INVESTIGATIONS INTO THE UNFOLDING OF A DNA QUADRUPLEX:
EFFECT OF LOOP SEQUENCE

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF SCIENCE

IN THE GRADUTATE SCHOOL OF THE

TEXAS WOMAN'S UNIVERSITY

DEPARTMENT OF BIOLOGY

COLLEGE OF ARTS AND SCIENCES

BY

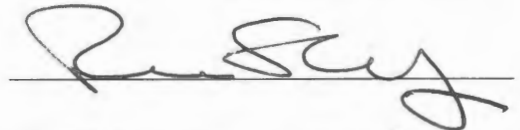DINESH KUMAR YADAV

DENTON, TEXAS

MAY, 2012

TEXAS WOMAN'S UNIVERSITY
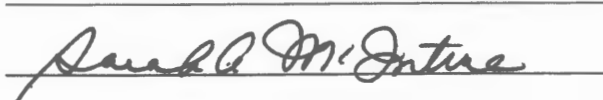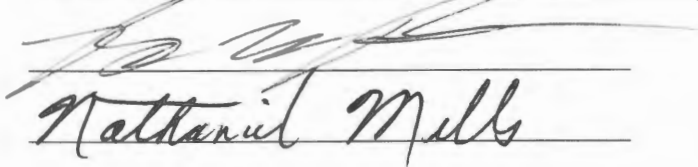
DENTON, TEXAS

January 9, 2012

To the Dean of the graduate school:

I am submitting herewith a thesis written by Dinesh Kumar Yadav entitled "Investigations into the Unfolding of a DNA Quadruplex: Effect of Loop Sequence." I have examined this thesis for form and content and recommended that it be accepted in partial fulfillment of the requirements for the degree of Master with a major in Biology.

Dr. Richard Sheardy, Major Professor

We have read this thesis and recommended its acceptance:

Department Chair

Accepted:

Dean of the Graduate School

DEDICATION

To my Mother

I am dedicating this thesis to my beloved mother for her blessings, care and support for my growth.

ACKNOWLEDGEMENTS

It is great pleasure that I record my deep sense of gratitude and sincere thanks to Professor Richard D. Sheardy for his support, guidance and patience throughout course of this research. He played an ideal role of a friend, philosopher and guide in shaping up the research of this thesis. I am grateful to Dr. Nathaniel Mills, Dr. Mark Britt and Dr. Brian Beck for their willingness to serve on my thesis committee. I am proud to say that I had a best committee to show me how to look beyond the scope of the biologist's perspective and find the clarity I need for the intelligible expression of my work.

I also wish to thank all other members of the research group, especially Brenna Tucker for her willingness to help and cooperate at all times.

Finally, and most importantly, I would like to record the contributions made by my parents in my education, whatever I have achieved in my life because of their love, understanding and support for me. Last but most certainly not least, I want to thank my wife, Neetu Jain who put up with a lot over the course of my graduate student life but gave me much needed support. This thesis is devoted to all my family members.

ABSTRACT

DINESH KUMAR YADAV

INVESTIGATIONS INTO THE UNFOLDING OF A DNA QUADRUPLEX:
EFFECT OF LOOP SEQUENCE

MAY, 2012

Understanding of the human telomere and telomerase is expected to provide

major insights into genome stability, cancer, and telomere-related diseases.

Quadruplexes have been received as a potential target for anti-cancer therapy because of

their ability to inhibit telomerase. Hence, it is very important to have a firm

understanding of the different structures that they can form as well as their respective

stabilities in order to use quadruplexes in developing new cancer therapeutics. Previous

studies have indicated that the unfolding of the DNA quadruplex formed from

$(TTAGGG)_4$ in potassium solution proceeds via a three state mechanism. Our interest is

the nature of the intermediate and the mechanism of the unfolding. Results of circular

dichroism (CD) suggest that unknown intermediate is a double hairpin which would form

if the backside loop of the folded quadruplex, which corresponds to the third TTA

segment of $(TTAGGG)_4$, acts as a "hinge." To investigate this further, we characterized

the unfolding of DNA quadruplexes formed where each loop of the quadruplex was

permuted to all possible combinations of A and T.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS AND SYMBOLS

Symbols

A……………………………………………………….................absorbance

ε……………………………………………….molar extinction coefficient

∂…………………………………………………………path length

c………………………………………………………..molar concentration

Θ…………………………………………………….................ellipticity

Θ_l ……….….………………………………………….left handed absorption

Θ_r …………………………………………………right handed absorption

$[\Theta]$………………………………………………………… molar ellipticity

CHAPTER I

INTRODUCTION

DNA, a fundamental building block for an individual's entire genetic makeup, was first reported by Watson and Crick in 1953 as a double helix (1). Bases from the two strands were paired by complimentary hydrogen bonding in the helix interior and the sugar phosphate backbones extended along the outside, minimizing electrostatic repulsions and the strands ran in anti parallel directions. Biochemical studies of base composition ratio suggested that the number of purine bases equals the number of pyrimidine bases, i.e. adenine equals thymine and guanine equals cytosine (2). The difference in hydrogen bonding between these pairs leads to differences in stability of the helix. X-ray diffraction studies of various forms of nucleic acid fibers resulted in several characteristic and distinct diffraction patterns. From these different patterns several classes of double helical models have been proposed like A, B and Z forms of oligonucleotide duplexes (Figure 1.1). A and B are right-handed and can occur with any sequence, Z is left handed and occurs with alternating Pu-Py sequences, mainly GC (3,4). In aqueous solutions and in cells, DNA is essentially in B-form.
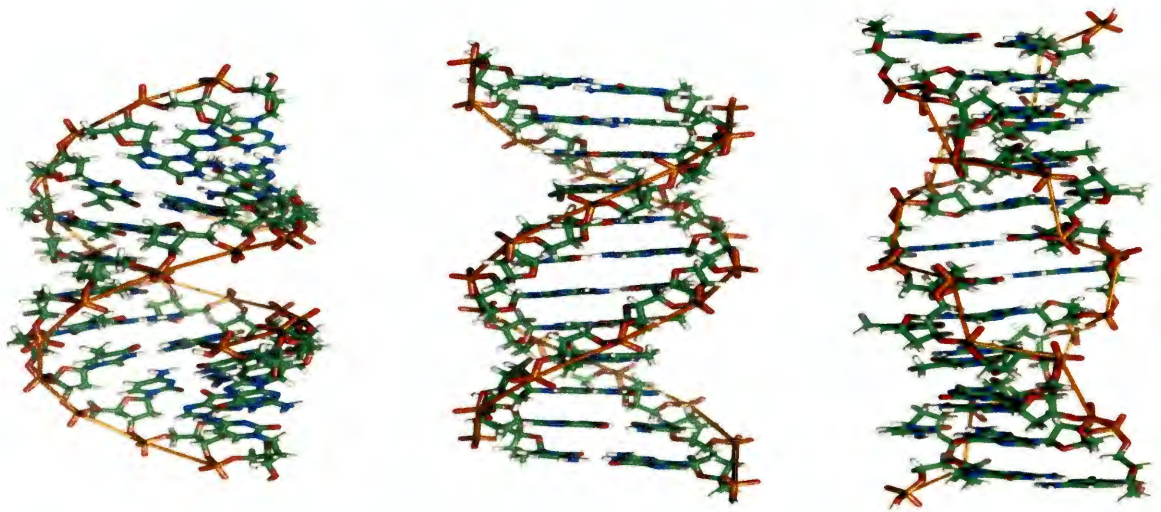
FIGURE 1.1: Oligonucleotide duplexes. From left to right showing A, B and Z form of DNA helix (37).

The shapes of DNA and RNA control their assembly with other cellular macromolecules and their linear sequence of nucleotides encodes genetic information to specify the composition of proteins. Different types of interactions control the structural changes observed in DNA like the charged phosphate groups of the backbone which are mutually repulsive (5). The formation of hydrogen bonds between bases and strong stacking interaction between the planar surfaces of the base pairs are implicated in the formation of multi-stranded structures (double or triple stranded helices, quadruplexes) Figure 1.2.

A) Watson-Crick
DNA Double Helix

B) DNA Triplex

C) DNA Quadruplex
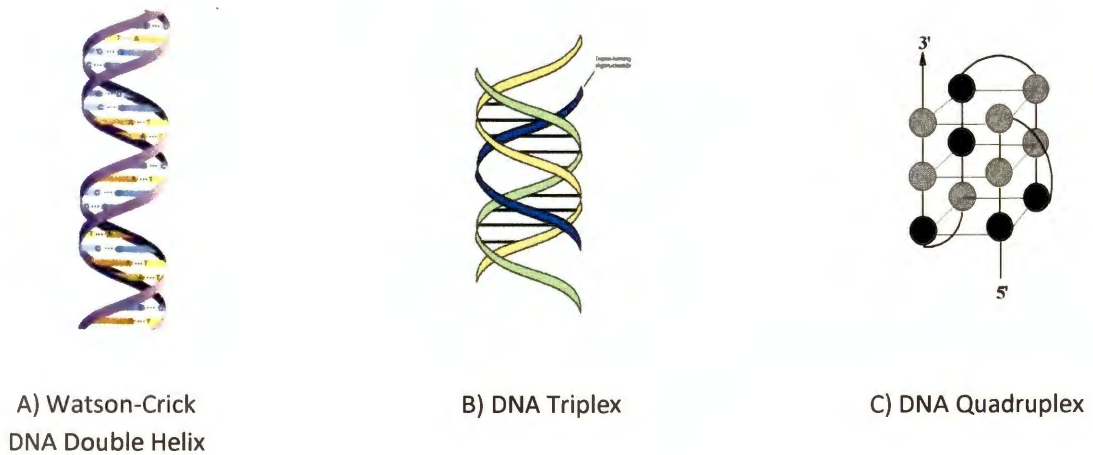
FIGURE 1.2: Selected DNA structures. DNA is highly polymorphic and can form a variety of secondary structures (33, 38).

The ends of human chromosomes have unique properties of capping. Muller named these ends "telomeres" (from Greek word *telo*, meaning "end" and *mere*, meaning "parts"), based on their position on chromosomes. A chromosome is a long DNA molecule, and associated proteins, which carries the heredity information of an organism. It is a single piece of coiled DNA containing many genes, regulatory elements and many nucleotides. Later in his mutagenesis experiment on *Drosophila,* Muller finds that the ends of the chromosomes were strangely resistant to the effects of the mutagenic X-rays. He therefore hypothesized that the terminal gene must have a special function of sealing the end of the chromosome (6). It is known now that Muller was only partially correct. Telomeres do indeed play an essential role in stabilizing the ends of

3

chromosomes, but they do not contain active genes. Instead, telomeres contain an array of highly repeated DNA sequences and specific binding proteins that form a unique structure at the end of the chromosome (7). This property of "capping" natural chromosome ends was the first essential function ascribed to telomeres. In the early 1970s, it was realized that due to intrinsic requirements and polarities of the normal DNA replication machinery, a DNA end cannot be fully replicated. Thus, telomeres must provide the basis for a second essential function, namely the complete replication of the DNA at chromosomal termini (8).

