetate to TBA was followed by CD measurements, as displayed in Figure 1. In the absence of metal ion, the CD spectrum is of relatively low amplitude, characteristic of oligonucleotides possessing little or no structure (Figure 5A in Chapter 2). Addition of Ba^{2+} results in a significant change in the CD spectra. As seen in many cases of quadruplexes formed with the addition of divalent cations, the long wavelength of CD is shifted, relative to that observed in the presence of K^+, by almost 10 to 20 nm to about 301 – 312 nm . The other significant difference between monovalent cations and divalent cations is that much lower concentrations of divalent cations are needed to complete the transition (see Chapter 3). Addition of micromolar range concentration of Ba^{2+} results in a long wavelength maximum at 301 nm (Figure 1). As we follow the long CD wavelength during the titration, the nice isosbestic points are good indications of an all or none model for the formation of this quadruplex.

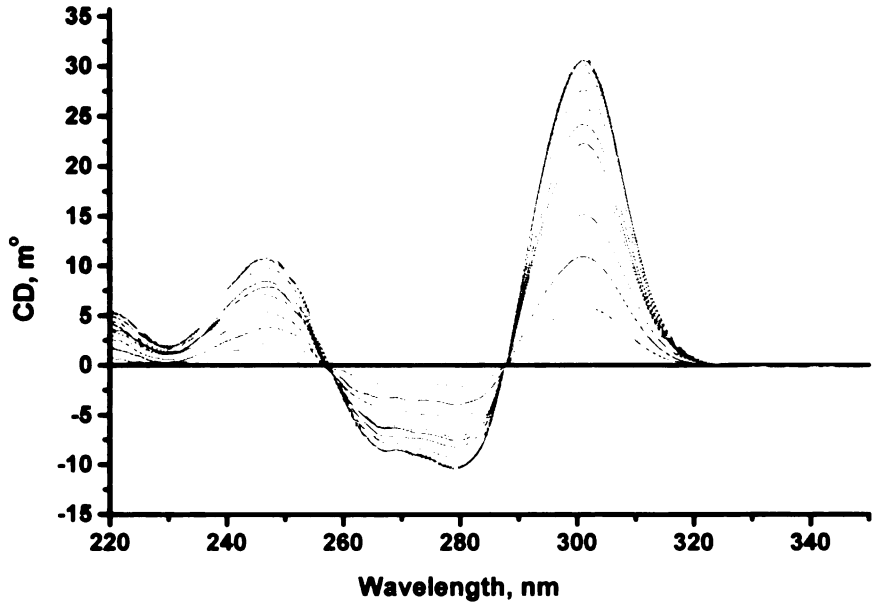

Figure 1: CD spectra of TBA with the addition of barium acetate, both in 10 mM tris acetate buffer, pH 6.8, at 30°C

Figure 2 represents the results of the CD titration of TBA by the addition of barium acetate, followed at 301 nm. At 10 °C, it can be seen that the binding of Ba^{2+} is very tight and the binding constant can only be estimated to be about 1.2 × 10^7. Figure 3 and Figure 4 are the same CD titrations repeated at 20 °C and 30 °C.