Telomeric DNA consists of tandem repeats of short guanine rich sequences, such as d(TTAGGG) in mammals or d(TTTTGGGG) in protozoa (9, 10) . Replication of eukaryotic chromosomes by DNA polymerase cannot fully copy the ends of the telomeric DNA, as the polymerase is unable to fully replicate the extreme 3' end of a DNA sequence in the absence of a template strand, known as "end replication effect." Consequently, each replication cycle of the cell results in the deletion of the few nucleotide bases at the 3' end and telomeric DNA progressively shortens in the absence of any compensating mechanism. The shortening, by approximately 50-100 bases per round of the cell division, continues until the critical state of replicative senescence is reached, when no further replication takes place.

Telomere maintenance mechanism is provided by a specific enzyme called telomerase, first identified in ciliates. Activation of a telomere maintenance mechanism

seems indispensible for immortalization of human cells. Instead of the telomere

shortening over time in cells that contain telomerase, the length of the telomere is

maintained and the cell never reaches apoptosis. In human cells, telomerase

functions as a reverse transcriptase to add multiple copies of the 5'-GGTTAG-3'

motif to the end of the G-strand of the telomere. Telomerase is normally active in

germ cells but not in somatic cells. In the majority of the tumor cells (80-90%),

this enzyme is over expressed. Even in normal stem cells, the level of telomerase

activity is low or absent (11). The understanding of the human telomere and

telomerase is expected to provide major insights into genome stability, cancer,

and telomere-related diseases. Quadruplexes have been proposed as potential

targets for anti-cancer therapy because of their ability to inhibit telomerase (8, 12-

14). Hence, it is very important to have a firm understanding of the different

structures that they can form, as well as their respective stabilities, to use

quadruplexes in developing new cancer therapeutics (15).

Quadruplex forming sequences can be found in many different DNA

sequences within eukaryotic cells. Telomeric DNA sequences comprise G-rich

tandem repeats, i.e., are not pure G sequences, and have short non-G tracts

regularly interspersing the G repeats. The sequence known as the human

telomere (TTAGGG), found at the end of human chromosomes, is one such

example. The formation of these quadruplex structures at telomere ends is

possible since the terminal nucleotides at the 3' ends of telomeric DNA are single

stranded, where the single-strand overhang is 100-200 nucleotides long. Promoter regions of several oncogenes including c-myc and Bcl-2 as well as the gene for VEGF (vascular endothelial growth factor) have also been reported to have quadruplex forming sequences (16-18).

Quadruplexes can be formed from one, two or four separate strands of DNA (or RNA) and can display a wide variety of topologies, which are in part a consequence of various possible combinations of strand direction, as well as variations in loop sizes and sequences (19). Quadruplexes can be defined in general terms as structures formed by at least two stacked G-tetrads, which are held together by loops arising from the intervening mixed-sequence nucleotides that are not usually involved in tetrads themselves.

The unique interaction of G-quadruplexes with specific monovalent cations consists of chelation of the ion, with varying degrees of affinity depending on the identity of the cations, between two adjacent guanine quartets. The actual interaction appears to depend on the balance between the energetics of coordination of the cations to G-quartets and the energetics of removal of water molecules ordinarily bound to the cations. The most stable G-quadruplexes are formed in the presence of potassium, with the order of stabilizing ability of group1a cations following the order K>Rb>Na>Li=Cs (and, for group 2b cations, Sr>Ba>Ca>Mg).

Quadruplexes can be formed by one, two or four molecules of oligonucleotide, which are referred to as monomer, dimer and tetramer structures, respectively. The G

bases of the tetrads have an anti conformation when the strand orientation is parallel but alters between syn-anti-syn when any strand runs anti-parallel to the others. Monomer and dimer quadruplex formations have been classified further based on the positioning of their loop regions into chair (lateral loop) or basket (diagonal loop) forms. The relative strand orientation (5' to 3' polarity) of the four strands of the quadruplex may be parallel, anti parallel or mixed. The sequence, as well as the length of the loop regions, influences how the structure is folded into a quadruplex.

Intramolecular G-quadruplexes are formed by the stacking of several G-tetrads and are constituted by four backbone strands with several connecting loops. Four guanines are arranged in a plane to form the G-tetrad by hydrogen bonds and are stabilized by cations.
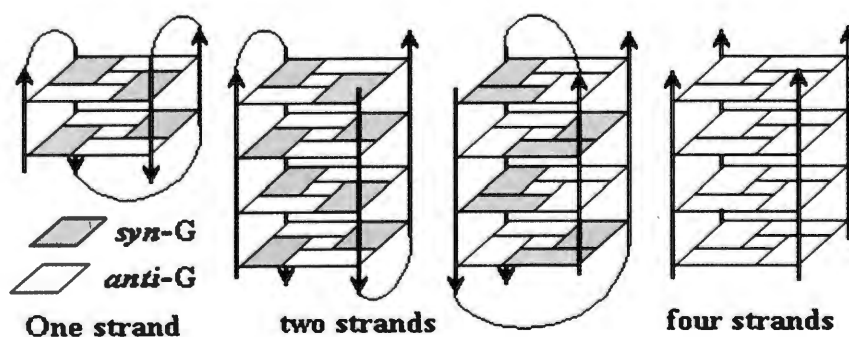
FIGURE 1.3: Molecularities of quadruplexes. Molecularity of the quadruplex is the number of associated strands in the structure. Here molecularties of 1, 2 and 4 are shown with different strand orientation indicated by arrow (20).

7

The molecular basis for the association was subsequently determined by fiber diffraction and biophysical studies using the concept that the Hoogsteen hydrogen-bonded guanine tetrad (also known as G-quartet) is the basic structural motif.
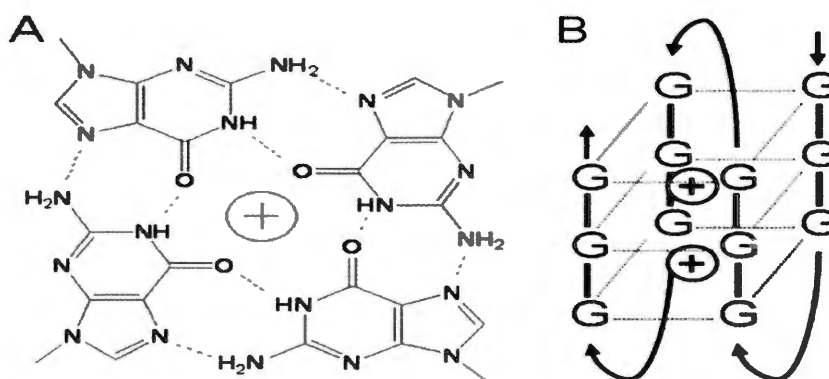
FIGURE 1.4: G-tetrad and quadruplex. (A) Guanine tetrad having 8- Hoogsteen hydrogen bonds and (B) Sequences containing four groups of three guanines are able to fold into a G-quadruplex (22).

The Hoogsteen base pairing occurs between the N1 and N2 of one G base with the O6 and N7 of the neighboring G base. This arrangement allows for the creation of eight hydrogen bonds within each tetrad increasing the stability of these structures (20-32).

To learn more about quadruplex structure and its folding/unfolding pattern in potassium solution, a library of the DNA sequences has been created. Previous studies

indicated that the unfolding of the quadruplex formed from $(TTAGGG)_4$ in potassium solution proceeded via a three state mechanism that is supported by calorimetric determinations. The observed enthalpic and entropic contributions to the total free energy for each transition were essentially independent of the $[K^+]$. However, the first transition had a higher enthalpy and entropy than the second transition.
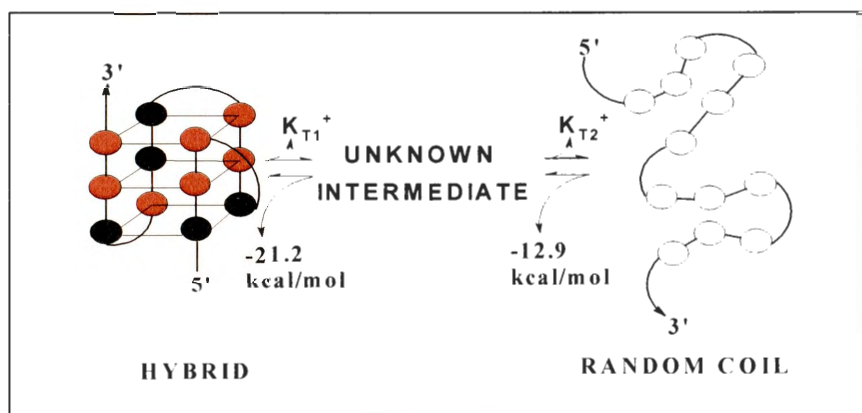
FIGURE 1.5: Proposed model for unfolding (33).

In addition, the number of K^+ ions released from the DNA into the bulk solution was higher for the first transition than for the second transition. These observations led to the working model for the transition. If the backside loop of the folded quadruplex (which corresponds to the third TTA segment of $(TTAGGG)_4$) acts as a "hinge," then the unknown intermediate in Figure 1.5 would be a double hairpin (33, 34).

Information on structure comes directly from the CD determinations while thermodynamic information can be collected from both UV/Vis and CD optical melts (via Van't Hoff analysis) and directly from differential scanning calorimetry. DNA absorbs light in the range of 240-280 nm with a maximum absorption around 260 nm. DNA absorption can be attributed to the purine and pyrimidine bases of the structure. We can be sure of this because the ribose-phosphate backbone of the DNA does not affect the absorption in this particular range. The wavelength of maximum absorbance (λ_{max}) and the extinction coefficient (ε) at λ_{max} are dependent upon several different factors including the base composition, the base-pairing interactions present, the salt concentration of the solution, and the pH of the solution. DNA absorbs light so the Beer-Lambert Law can be used to determine the concentration of the DNA in a given sample. It states that:

$$A = \partial * \varepsilon * c \qquad\qquad \text{(Equation 1.1)}$$

Where A = absorbance, ε = the molar extinction coefficient, ∂ = the path length (width of the cuvette) and c = molar concentration.

Equation (1.1) can be solved for DNA concentration when its absorbance in a sample at 260 nm, molar extinction coefficient and path length is known. DNA optical melting refers to monitoring the changes in absorbance of the sample at a particular wavelength while there is an increase in temperature. While plotting absorbance at 260 nm against the temperature, a sigmoidal curve can be obtained which shows a transition

10

between the folded and unfolded structures. The midpoint of the transition on the curve is directly related to the melting temperature of the sample. As the temperature increases, the DNA structure unfolds into a single strand and re-folds (anneal) again to its secondary structure while cooling for a reversible process. Some annealing processes are longer than others as they depend upon the length of the sequence and other environmental effects such as pH and salt concentrations.

CD spectroscopy has been widely used to diagnose the existence of G-quadruplexes and to study the effects of various loop sequences in solution conditions. For a CD signal to be observed, the molecule under study must be chiral due to lack of the symmetry and must be chromophoric. Circular dichroism can be used to determine types of secondary structures of nucleic acids, like a duplex or a quadruplex as well as proteins. CD is also a technique commonly used as a tool for studying nucleic acid – ligand interactions. It is technique that measures the difference in absorption of left and right circularly polarized light. When circularly polarized light hits a chiral sample, the right and left hand components of the light are absorbed differently. The difference between the left handed absorption and the right handed absorption is a phenomenon known as ellipticity:

$$\Theta = \Theta_l - \Theta_r \qquad \qquad \text{(Equation 1.2)}$$

Where, Θ = ellipticity, Θ_l = left handed absorption and Θ_r = right handed absorption

CD spectra are most often presented as a plot of molar ellipticity versus wavelength. Molar ellipticity is a preferred unit of measurement because it allows comparison of different data based on their concentration. The relationship between observed ellipticity and molar ellipticity is illustrated by the following equation:

$$[\Theta] = \Theta/c\partial \qquad \text{(Equation 1.3)}$$

Where, $[\Theta]$ = molar ellipticity, Θ = ellipticity, c = the concentration of the sample and ∂ = the path length of the cuvette.

Our working hypothesis is that the third TTA in the sequence (which forms the loop at the upper back of the quadruplex in the Figure 1.5) acts as a hinge in the first transition allowing the folded structure to unfold to a double hairpin. To test this hypothesis, we characterized sequences possessing all permutations of T and A in all the loops, i.e. $(XXXGGG)_4$ where X = A or T. The specific aims of this research are: 1) to investigate the effect of loop sequence context on the structure of intramolecular DNA quadruplexes by circular dichroism (CD) spectroscopy; and 2) to investigate the effect of loop sequence context on the thermal stability of intramolecular DNA quadruplexes by CD optical melting. Our approach is to systematically study DNA oligomers with different permutations of T and A in the second, third and fourth segments of $(TTAGGG)_4$.

12

MATERIALS AND METHODS

Buffer preparation

Potassium standard buffers (150 mM K^+, 10 mM phosphate, and 0.1 mM EDTA) at pH 7 were prepared by dissolving 0.680 g KH_2PO_4 (VWR International Lot # 46032627), 0.870 g K_2HPO_4 (VWR International Lot # 46205641), 10.1 g KCl (VWR International Lot # 46100627), and 0.037 g EDTA (EMD Chemicals Lot # 45166714) in 1 L of deionized water. 10X potassium buffers (150 mM K^+, 100 mM phosphate, and 1.0 mM EDTA) were also prepared by combining 6.80 g KH_2PO_4, 8.71 g K_2HPO_4, and 0.372 g EDTA in 1 L of deionized water. All buffers were filtered through a 0.45 μm Millipore filter and degassed before being stored for use.

Determination of DNA concentration

DNA concentrations were determined by UV/Vis Varian Cary 100 Bio model (Varian Associates, Palo Alto, CA) spectrometer at 260 nm. The reconstituted DNA sample was diluted 10 fold in 150 mM potassium buffer and scanned between 320 nm and 220 nm at 25°C and 95°C with subtraction of the buffer baseline. Samples were run in 10 mm square quartz cuvettes and

concentrations were calculated by Beer's Law using the molar extinction coefficient

provided by the DNA oligomers supplier (Table 2.1-2.3).

TABLE 2.1: Molar extinction coefficient (ε) of the oligonucleotides (with all possible base combinations for loop 2)

	Oligomer	Sequence	Ext. coefficient
Loop 2	**HTEL1**	TTAGGGTTAGGG**TTA**GGGTTAGGG	243717.6
	HTEL 1HA	TTAGGGTTAGGG**TAT**GGGTTAGGG	243717.6
	HTEL 1HB	TTAGGGTTAGGG**ATT**GGGTTAGGG	243717.6
	HTEL 1HC	TTAGGGTTAGGG**TTT**GGGTTAGGG	237816.9
	HTEL 1HD	TTAGGGTTAGGG**AAT**GGGTTAGGG	249618.3
	HTEL 1HE	TTAGGGTTAGGG**ATA**GGGTTAGGG	249618.3
	HTEL 1HF	TTAGGGTTAGGG**TAA**GGGTTAGGG	249618.3
	HTEL 1HG	TTAGGGTTAGGG**AAA**GGGTTAGGG	255519.0

TABLE 2.2: Molar extinction coefficient (ε) of the oligonucleotides (with all possible base combinations for loop 1)

	Oligomer	Sequence	Ext. coefficient
Loop 1	**HTEL1**	TTAGGG**TTA**GGGTTAGGGTTAGGG	243717.6
	HTEL-1LA	TTAGGG**TTT**GGGTTAGGGTTAGGG	238700.0
	HTEL-1LB	TTAGGG**TAT**GGGTATGGGTTAGGG	244600.0
	HTEL-1LC	TTAGGG**ATT**GGGATTGGGTTAGGG	243200.0
	HTEL-1LD	TTAGGG**AAT**GGGTTTGGGTTAGGG	247100.0
	HTEL-1LA2	TTAGGG**ATA**GGGTTTGGGTTAGGG	249100.0
	HTEL-1LA3	TTAGGG**TAA**GGGTTTGGGTTAGGG	248500.0
	HTEL-1LA4	TTAGGG**AAA**GGGTTTGGGTTAGGG	251000.0

14

TABLE 2.3: Molar extinction coefficient (ε) of the oligonucleotides (with all possible base combinations for loop 3)

	Oligomer	Sequence	Ext. coefficient
Loop 3	HTEL1	TTAGGGTTAGGGTTAGGGTTAGGG	243717.6
	HTEL-3LA	TTAGGGTTAGGGTTAGGG**TAT**GGG	244600.0
	HTEL-3LB	TTAGGGTTAGGGTTAGGG**ATT**GGG	243200.0
	HTEL-3LC	TTAGGGTTAGGGTTAGGG**TTT**GGG	238700.0
	HTEL-3LD	TTAGGGTTAGGGTTAGGG**ATA**GGG	249100.0
	HTEL-3LD	TTAGGGTTAGGGTTAGGG**TAA**GGG	248500.0
	HTEL-3LD	TTAGGGTTAGGGTTAGGG**AAT**GGG	247100.0
	HTEL-3LD	TTAGGGTTAGGGTTAGGG**AAA**GGG	251000.0

Sample preparation

HPLC purified oligonucleotides were purchased from Biosynthesis Inc (Lewisville, Texas). All sequences were spun at 14000 rpm for 5 minutes before reconstitution in 1 ml filtered buffer (150mM K$^+$). The tube was vortexed for 30 seconds and placed in a heating block. The temperature of the heating block was set to 95°C. When the temperature reached 95°C for 5 minutes, the heating block was turned off. After cooling to room temperature, the tube was stored for 24 hrs at 5°C.

Circular dichroism spectrophotometry

Circular dichroism spectra were collected with DESA Rapid Scanning Monochromator Spectrophotometer (Model Olis RSM 1000) attached with a temperature controller (Julabo CF31) and a nitrogen purging unit. Nitrogen is

purged through the entire three chambers (lamp, sample and monochromator) at 25-30 standard cubic feet per hour (SCFH). Reconstituted samples were kept at room temperature at least for 1 hour prior to every CD studies. Data was collected from 25°C to 95 °C at an interval of 5°C with an integration time of 3 seconds. After equilibration for 5 minutes at each temperature, CD spectra were collected in the range 220 nm to 320 nm for integration time 3 seconds. A baseline of the 150 mM potassium buffer was obtained from absorbance 320 – 220 nm using an integration time of 3 seconds and was subtracted from each spectrum. Samples were run in a 1 mm circular quartz cuvette. Data were analyzed using Olis Global Works and SigmaPlot version 11.

CHAPTER III

RESULTS AND DISCUSSION

To understand the contribution of the loops to the structure and stability of

the folded quadruplexes, we chose to study the variations of the model sequence

HTEL1 (TTAGGGTTAGGGTTAGGGTTAGGG). We can consider HTEL1 as

wild type because all other sequences are copies of this with few variations

(Figure 3.1) in the loop sequence.
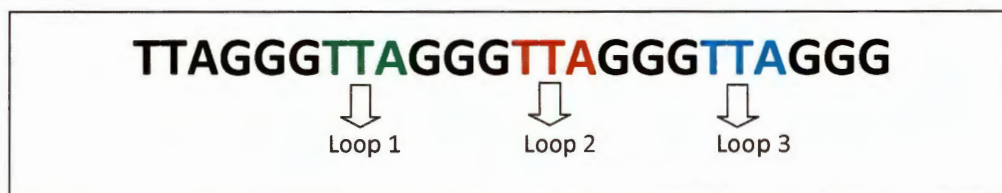
FIGURE 3.1: HTEL1. Segments of oligomer defined as loop 1, loop 2 and loop 3.

Representative CD spectra of HTEL1 at selected temperatures are presented in

Figure 3.2. The quadruplex formed from this sequence has signature CD spectra and can

be identified by a peak at 290nm, a shoulder near 270nm and a trough at 240nm.

Solution NMR structure of HTEL1 confirms that in a standard potassium (150 mM)

Buffer, HTEL1 exists in a quadruplex form and this result was confirmed with CD later

(31).