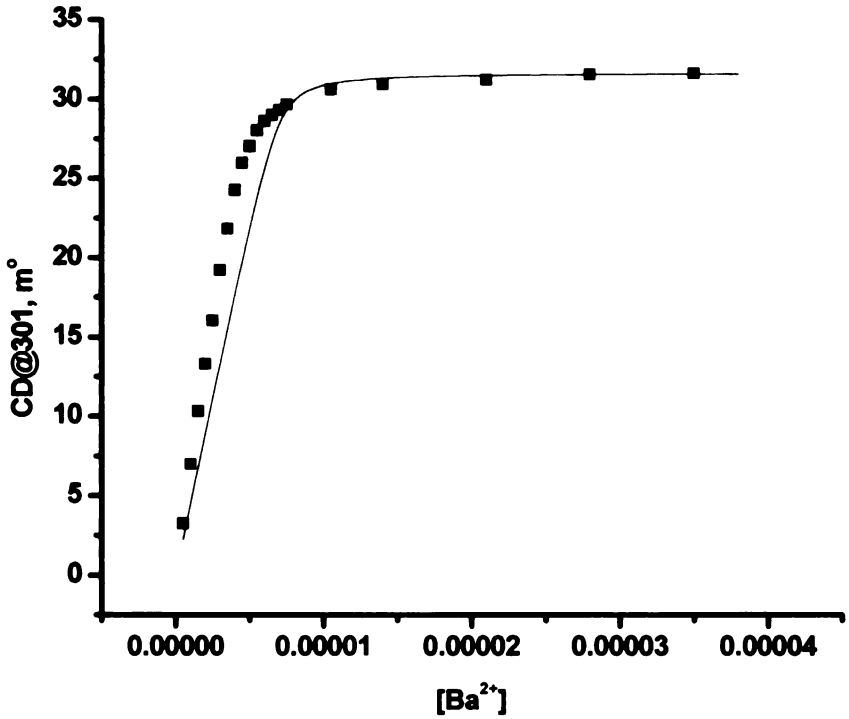

Figure 2: CD titration of TBA with barium acetate at 10 °C

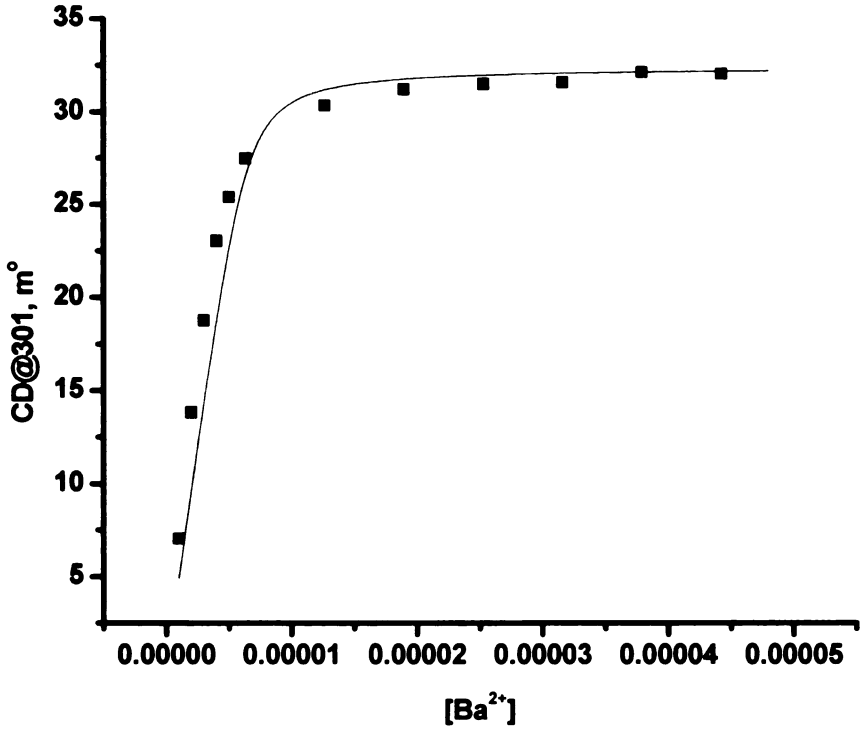

Figure 3: CD titration of TBA with barium acetate at 20 °C

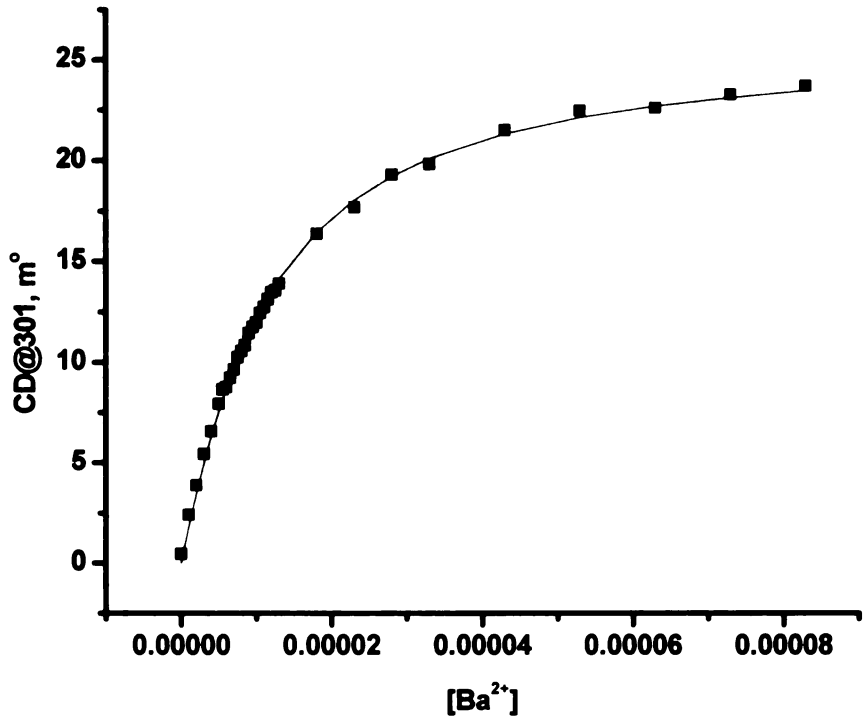

Figure 4: CD titration of TBA with barium acetate at 30 °C

Binding constants, K, can be estimated from fitting the CD titration curves above. The estimated K values are presented in Table 2 along with K values obtained from ITC experiments below.