17
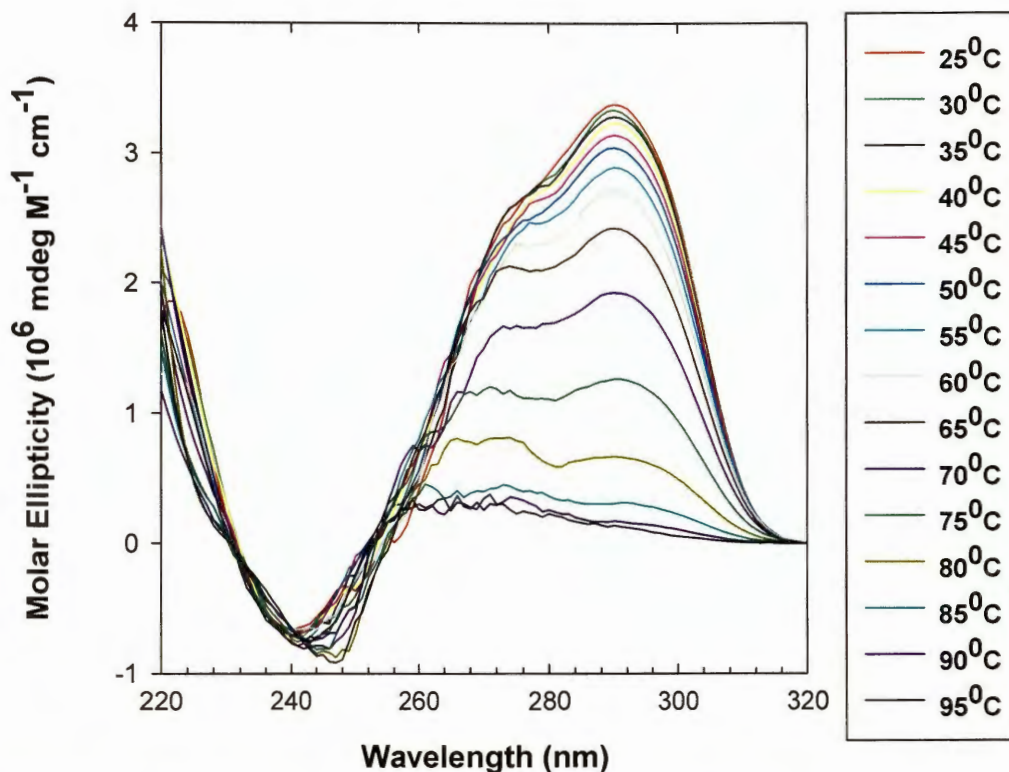
FIGURE 3.2: CD spectra of HTEL1 at various temperatures. Spectra collected from 25°C to 95°C in an interval of 5°C. HTEL1 sample was prepared in potassium standard buffer at room temperature.

The peak at 290 nm is attributed to the stacking of two tetrads with opposite H-bond orientation polarities while the peak at 270 nm has been attributed to the stacking of two tetrads with same H-bond polarities (33). Our CD spectrum of HTEL1 is also in accordance with NMR results. The peak at 290 nm and a shoulder at 270 nm characterize an anti parallel quadruplex while shoulder at 270 nm and trough at 240 nm characterize a parallel quadruplex.

A plot between temperature and corresponding molar ellipticity at 290 nm gives a melting profile for HTEL1 (Figure 3.3). In this graph we can assume a two state transition model to calculate a T_m value the phase transition temperature. As we have hypothesized that the unfolding of the quadruplexes follows a three state model, the calculated T_m value is only an approximation.
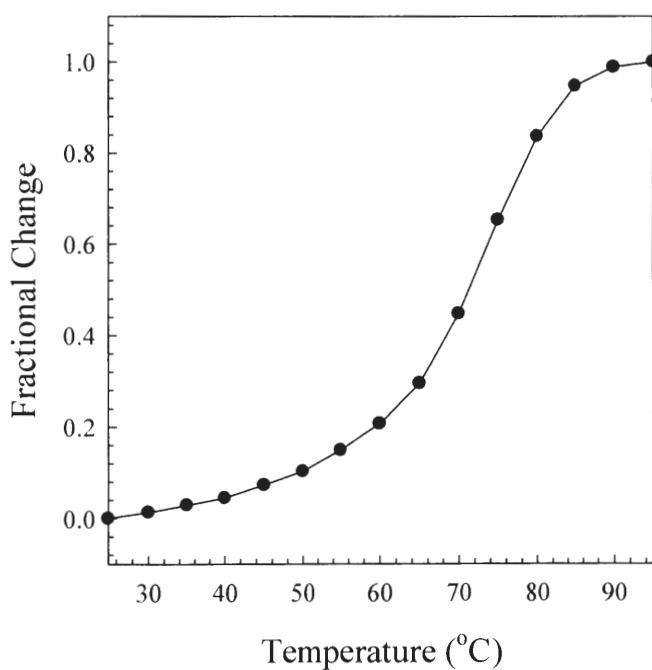
FIGURE 3.3: Melting curve for HTEL1. Molar ellipticities collected at various temperatures at a constant wavelength of 290 nm.

Loop 2

Our first specific aim is to identify the effect of modifying the sequence context of loop 2 on the structure and stability of the resultant quadruplexes. We selected eight oligomers by modifying the sequence of loop 2 by considering all permutations of A and T. The CD spectra of all DNA oligomers (for loop 2) in Table 3.1 at selected temperatures and in 150 mM potassium buffer are given in Figure 3.4.

TABLE 3.1: Loop 2 variants

	Oligomer	Sequence
	HTEL1	TTAGGGTTAGGG**TTA**GGGTTAGGG
	HTEL 1HA	TTAGGGTTAGGG**TAT**GGGTTAGGG
	HTEL 1HB	TTAGGGTTAGGG**ATT**GGGTTAGGG
Loop 2	**HTEL 1HC**	TTAGGGTTAGGG**TTT**GGGTTAGGG
	HTEL 1HD	TTAGGGTTAGGG**AAT**GGGTTAGGG
	HTEL 1HE	TTAGGGTTAGGG**ATA**GGGTTAGGG
	HTEL 1HF	TTAGGGTTAGGG**TAA**GGGTTAGGG
	HTEL 1HG	TTAGGGTTAGGG**AAA**GGGTTAGGG
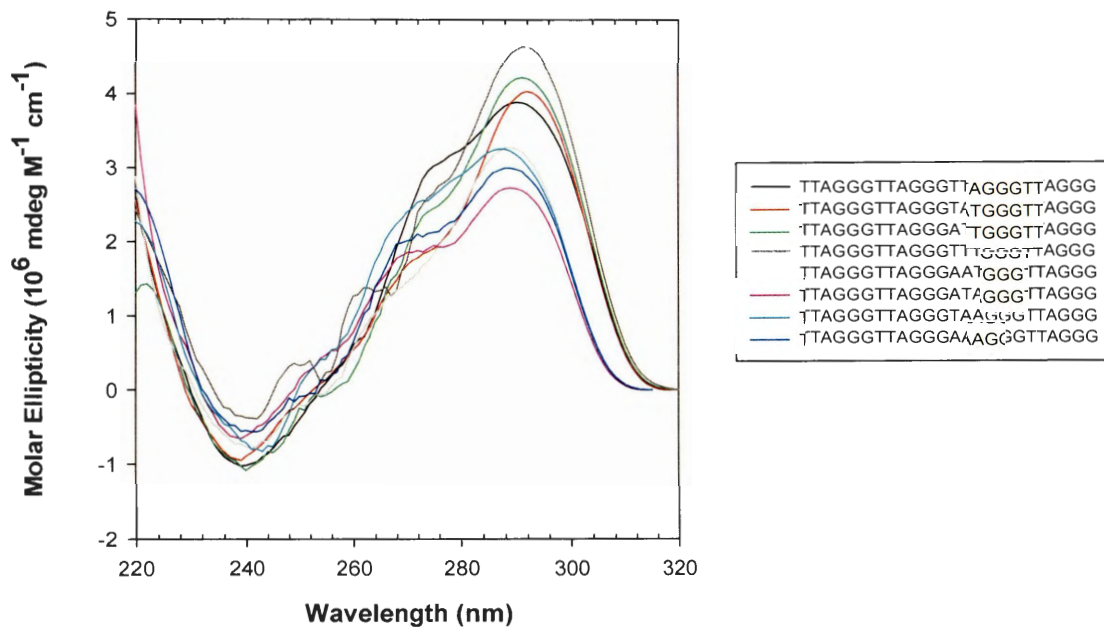
FIGURE 3.4: CD spectra of all loop 2 oligomers in potassium buffer at 25°C.

The spectra are similar to that of HTEL1 with peaks near 290 nm and shoulders near 270 nm. The CD spectra of the same oligomers at 95°C (Figure 3.5) indicate that these oligomers are all fully unfolded at that temperature.
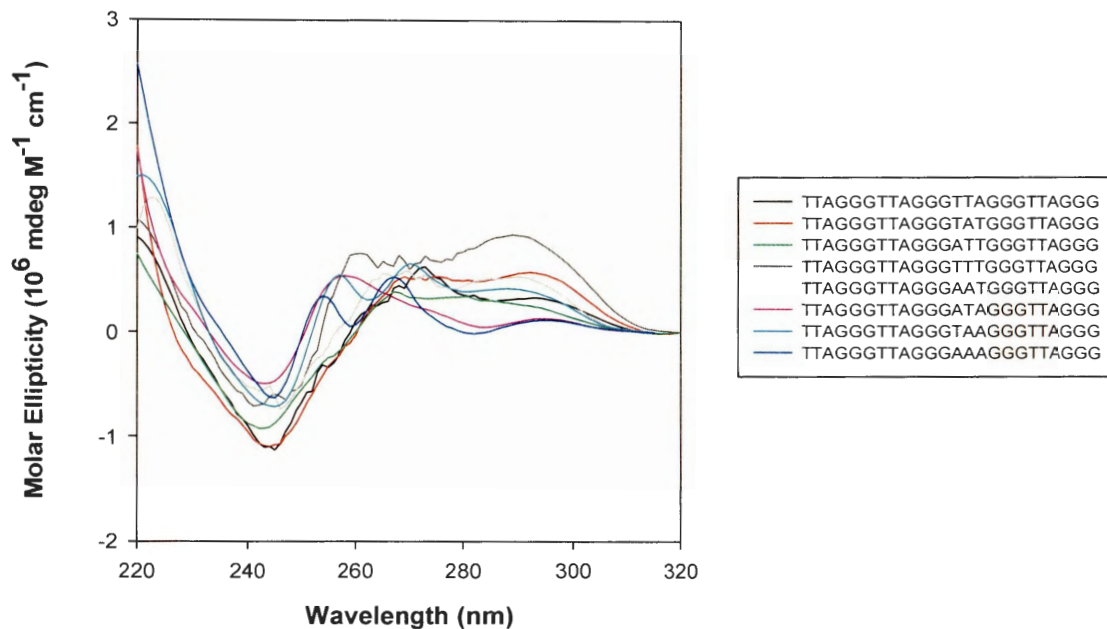
FIGURE 3.5: CD spectra of the loop 2 oligomers at 95°C. All the oligomers are prepared in potassium buffer.