ITC experiments were carried out to measure the heats of quadruplex formation at different temperatures. Figure 5 and Figure 6 are representative ITC data, involving multistep addition of divalent metals into TBA in 10 mM tris acetate. Control experiments were carried out to take into account the heat changes associated with processes other than the formation of quadruplexes.

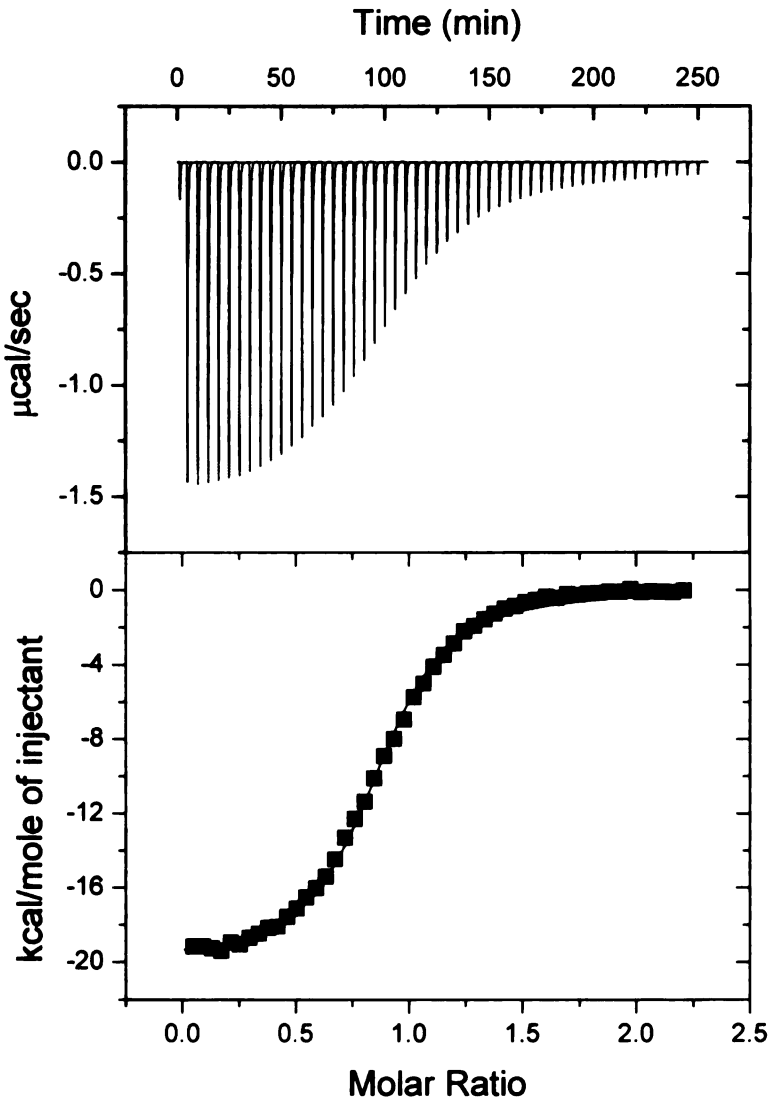

Figure 5: Representative ITC data of 615 mM barium acetate titrated into 51 µM TBA in 10 mM Tris-acetate, pH 6.8