Thus, CD spectra of these structures are in agreement with the formation of quadruplex. CD spectra for these sequences have also been collected over a range of temperatures to establish an optical melting profile and obtain a melting temperature (T_m) for each of the quadruplexes. Molar ellipticities at 290 nm are plotted against temperature giving a sigmoidal curve that can be used to derive T_m values as given in Table 3.2.
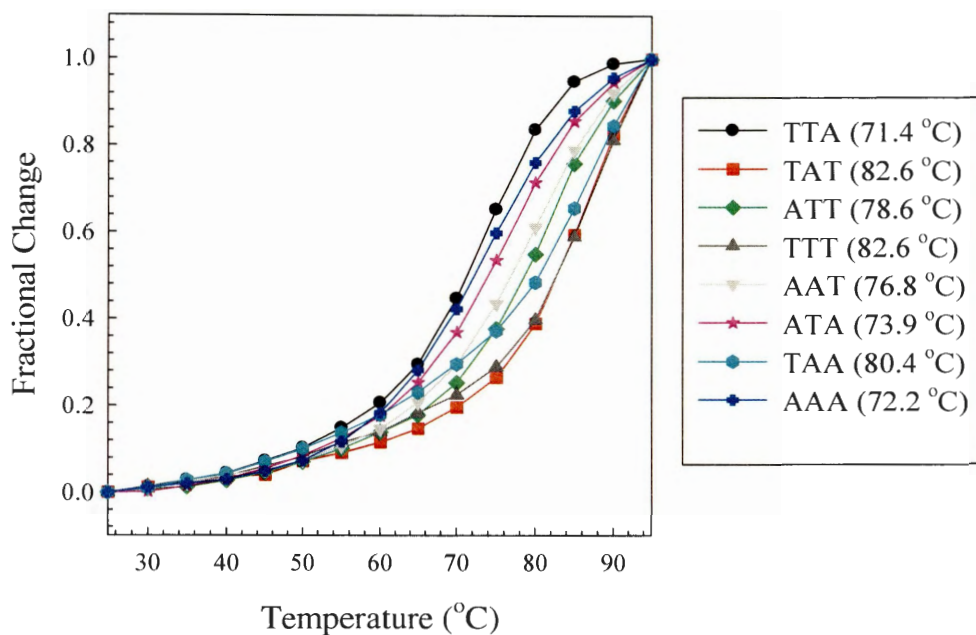
FIGURE 3.6: Melting profiles of all the loop 2 sequences at 290 nm.

Close inspection of each melting curve gives a hint of a multiphase transition. The data show systematic deviations from a simple single sigmoidal shape. From these plots we can see that when loop 2 is all replaced with T bases, the stability in relation to melting temperature has increased. Table 3.2 shows listing of T_m values for all the loop 2 oligomers.

TABLE 3.2: Number of "T" and number of "A" bases in the oligomers, T_m, hyperellipticity values along with wavelength for peak, shoulder and trough.

|	# T	# A	λ1 (Trough)	λ2 (Shoulder)	λ3 (Peak)	Loop Seq.	Melting Point (°C)	Hyper ellipticity
HTEL1	8	4	240	272	291	TTA	71.4	10.3
HTEL 1HA	8	4	239	268	292	TAT	82.6	5.99
HTEL 1HB	8	4	240	272	291	ATT	78.6	8.40
HTEL 1HC	9	3	242	272	291	TTT	82.6	1.23
HTEL 1HD	7	5	241	268	292	AAT	76.8	11.8
HTEL 1HE	7	5	239	270	292	ATA	73.9	20.7
HTEL 1HF	7	5	241	269	290	TAA	80.4	6.61
HTEL 1HG	6	6	239	270	292	AAA	72.2	35.2

On the basis of Table 3.2, the sequences are arranged according to their melting point. These findings, however, cause questions to arise about the validity of the T_m values obtained from the optical melting profiles which assume a two state transition. The T_m values established through the optical melts can only be used to determine a relative stability between the strands.

TAT (82.6°C) ≈ TTT (82.6°C)

>TAA (80.4°C) >ATT (78.6°C)

>AAT (76.8°C) >ATA (73.9°C) >AAA (72.2°C) >TTA (71.4°C)

As can be seen, an increase in the number of A bases in the loop 2 decreases the stability in relation to melting temperature except –TAA–. Even though –TAA– has 2 A bases, its stability is more in comparison to –ATT– and –TTA–. It appears that a T base at

24

the middle position has no impact in the unfolding or stability of the oligomer as we see melting temperatures of –TTT- and –TAT- are almost the same. A possible reason could be the distance of that middle base from the G-tetrads or position of the middle base as it is sandwich between two other bases in the loop. The same difference is seen with –ATA- and –AAA-. Apparently, the position of T base is also having some impact upon the stability of the oligomer when comparing loop sequence –TAA- and – AAT-. This impact is opposite when we compare loop sequence –ATT- and –TTA-. The exception sequence –TAA- can be a complimentary sequence of the overhang part (-TTA-) of oligomer when reversed (-ATT-). Perhaps that overhang part is interacting with loop 2 in the unfolding process. In the unfolding process, loop 2 opens and makes double hairpin structure. This unfolding process exposes loop 2 to form hydrogen bonds with complimentary sequence of overhang part. Those hydrogen bonds provide more stability to that structure.

The above results suggest three possibilities which may cause difference in the unfolding and stability of the oligomers: (1) loop sequence context; (2) the loops are interacting with the G-tetrads some way causing a difference in stability; or (3) the loops are interacting to each other and these interactions are either increasing stability or decreasing it. Loop 2 has a major effect in the stability which shows the importance of position of this loop. The position of loop 2 is very compact as this is the middle loop above G-tetrads. This particular position

of the loop provides a protection for H-bond between guanine bases at one side and allows H-bonds at opposite side to break first. This selective breaking of the bonds provides space for escape of cations. Removal of cations clears the path for the breaking of other H-bonds. In the unfolding of quadruplexes, base-base interactions between all three loops and G-tetrads are very important and to understand it better, we also studied the influence of loop 1 and loop 3 on the structure and stability.

Loop 1

The following oligomers in Table 3.3 of the loop 1 sequence have been studied to see the effect of base-base interaction with loop 1.

TABLE 3.3: Loop 1 variants

	Oligomer	Sequence
Loop 1	**HTEL1**	TTAGGG**TTA**GGGTTAGGGTTAGGG
	HTEL-1LA	TTAGGG**TTT**GGGTTAGGGTTAGGG
	HTEL-1LB	TTAGGG**TAT**GGGTTAGGGTTAGGG
	HTEL-1LC	TTAGGG**ATT**GGGTTAGGGTTAGGG
	HTEL-1LD	TTAGGG**AAT**GGGTTAGGGTTAGGG
	HTEL-1LA2	TTAGGG**ATA**GGGTTAGGGTTAGGG
	HTEL-1LA3	TTAGGG**TAA**GGGTTAGGGTTAGGG
	HTEL-1LA4	TTAGGG**AAA**GGGTTAGGGTTAGGG

FIGURE 3.7: CD spectra of the loop 1 oligomers at 25°C. All the oligomers are prepared in potassium buffer at room temperature.

The CD spectra of the oligomers given in Table 3.3 have peaks at 290 nm, shoulders at 270 nm and troughs at 240 nm. These spectra indicate that all the oligomers have similar loop orientations, stacking of the G-tetrads and H-bond polarities. Oligomers with –TTT- and -TAT- loop sequence have another peak near 250 nm. That extra peak at 250 nm gives a hint about different folding from the other oligomers. This may be most likely due to a chair type anti parallel quadruplex.

Legend (right of plot):
- TTAGGGTTAGGGTTAGGGTTAGGG
- TTAGGGTTTGGGTTAGGGTTAGGG
- TTAGGGTATGGGTTAGGGTTAGGG
- TTAGGGATTGGGTTAGGGTTAGGG
- TTAGGGAATGGGTTAGGGTTAGGG
- TTAGGGATAGGGTTAGGGTTAGGG
- TTAGGGTAAGGGTTAGGGTTAGGG
- TTAGGGAAAGGGTTAGGGTTAGGG

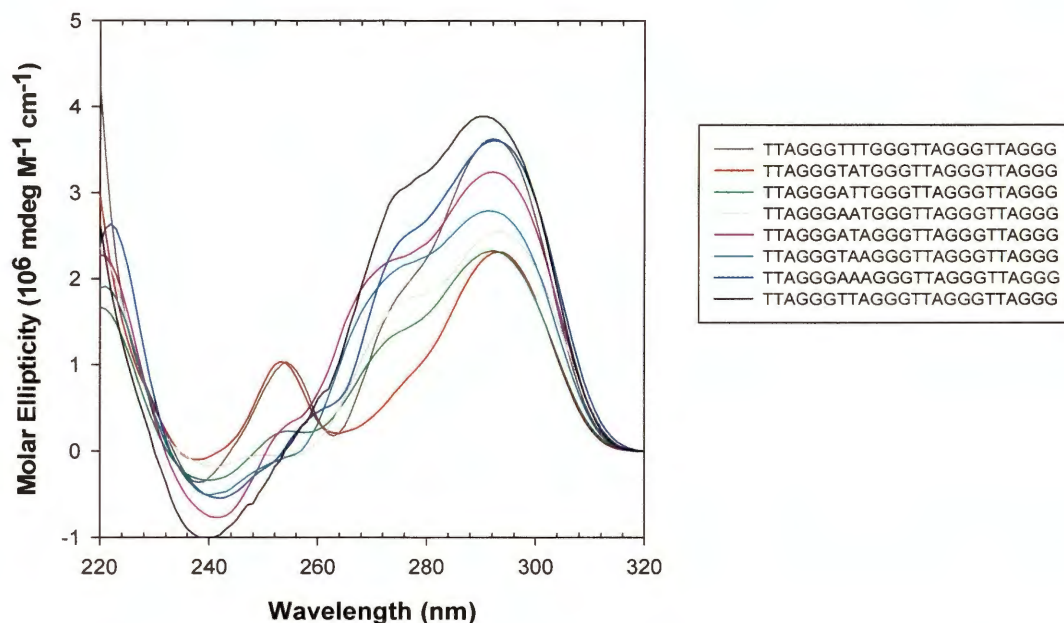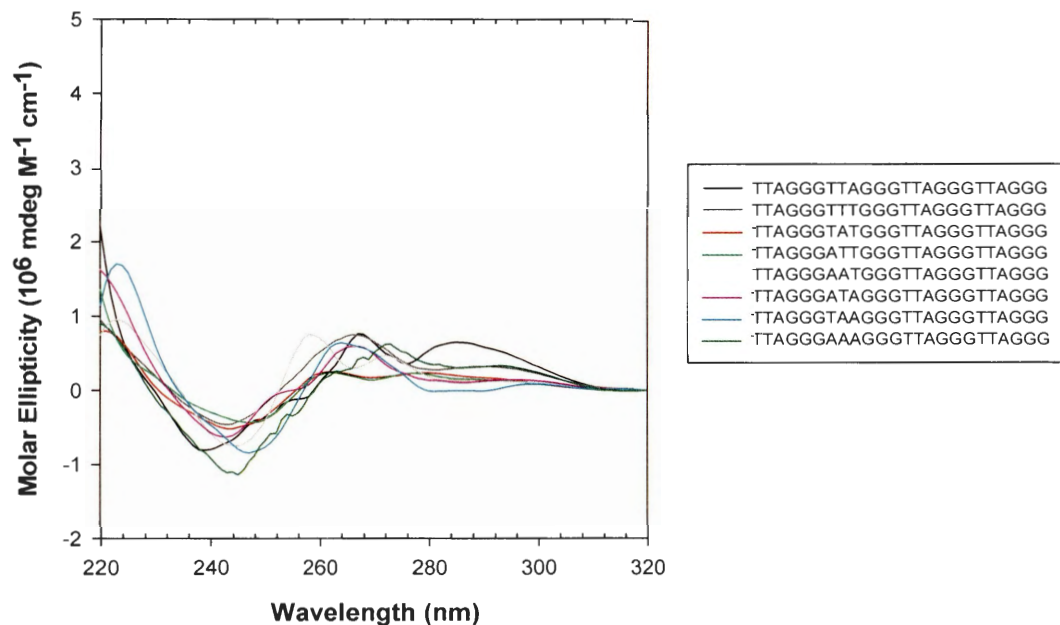FIGURE 3.8: CD spectra of the loop 1 oligomers at 95°C. Samples are prepared in potassium buffer at room temperature.