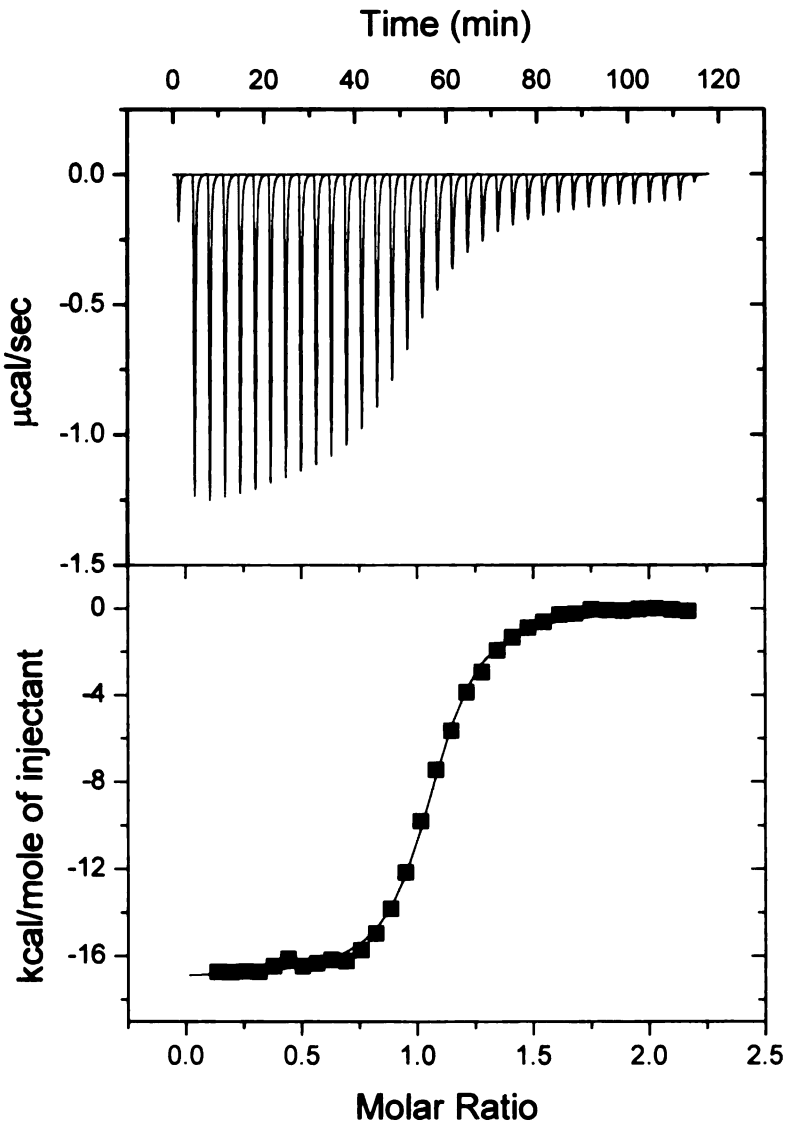

Figure 6: Representative ITC data of 200 mM lead nitrate titrated into 18.7 µM TBA in 10 mM Tris-acetate, pH 6.8

metal	[TBA] (μM)	Temperature (°C)	N	K (×10^5)	ΔH (×10^4)	ΔS
Ba^{2+}	51.22	15	0.83	23.71	-2.04	-41.61
Ba^{2+}	51.22	20	0.86	16.65	-2.24	-48.07
Ba^{2+}	51.22	25	0.88	8.34	-2.50	-56.60
Ba^{2+}	51.22	30	0.89	3.32	-2.83	-68.06
Pb^{2+}	21.22	15	1.23	25.40	-1.37	-18.28
Pb^{2+}	21.22	20	1.23	15.31	-1.92	-37.11
Pb^{2+}	21.22	25	1.00	15.39	-2.44	-53.57
Pb^{2+}	21.22	30	0.61	10.70	-2.72	-61.96
Pb^{2+}	21.22	35	0.58	7.76	-2.78	-63.16
Sr^{2+}	16.6	25	0.75	12.54	-2.04	-40.45
Sr^{2+}	19.5	25	0.57	7.66	-2.38	-52.81
Sr^{2+}	19.5	25	0.69	9.57	-2.02	-40.53

Table 1 Thermodynamic parameters obtained from ITC experiments.