Figure 3.8 shows CD spectra of all the oligomers at 95°C given in Table 3.3. As can be seen here, the loop 1 series also have signature CD spectra similar to HTEL1. Differences in band intensities may be attributed to loop sequence as we know that the DNA is completely unfolded at 95°C. Melting profiles of all the oligomers are shown in Figure 3.9 and their corresponding T_m values are given in Table 3.4.
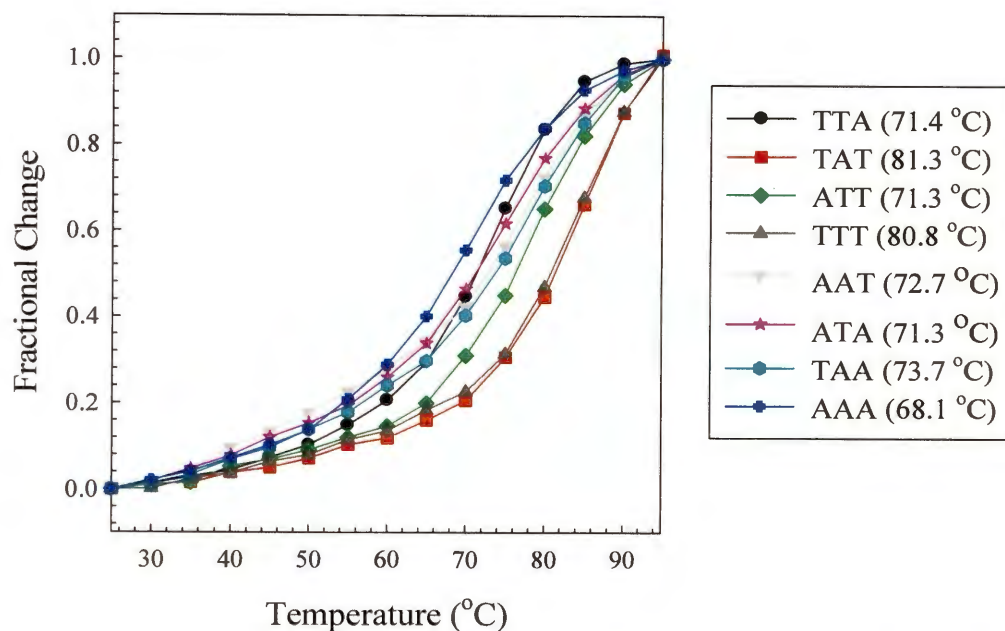
FIGURE 3.9: Melting profile for all loop 1 sequences. The T_m value for each melt is estimated at 290 nm.

TABLE 3.4: Number of "T" and number of "A" bases in the oligomers, T_m, hyperellipticity values along with wavelength for peak, shoulder and trough.

	# T	# A	λ1 (Trough)	λ2 (Shoulder)	λ3 (Peak)	Loop Seq.	Melting Point (°C)	Hyper ellipticity
HTEL1	8	4	240	272	291	TTA	71.4	10.3
HTEL-1LA	9	3	238	273	292	TTT	80.8	5.60
HTEL-1LB	8	4	238	273	293	TAT	81.3	6.44
HTEL-1LC	8	4	239	274	292	ATT	71.3	13.3
HTEL-1LD	7	5	241	274	293	AAT	72.7	18.7
HTEL-1LA2	7	5	241	272	292	ATA	71.3	24.6
HTEL-1LA3	7	5	243	273	291	TAA	73.7	20.1
HTEL-1LA4	6	6	242	275	292	AAA	68.1	32.4

TAT (81.3°C) ≈TTT (80.8°C)

>TAA (73.7°C) ≈ AAT (72.7°C) >TTA (71.4°C) ≈ATT (71.3°C) ≈ATA (71.3°C)

>AAA (68.1°C)

The data in Table 3.4 show a relationship similar to Table 3.2, which indicates that the more number of T bases, the more stability and the –AAA- sequence is less stable.

Loop 3

To get a better understanding of loop-loop interactions and loop-G-tetrads interaction, we also studied loop 3 sequences. CD spectra of eight oligomers of loop 3 (Table 3.5) at 25°C are given in Figure 3.10.

TABLE 3.5: Loop 3 variants

	Oligomer	Sequence
	HTEL1	TTAGGGTTAGGGTTAGGG**TTA**GGG
	HTEL-3LA	TTAGGGTTAGGGTTAGGG**TAT**GGG
	HTEL-3LB	TTAGGGTTAGGGTTAGGG**ATT**GGG
	HTEL-3LC	TTAGGGTTAGGGTTAGGG**TTT**GGG
Loop 3	**HTEL-3LD**	TTAGGGTTAGGGTTAGGG**TAA**GGG
	HTEL-3LE	TTAGGGTTAGGGTTAGGG**AAT**GGG
	HTEL-3LF	TTAGGGTTAGGGTTAGGG**ATA**GGG
	HTEL-3LG	TTAGGGTTAGGGTTAGGG**AAA**GGG

FIGURE 3.10: CD spectra of loop 3 sequences at 25°C. All the oligomers are prepared in potassium buffer at room temperature.

All the CD spectra in Figure 3.10 are similar with CD spectra of HTEL1. They all have peaks near 290 nm, shoulders near 270 nm and troughs near 240 nm. The oligomer with loop 3 sequence –TAT- displays another peak near 250 nm which suggests the possibility of chair type anti parallel structure. CD spectra of loop 3 sequences at 95°C are also given in Figure 3.11. Differences in the band intensities of spectra are attributed to the loop sequences as DNA unfolds completely at 95°C.

31

Legend:
- TTAGGGTTAGGGTTAGGG**TAT**GGG
- TTAGGGTTAGGGTTAGGG**ATT**GGG
- TTAGGGTTAGGGTTAGGG**TTT**GGG
- TTAGGGTTAGGGTTAGGG**TTA**GGG
- TTAGGGTTAGGGTTAGGG**TAA**GGG
- TTAGGGTTAGGGTTAGGG**AAT**GGG
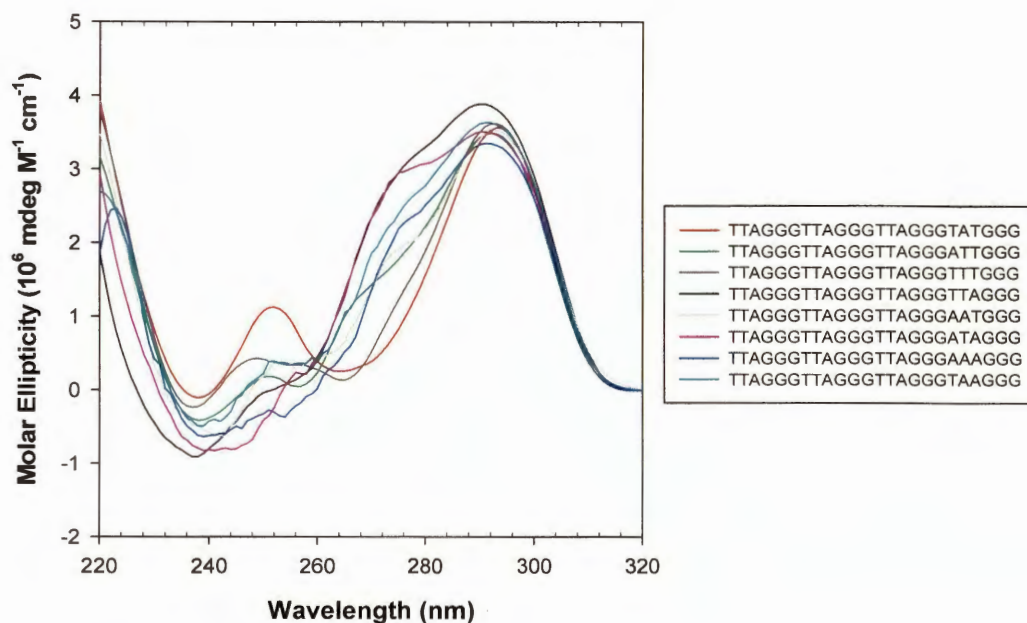- TTAGGGTTAGGGTTAGGG**ATA**GGG
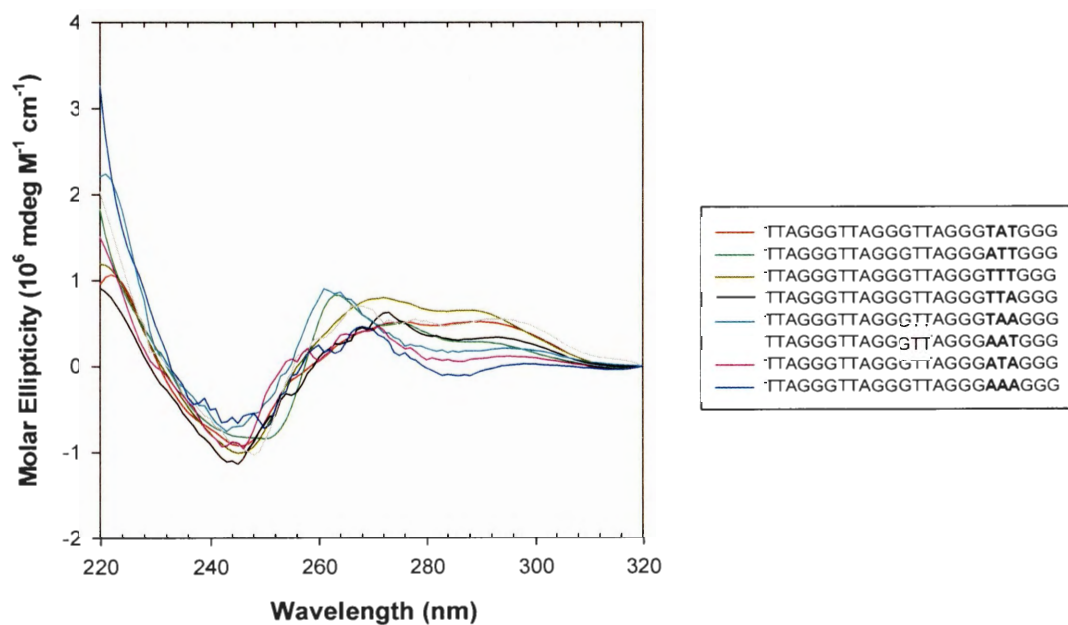- TTAGGGTTAGGGTTAGGG**AAA**GGG

FIGURE 3.11: CD spectra of Loop 3 sequences collected at 95°C. All the oligomers are prepared in potassium buffer at room temperature.