Table 1 lists all the thermodynamic parameters obtained from ITC experiments. Note that when Ba^{2+} or Pb^{2+} is titrated into TBA, the binding constant decreases as temperature increases.

As in the case of titrating monovalent cation, such as K^+ as described in Chapter 3, the heat evolved during quadruplex formation of TBA with divalent cations increases as the temperature increases. Once again, it is shown that ΔH is temperature dependent. ΔC_p is determined by measuring the ΔH as a function of temperature in the ITC experiments above. ΔC_p was determined from the

slope of the linear regression of the ΔH vs. temperature plots as shown in Figure 7 and Figure 8.

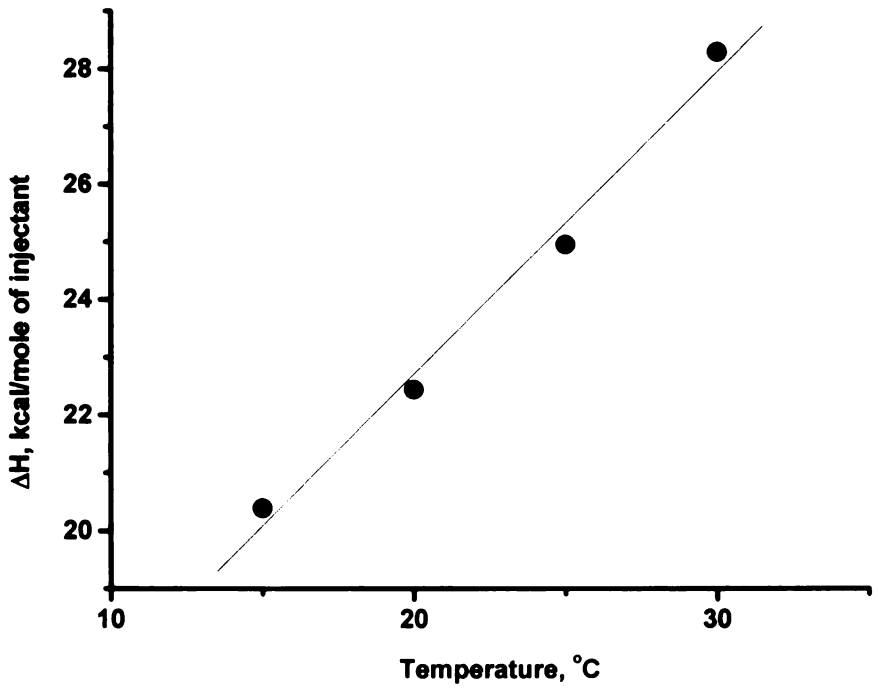

Figure 7: ΔH vs. temperature plot for TBA-barium quadruplex. ΔCp = 524 $calK^{-1}mol^{-1}$

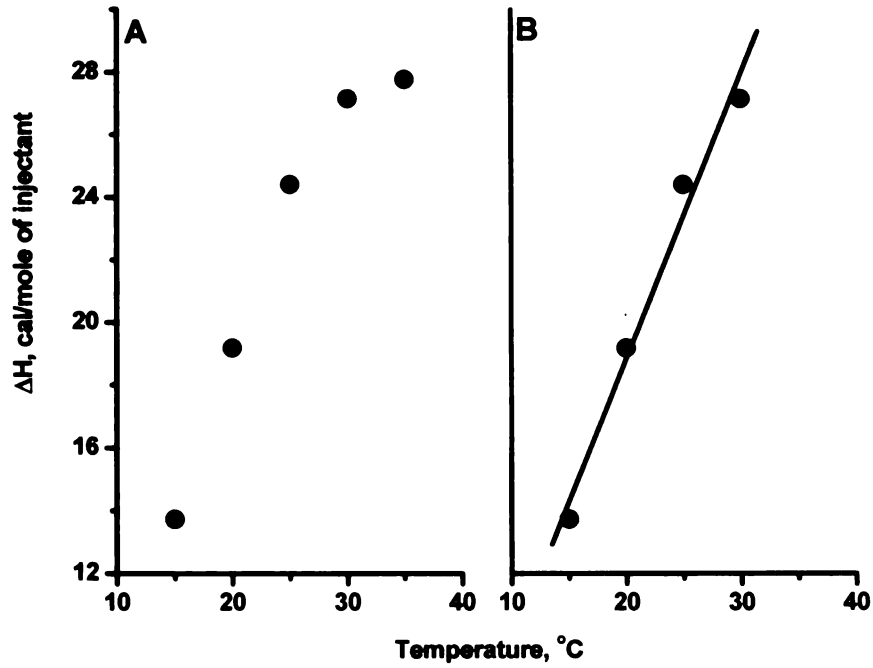

Figure 8: ΔH vs. temperature plot for TBA-lead quadruplex. ΔCp = 910 $calK^{-1}mol^{-1}$

Temperature (°C)	K (CD)	K (ITC)
10	1.2×10^7	-
15	-	2.37×10^6
20	4×10^6	1.66×10^6
25	-	8.34×10^5
30	1.17×10^5	3.32×10^5

Table 2. K for binding of Ba^{2+} and TBA.

All the binding constants were used to make a van't Hoff plot in Figure 9.

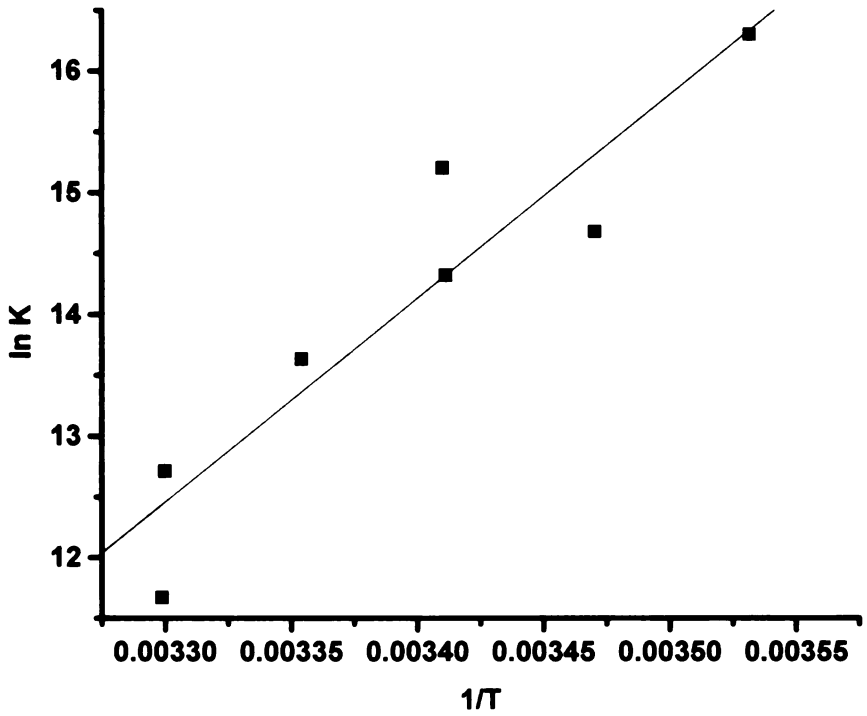

Figure 9: van 't Hoff plot for binding constants of Ba^{2+} to TBA obtained from CD and ITC

In Figure 9, the slope from the van't Hoff plot is 11, 468 K. From the van't Hoff equation, namely

$$\ln(K) = -\frac{\Delta H^{\circ}}{R}\left[\frac{1}{T}\right] + \frac{\Delta S^{\circ}}{R}$$

The slope ΔH can be estimated from multiplying the slope with negative R, which is 1.98 cal K^{-1} mol^{-1}. Thus, we obtain

ΔH = - 1. 98 cal K^{-1} mol^{-1} × 11,468 K = -22.7 kcal mol^{-1}

This value is in a good agreement with the average ΔH values from ITC data of Ba^{2+} titrated into TBA, which is about -24 kcal mol^{-1}

4.5 DISCUSSION

Thermodynamic characterization of the quadruplex formation process in solution is crucial in understanding the stability of the structures in solution. One of the objectives of thermodynamic studies is to obtain precise equilibrium constants and changes of thermodynamic quantities including enthalpies and entropies [17]. CD titration and ITC are two common techniques used in thermodynamic studies.