Melting profile of all the oligomers of loop 3 have been shown in Figure 3.12. T_m values are calculated and given in the Table 3.6.
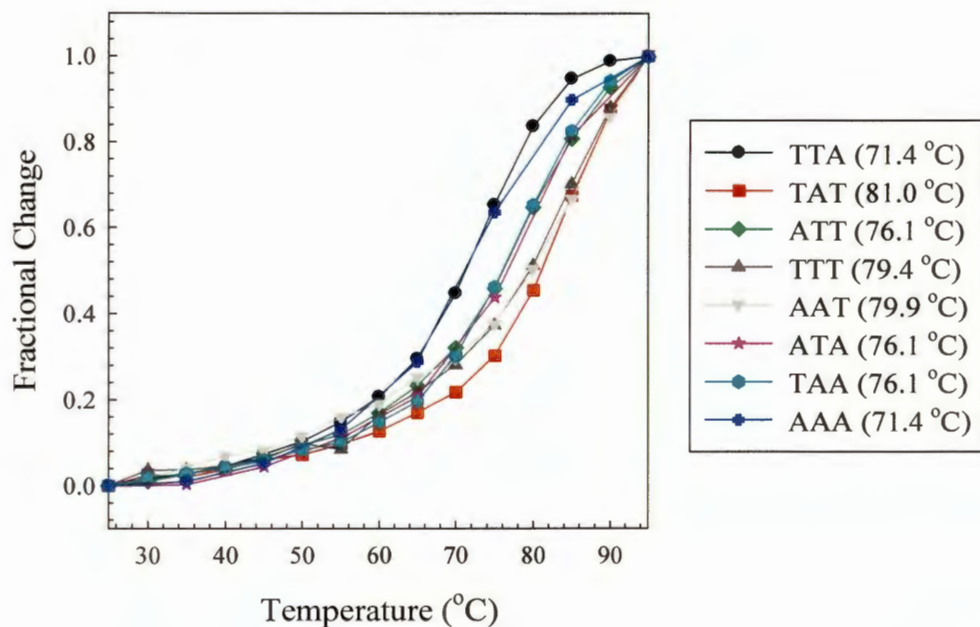
FIGURE 3.12: Melting curve for loop 3 sequences at 290 nm.

TABLE 3.6: T_m, hyperellipticity and wavelength (of peak, shoulder and trough)

	# T	# A	$\lambda 1$ (Trough)	$\lambda 2$ (Shoulder)	$\lambda 3$ (Peak)	Loop Seq.	Melting Point (°C)	Hyper Ellipticity
HTEL1	8	4	240	272	291	TTA	71.4	10.3
HTEL-3LA	8	4	238	-	293	TAT	81	6.16
HTEL-3LB	8	4	238	268	292	ATT	76.1	11.7
HTEL-3LC	9	3	237	275	293	TTT	79.4	5.34
HTEL-3LD	8	5	240	273	291	TAA	76.1	18.5
HTEL-3LE	7	5	242	271	292	AAT	79.9	9.47
HTEL-3LF	7	5	243	275	290	ATA	76.1	38.5
HTEL-3LG	7	6	243	274	291	AAA	71.4	86.2

TAT (81.0°C) ≈ TTT (79.4°C) ≈ AAT (79.9°C)

> ATA (76.1°C) ≈ ATT (76.1°C) ≈ TAA (76.1°C)

> TTA (71.4°C) ≈ AAA (71.4°C)

In the loop 2 and loop 1 variants we saw that T_m values increases with more number of T bases but in loop 3 it is very random. Five sequences have almost same T_m value in regardless of the number of T or A bases.
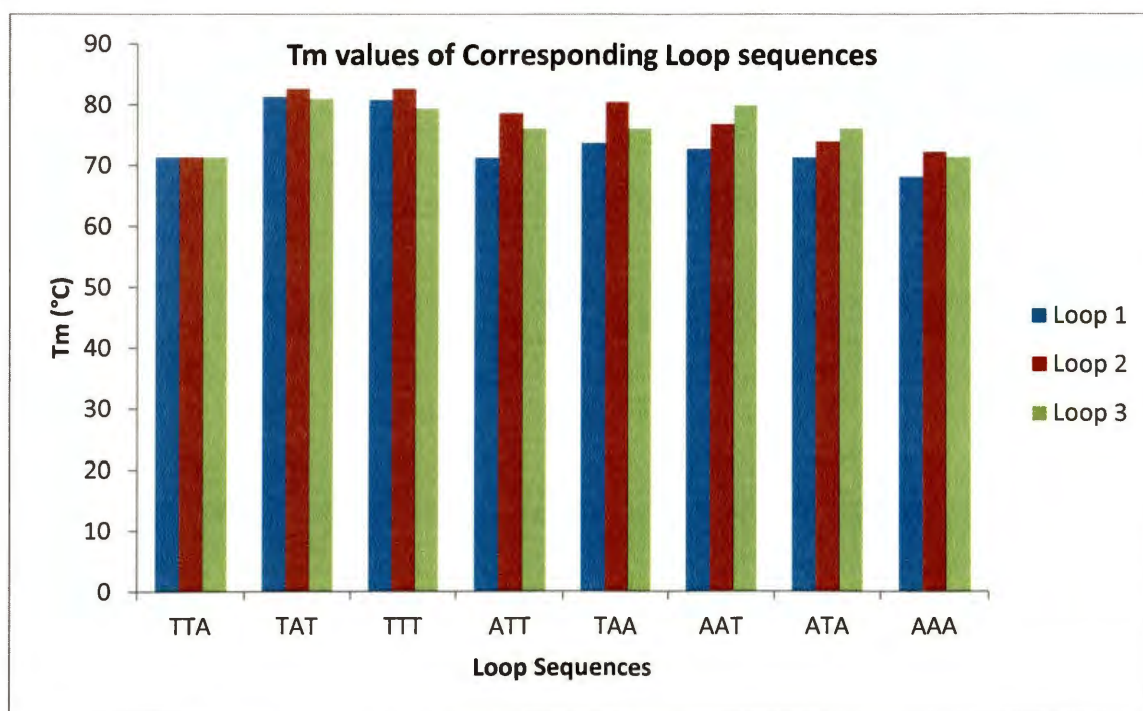

FIGURE 3.13: Comparison of T_m values (°C) between all three loops.

The sequence with all A bases has the lowest T_m while the sequence with –TAT- has the highest, similar to loop 2 and loop 1. Loop sequences of –ATT-, -TAA-, -AAT- and –ATA- give similar T_m values. The reason could be the non interaction of this lateral loop with G-tetrads and other loops. A comparative study of T_m values for three loops is given in Figure 3.13.

It is very clear from the bars given in Figure 3.13 that loop 2, in general, results in higher values of T_m than the other two. From this we can conclude that position of loop 2 plays a major role in the conformational stability of the quadruplexes. Further, this stability differs with different sequences within loop 2 as they all have a different T_m values. Furthermore, loops with –TAT- and –TTT- sequences have the highest stabilities whereas the wild type sequence, -TTA-, has, in general the least stability.

CHAPTER IV

SUMMARY AND CONCLUSION

In this study, we analyzed the effects of the base substitutions in all loops for intramolecular quadruplexes related to the human telomeric sequence (TTAGGG)$_4$. Previous studies have indicated that permuting all –TTA- segments influences the conformation and stability of the resultant folded structure (34). In this study, we systematically permuted the -TTA- segments that comprise the loops of the folded structure. The results presented clearly demonstrate that even small changes in the sequence influences both the structure and stability of the resultant quadruplex and that the changes are dependent upon both the sequence context of the permutation and the location of the permutation.

In general, all sequences fold into quadruplexes similar to the wild type as indicated by the similarities in CD signatures. This is strictly true for the loop 2 variants, all of which have a peak at around 290 nm, a shoulder at 270 nm and a shallow trough at 240 nm. Loop 1 and loop 3 variants also have similar CD spectra, and therefore similar structures to the wild type sequence, with the exception of the –TTT- and –TAT- variants. These sequences also have an additional minor peak at around 250 nm and are missing the shoulder at 270 nm. These differences can be attributed to two possibilities: 1) the folded structure has a single conformation different from the wild type sequence

36

or 2) the folded structure is a mixture of 2 or more conformations. Without high resolution structural analysis (such as NMR), we cannot differentiate the two possibilities at this time.

Although most of the folded structures have similar CD spectra, the spectra are not identical. The slight variations in peak intensities and wavelengths are due to the sequence context variations. First and foremost, A bases give different CD spectra than T bases (35). Thus, changing even a single T to an A (or vice versa) will change the CD signature. The second factor that contributes to the variations is due to the stacking potential of the bases. The stacking of the bases allows for the transition dipole moments to interact. The nature of the interaction (orientation of the bases, overlap of the π skeletons, etc) influences the resultant spectrum (35). Thus, the observed variations in peak intensities and wavelengths will also depend upon: 1) how the bases in the loops are stacking with each other; and 2) how the bases in the loops are stacking with the G-tetrads (see below). Again, without high resolution structural data, we cannot assess base stacking from CD spectra alone.

The various permutations in sequence context also influences apparent stability, based upon the T_m values reported, as a function of base content of the permutation and location of the permutation. In general, the loop 2 variants are the most stable within a sequence permutation. Within each loop, variants with a

T base in the first and third positions (-TAT- and –TTT-) are the most stable. Finally, the wild type sequence (-TTA-) is the least stable oligomer.

In studies examining DNA hairpins of sequence (CGATACX$_4$GTATCC, where X =A, T, G, or C), it was demonstrated that the stability of the hairpin stem was influenced by the sequence context of the hairpin loop (36). The stability hierarchy was dependent upon the end-loop sequence in the following order: $T_4 > C_4 > G_4 > A_4$. In other words, pyrimidine bases are more stabilizing that purine bases. Consider, for example, loop 3. Loop 3 could be modeled as a DNA hairpin with an X_3 loop where X = T and/or A. If this hairpin then behaved as a normal DNA hairpin, it would be most stable if X was T and only T and least stable if X was A and only A. Further extrapolation would suggest that if the first base and the third base were both T, regardless of the identity of the second base, that the hairpin would be more stable than if all A bases, or even –ATA-. This is exactly what is observed. All sequence variants with the –TTT- and –TAT- permutations are the most stable. Further, variants with A bases in the first and third position (-AAA- and –ATA-) are not as stable. Our previous study has shown that when all bases in all loops are –AAA-, the resultant folded structure is the least stable of all others (34).