CD titrations rely on the CD signal changes observed upon quadruplex formation induced by binding of metal ions. It is possible that other modes of metal-DNA interactions may exist but without inducing changes in the ellipticity of the CD spectrum. Such interactions would not be detected by CD titration experiments. CD signals are measured as a small percentage difference in absorption of left- and right-polarized light. Hence signal increases with sample concentration. However, the concentration range that can be used is restricted by the amount of light passing through the sample to the detector. If the sameple concentration is too high, then only limited light will be detected and the difference in absorption will be swamped by noise in the electronics.

For the purposes of our studies and the concentration range we covered, binding constants measured by ITC are relatively more reliable than by those estimated from CD titration. ITC experiments allow a wider range of DNA concentrations In VP-ITC, a good approximation for the concentration (C) range of micromolecules in the sample cell given a binding constant (K), is

$K \times C = 5\text{-}500$

In Table 1, the highest K value for Ba^{2+} is 2.4×10^6 and concentration of TBA is 5.1×10^{-5} M. This leads to a value of

$K \times C = 122$

As for Pb^{2+}, the highest K value is 2.5×10^6 and the concentration of TBA is 2.1×10^{-5} M, resulting in a value for

$K \times C = 53$

This is an even better condition for the ITC experiments as the ideal concentration according to the ITC instrument manual is when $K \times C = 5\text{-}50$.

Unlike ITC, in CD experiments the corresponding DNA concentration for such tight binding to be accurately estimated should be in a range around 1/K, i.e. considerably smaller than in ITC. The uncertainty in Ba^{2+} concentration is also significantly greater at such low concentrations.

We can see that ΔH is not constant with respect to temperature, as discussed in chapter 3. In chapter 3, the ΔC_p measured for TBA in 25 mM KCl is 458 cal K^{-1} mol^{-1} and ΔC_p is 340 cal $K^{-1}mol^{-1}$ for TBA in 140 mM KCl. In our studies with barium and lead, the ΔC_p values are relatively higher. ΔC_p is 524 cal $K^{-1}mol^{-1}$ for TBA in Ba^{2+} and ΔC_p 910 cal $K^{-1}mol^{-1}$ is for TBA in Pb^{2+}. These results clearly indicated that ΔC_p is not zero as assumed in most studies.

4.6 REFERENCE

1. Xu, Y., Y. Noguchi, and H. Sugiyama, *The new models of the human telomere d[AGGG(TTAGGG)(3)] in K(+) solution.* Bioorg Med Chem, 2006. **14**(16): p. 5584-5591.
2. Luu, K.N., et al., *Structure of the human telomere in K+ solution: an intramolecular (3 + 1) G-quadruplex scaffold.* J Am Chem Soc, 2006. **128**(30): p. 9963-70.
3. Dai, J., et al., *An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution.* J Am Chem Soc, 2006. **128**(4): p. 1096-8.
4. Ambrus, A., et al., *Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization.* Biochemistry, 2005. **44**(6): p. 2048-58.
5. Phan, A.T., Y.S. Modi, and D.J. Patel, *Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter.* J Am Chem Soc, 2004. **126**(28): p. 8710-6.
6. Smirnov, I.V., et al., *Pb EXAFS studies on DNA quadruplexes: identification of metal ion binding site.* Biochemistry, 2002. **41**(40): p. 12133-9.
7. Basu, S., et al., *Direct detection of monovalent metal ion binding to a DNA G-quartet by 205 Tl NMR.* J. Am. Chem. Soc., 2000. **122**(13): p. 3240-3241.

8. Hardin, C.C., A.G. Perry, and K. White, *Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids.* Biopolymers, 2001. **56**(3): p. 147-94.

9. Mondragon-Sanchez, J.A., et al., *Tetraplex structure formation in the thrombin-binding DNA aptamer by metal cations measured by vibrational spectroscopy.* J Biomol Struct Dyn, 2004. **22**(3): p. 365-73.

10. Pinnavaia, T.J., et al., *Alkali metal ion specificity in the solution ordering of a nucleotide, 5'-guanosine monophosphate.* J Am Chem Soc, 1978. **100**: p. 1978.

11. Potaman, V.N. and V.N. Soyfer, *Divalent metal cations upon coordination to the N7 of purines differentially stabilize the PyPuPu DNA triplex due to unequal Hoogsteen-type hydrogen bond enhancement.* J Biomol Struct Dyn, 1994. **11**(5): p. 1035-40.

12. Smirnov, I. and R.H. Shafer, *Lead is unusually effective in sequence-specific folding of DNA.* J Mol Biol, 2000. **296**(1): p. 1-5.

13. Kankia, B.I. and L.A. Marky, *Folding of the thrombin aptamer into a G-quadruplex with Sr(2+): stability, heat, and hydration.* J Am Chem Soc, 2001. **123**(44): p. 10799-804.

14. Gray, D.M., S.H. Hung, and K.H. Johnson, *Absorption and circular dichroism spectroscopy of nucleic acid duplexes and triplexes.* Methods Enzymol, 1995. **246**: p. 19-34.

15. Lu, M., Q. Guo, and N.R. Kallenbach, *Thermodynamics of G-tetraplex formation by telomeric DNAs.* Biochemistry, 1993. **32**(2): p. 598-601.

16. Williamson, J.R., *G-quartet structures in telomeric DNA.* Annu Rev Biophys Biomol Struct, 1994. **23**: p. 703-30.

17. van Holde, K.E. Physical Biochemistry. 1985, Englewood Cliffs: Prentice-Hall, Inc. 235.

CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS

Specific substitution of ribo- for deoxyriboguanosines in the thrombin-binding aptamer (TBA), and vice versa, was demonstrated as a means of controlling the overall folding pattern of this particular quadruplex. Future studies will address the issue of whether this approach can be applied to other quadruplex-forming sequences and thus serve as a general method for engineering the quadruplex fold. Preliminary results on several other quadruplexes suggest that such an approach can indeed be extended to other sequences, although not without additional complications that will warrant detailed investigation.

As the studies presented here demonstrate that individual nucleoside conformational preferences have a profound impact on the overall quadruplex fold, another area of future study concerns the possibility of creating novel quadruplex folds. For example, substituting brominated or methylated deoxyguanosine, which prefer the *syn* conformation about the glycosidic bond, for deoxyguanosine may force the TBA structure from the normal *syn, anti, syn, anti, syn, anti, syn, anti* alternation in the guanine stretches to *anti, syn, anti, syn, anti, syn, anti, syn* alternation.

Another important aspect of quadruplex structure that has emerged from the investigation of quadruplex refolding is the observation that RNA quadruplexes all appear to fold in a parallel fashion, again most likely due to the propensity for rG to possess the *anti* conformation about the glycosidic bond. This raises the question of the generality of this observation, i.e. whether all RNA quadruplexes inherently fold into parallel quadruplexes. Due to the growing importance of RNA

quadruplexes in biological systems, the question of their structure is an important one worthy of future studies.

Many guanine-rich sequences form a mixture of different quadruplex structures under certain experimental conditions, making it difficult to study the properties of each structure individually. Judicious substitution of dG by rG and/or 8-brdG may allow selection of a single conformation under such conditions. Being able to control the equilibrium of one fold over the others will enable us to isolate single conformations from mixtures, thereby facilitating analysis of their structure and stability.

Finally, the studies on the binding of divalent cations to TBA may be extended to other quadruplex-forming sequences and, ultimately, to other binding partners, such as quadruplex-specific ligands or proteins.

UCSF LIBRARY

UC San Francisco LIBRARY
UNIVERSITY OF CALIFORNIA

7539326
3 1378 00753 9326

For reference
Not to be taken from the room.