In the proposed mechanism for the unfolding of the wild type sequence, we proposed a three state transition whereby loop 2 would act as a hinge in the first step for the formation of a double hairpin. In the second step, the hairpins unfold to the single

strand. One of the aims of the overall project is to verify that mechanism. The results presented do not necessarily support or oppose the mechanism. However, it is interesting that the loop 2 variants were, in general, the most stable of all investigated. What is perhaps more interesting, biologically, is that the wild type sequence is the least stable of all. Since the unfolding of the quadruplexes formed at the ends of chromosomes in the human telomere is necessary for DNA replication, it should not be surprising that mother nature selected –TTAGGG- for the human telomere repeat. To fully answer the questions raised by these results, calorimetric studies of all oligos need to be carried out.

REFERENCES

1. Watson, J.D. and Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids – A Structure for Deoxyribose Nucleic Acid. *Nature*, **171**, 737-738.

2. Zamenhof, S., Brawermann, G. and Chargaff, E. (1952). On the Deoxypentose Nucleic Acids from Several Microorganisms, *Biochim. Biophys. Acta.*, **9**, 402-405.

3. Pohl, F. M. and T. M. Jovin. (1972). Salt-Induced Co-Operative Conformational Change of Synthetic DNA: Equilibrium and Kinetic Studies with Poly(dG-dC). *J. Mol. Biol.*, **67**, 375-396.

4. Wang, A. J.-H., Quigley, G. J., Kolpak, F. J., Crawford, J. L., Van Boom, J. H., Vander, M. G. and Rich, A. (1979). Molecular Structure of a Left-Handed Double Helical DNA Fragment at Atomic Resolution. *Nature*, **282**, 680–686.

5. Belmont, P., Constant, J. F. and Demeunynck, M. (2001). Nucleic Acid Conformation Diversity: from Structure to Function and Regulation. *Chem. Soc. Rev.*, **30**, 70-81.

6. Muller, H. J. (1938) The Remaking of Chromosomes. *Collecting Net*, **13**, 182-198.

7. Wellinger, R. J., Sen, D. (1997). The DNA Structure at the Ends of Eukaryotic Chromosomes. *European Journal of Cancer*, **33**, 735-749.

8. Olaussen, K. A., Dubrana, K., Domont, J., Spano, J. P., Sabatier, L., Soria, J. C. (2006). Telomeres and Telomerase as Target for Anticancer Drug Development. *Critical Review in Onco. Hemato.*, **57**, 191-214.

9. Neidle, S. and Parkinson, G. N. (2003). The Structure of Telomeric DNA. *Current Opinion in Structural Biology*, **13**, 275-283.

10. Williamson, J. R. (1994). G-Tetrad Structures in Telomeric DNA. *Annual Reviews in Biophysical and Biomolecular Structure,* **23,** 703-730.

11. Dunham, M. A., Neumann, A. A., Fasching, C. L. and Reddel, R. R. (2000). Telomere Maintenance by Recombination in Human Cells. *Nat Genet*, **26**, 247-250.

12. Sun, D. and Hurley, L.H. (2001). Targeting Telomers and Telomerase. *Meth. Enzymol.*, **340**, 573 -592.

13. Han, H. and Hurley, L.H. (2000). G-quadruplex DNA: A Potential Target for Anti-cancer Drug Design. *TIPS,* **21**, 136-142.

14. Jenkins, T.C. (2000). Targeting Multi-stranded DNA Structures. *Current Medicinal Chemistry*, **7**, 99-115.

15. Neidle, S. and Read, M.A. (2001). G-quadruplexes as Therapeutic Agents. *Biopolymers*, **56**, 195-208.

16. Phan, A. T., Modi, Y.S., and Patel, D. J. (2004). Propeller Type Parallel-Stranded G-Quadruplex in the Human c-MYC Promoter. *J. Am. Chem. Soc.*, **126**, 8710-8716.

17. Guo, K., Gokhale, V., Hurley, L. H. and Sun, D. (2008). Intramolecularly Folded G-quadruplex and i-motif Structures in the Proximal Promoter of the Vascular Endothelial Growth Factor Gene. *Nucleic Acid Research*, **36**, 4598-4608.

18. Dai, J., Dexheimer, T.S., Chen, D., Carver, M., Ambrus, A., Jones, R.A., Yang, D. (2006). An Intramolecular G-quadruplex Structure with Mixed Parallel/antiparallel G-strands Formed in the Human BCL-2 Promoter Region in Solution. *J. Am. Chemical Soc.,* **126**, 1096-1108.

19. Phan, A. T., and Mergny, J. L. (2002). Human Telomeric DNA: G- Quadruplex, I-motif and Watson-Crick Double Helix. *Nucleic Acids Res.,* **30**, 4618-4625.

20. Venczel, E.A. and Sen, D. (1993). Parallel and Antiparallel G-DNA Structures from a Complex Telomeric Sequence. *Biochemistry,* **32**, 6220-622

21. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. and Neidle, S. (2006). Quadruplex DNA: Sequence, Topology and Structure. *Nucleic Acids Research,* **34**, 5402-5415.

22. Ambrus, A., Chen, D., Dai, J., Bialis, T., Jones, R. and Yang, D. (2006). Human Telomeric Sequence forms a Hybrid-type Intramolecular G-quadruplex with Mixed Parallel/antiparallel Strands in Potassium Solution. *Nucleic Acids Research,* **34**, 2723-2735.

23. Luu, K.N., Phan, A.T., Kuryavyi, V., Lacroix, L. and Patel, D. (2006). Structure of the Human Telomer in K^+ Solution: an Intramolecular (3+1) G-quadruplex Scaffold. *Journal of the American Chemical Society,* **128**, 9963-9970.

24. Xu, Y., Noguchi, Y., and Sugiyama, H. (2006). The New Models of the Human Telomere d[AGGG(TTAGGG)₃] in K^+ Solution. *Bio-organic and Medicinal Chemistry,* **14**, 5584-5591.

25. Phan, A. T., Kuryavyi, V., Luu, K. N., and Patel, D. J. (2007). Structure of Two Intramolecular G-quadruplexes Formed by the Natural Human Telomere Sequence in K^+ Solution. *Nucleic Acids Research,* **35**, 6517-6525.

26. Lim, K. W., Amrane, S., Bouaziz, S., Xu, W., Mu, Y., Patel, D. J., Luu, K. N., Kim, N. and Phan, A. T. (2009). Structure of the Human Telomere in K^+ Solution: A Stable Basket-Type G-quadruplex with only Two G-tetrad Layers. *Journal of The American Chemical Society,* **131**, 4301-4309.

27. Kuryavyi, V. and Patel, D. (2010). Solution Structure of a Unique G-quadruplex Scaffold Adopted by a Guanosine-rich Human Intronic Sequence. *Structure,* **18**, 73-82.

28. Zhang, Z., Dai, J., Veliath, E., Jones, R. A. and Yang, D. (2010). Structure of a Two-G-tetrad Intramolecular G-quadruplex Formed by a Variant Human Telomeric Sequence in K^+ Solution. *Nucleic Acids Research,* **38**, 1009-1021.

29. Viglasky, V., Bauer, L. and Tluckova, K. (2010). Structural Features of Intra- and Intermolecular G-quadruplexes Derived from the Human Telomere. *Biochemistry*, **49**, 2110-2120.

30. Guedin, A., Cian, A. D., Gros, J., Lacroix, L. and Mergny, J. L. (2008). Sequence Effects in Single Base Loops for Quadruplexes. *Biochemie,* **90**, 686-696.

31. Wang, Y., and Patel, D. J. (1993). Solution Structure of the Human Telomeric Repeat d[AG3(T2AG3)3] G-Tetraplex. *Structure,* **1**, 263-282.

25. Phan, A. T., Kuryavyi, V., Luu, K. N., and Patel, D. J. (2007). Structure of Two Intramolecular G-quadruplexes Formed by the Natural Human Telomere Sequence in K^+ Solution. *Nucleic Acids Research,* **35**, 6517-6525.

26. Lim, K. W., Amrane, S., Bouaziz, S., Xu, W., Mu, Y., Patel, D. J., Luu, K. N., Kim, N. and Phan, A. T. (2009). Structure of the Human Telomere in K^+ Solution: A Stable Basket-Type G-quadruplex with only Two G-tetrad Layers. *Journal of The American Chemical Society,* **131**, 4301-4309.

27. Kuryavyi, V. and Patel, D. (2010). Solution Structure of a Unique G-quadruplex Scaffold Adopted by a Guanosine-rich Human Intronic Sequence. *Structure,* **18**, 73-82.

28. Zhang, Z., Dai, J., Veliath, E., Jones, R. A. and Yang, D. (2010). Structure of a Two-G-tetrad Intramolecular G-quadruplex Formed by a Variant Human Telomeric Sequence in K^+ Solution. *Nucleic Acids Research,* **38**, 1009-1021.

29. Viglasky, V., Bauer, L. and Tluckova, K. (2010). Structural Features of Intra- and Intermolecular G-quadruplexes Derived from the Human Telomere. *Biochemistry,* **49**, 2110-2120.

30. Guedin, A., Cian, A. D., Gros, J., Lacroix, L. and Mergny, J. L. (2008). Sequence Effects in Single Base Loops for Quadruplexes. *Biochemie,* **90**, 686-696.

31. Wang, Y., and Patel, D. J. (1993). Solution Structure of the Human Telomeric Repeat d[AG3(T2AG3)3] G-Tetraplex. *Structure,* **1**, 263-282.

32. Dapic, V., Abdomerovic, V., Marrington, R., Peberdy, J., Rodger, A., Trent, J. O. and Bates, P. J. (2003). Biophysical and Biological Properties of Quadruplex Oligodeoxyribonucleotides. *Nucleic Acid Research,* **31**, 2097-2107.

33. Antonacci, C., Chaires, J. B., Sheardy, R. D. (2007). Biophysical Characterization of the Human Telomeric (TTAGGG)$_4$ Repeat in a Potassium Solution. *Biochemistry,* **46**, 4654-4660.

34. Tucker, B. A., Gabriel, S., Sheardy, R. D. (2011). A CD Spectroscopic Investigation of Intermolecular and Intramolecular DNA Quadruplexes. *A. Chem. Soc.,* **1082,** 51-67.

35. Cantor, C. R. & Schimmel, P. R. (1980). The Confirmation of Biological Molecules. *Biophysical Chemistry*, 409-432.

36. Paner, T. M., Amaratunga, M., Doktycz, M. J. & Benight, A. S. (1990). Analysis of Melting Transitions of the DNA Hairpins Formed form the Oligomer Sequence d[GGATAC(X)$_4$GTATCC] (X = A, T, G, C). *Biopolymers,* **29**, 1715-1734.

37. Wheeler, R. (2007). The Structure of A-, B- and Z-DNA Form. *Wikimedia Commons,* 1.

38. Seidman, M. M. & Glazer, P. M. (2003). The Potential for Gene Repair via Triple Helix Formation. *J. Clin. Invest.,* **112**, 487-494.