

2010

Stability of Nucleic Acid Secondary Structures and Their Contribution to Gene Expression

Maged A. Darwish
Seton Hall University

Follow this and additional works at: <https://scholarship.shu.edu/dissertations>

 Part of the [Biochemistry Commons](#)

Recommended Citation

Darwish, Maged A., "Stability of Nucleic Acid Secondary Structures and Their Contribution to Gene Expression" (2010). *Seton Hall University Dissertations and Theses (ETDs)*. 1435.
<https://scholarship.shu.edu/dissertations/1435>

**Stability of nucleic acid secondary structures and their
contribution to gene expression**

*This dissertation is submitted to the faculty of the Department of Chemistry and
Biochemistry in the School of Arts and Sciences in partial fulfillment of the requirements*

for the degree of

Doctor of Philosophy

at

Seton Hall University

400 South Orange Avenue

South Orange, New Jersey 07079

by

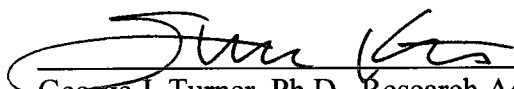
Maged A. Darwish

December 2009

CERTIFICATION

We certify that we have read this thesis and that in our opinion it is adequate in scientific scope and quality as a dissertation for the degree of Doctor of Philosophy.

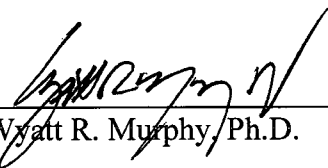
APPROVED



George J. Turner, Ph.D., Research Advisor
Stephen P. Kelty, Ph.D., Research Co-advisor



Cosimo Antonacci, Ph.D.



Wyatt R. Murphy, Ph.D.



Cecilia H. Marzabadi, Ph.D.



Stephen P. Kelty, Ph.D.
Chairman, Seton Hall University Department of Chemistry and Biochemistry

DEDICATION

For my family, thank you for your contributions, support and patience.

Table of Contents

1	Introduction.....	1
1.1	Nucleic Acids (DNA and RNA).....	1
1.1.1	Primary Structure of Nucleic Acids.....	3
1.1.2	Secondary Structure of Nucleic Acids.....	4
1.1.3	Tertiary Structure of Nucleic Acids.....	6
1.1.4	Factors contributing to the stability of nucleic acids.....	6
1.2	The Central Dogma of Molecular Biology.....	7
1.3	Halobacterium salinarum.....	8
1.4	Bacteriorhodopsin (bR).....	9
1.5	The bacterio-opsin gene (bop).....	11
1.5.1	Nucleic acid sequences can contribute to gene expression.....	11
1.5.2	Secondary structural element (Stem-loop).....	12
1.6	Research Purpose.....	12
2	Theory.....	13
2.1	Spectroscopy.....	13
2.1.1	Light.....	13
2.1.2	Electromagnetic radiation spectrum.....	13
2.1.3	Ultraviolet Absorption.....	14
2.1.4	Hyperchromicity.....	20
2.1.5	Melting temperature T_M	21
2.1.6	Circular Dichroism (CD) Spectroscopy.....	21
2.2	Spectroscopic Data Analysis.....	27
2.2.1	van't Hoff Analysis.....	27
2.2.2	Equilibrium Melting Curve.....	28
2.2.3	α Plots.....	29
2.2.4	Singular Value Decomposition (SVD).....	30
2.3	Differential Scanning Calorimetry (DSC).....	35
2.3.1	Principles of DSC.....	35
2.3.2	DSC Instrumentation.....	37
2.3.3	Data Analysis.....	39
2.4	Predictive Structural Analysis.....	40
3	Materials and Methods.....	42
3.1	Materials.....	42
3.1.1	DNA oligo.....	42
3.1.2	RNA oligo.....	42
3.1.3	Preparation of the Stem-loop.....	42
3.1.4	Buffers.....	43
3.1.5	Cleaning Solutions.....	44
3.2	Methods.....	45
3.2.1	Ultraviolet-visible spectroscopy (UV/Vis).....	45
3.2.2	Circular Dichroism Spectropolarimetry (CD).....	45
3.2.3	Differential Scanning Calorimeter (DSC).....	46
3.2.4	Singular-Value Decomposition (SVD).....	47
4	Results.....	48
4.1	Stem-loop Purity.....	48

4.2	Predictive Structural Analysis.....	49
4.2.1	Nucleic acid sequence analysis.....	49
4.2.2	<i>In silico</i> sequence analysis.....	49
4.3	Biophysical analysis of the DNA model sequence.....	54
4.3.1	UV Spectroscopy.....	54
4.3.2	Circular Dichroism Spectroscopy.....	68
4.3.3	Differential Scanning Calorimetry.....	96
4.4	Biophysical analysis of the RNA model sequence.....	105
4.4.1	UV Spectroscopy.....	106
4.4.2	Circular Dichroism Spectroscopy.....	109
5	Discussion.....	119
5.1	Predictive Structural Analysis.....	120
5.2	UV Spectroscopy.....	120
5.3	Circular Dichroism Spectroscopy.....	121
5.4	Differential Scanning Calorimetry.....	122
6	References.....	123

List of Abbreviations and Definitions

bR	bacteriorhodopsin
bop	bacterio-opsin gene
Bop	bacterio-opsin protein
UV	Ultraviolet region of the electromagnetic spectrum
VIS	Visible region of the electromagnetic spectrum
CD	Circular Dichroism
7-TM	heptahelical transmembrane
DSC	Differential Scanning Calorimetry
ΔC_p	Change in Heat Capacity
ΔH	Enthalpy
KPBS	Potassium Phosphate Buffered Solution
SVD	Singular Value Decomposition
TBE	Tris-Borate Buffer
T_M	Melting Temperature
1X	Working concentration for a given solution
10X	Concentrated stock solution 10 folds the working concentration
λ	wavelength
ν	frequency

List of Figures

Figure 1: Five-carbon sugars in DNA (2-deoxy-D-ribose) and RNA (D-ribose).....	2
Figure 2: The structure of common nucleotides in RNA and DNA	3
Figure 3: Nucleic acids primary structure (7)	4
Figure 4: DNA secondary structure (7)	5
Figure 5: RNA secondary structures.....	6
Figure 6: The Central Dogma of Molecular Biology (7).....	8
Figure 7: The three domains of life forms	8
Figure 8: Bacteriorhodopsin (bR) expression.....	10
Figure 9: Homotrimers aggregate and expand across the membrane forming the purple membrane	10
Figure 10: The wavelength of electromagnetic radiation	13
Figure 11: The electromagnetic spectrum.....	14
Figure 12: Energy levels of a molecule	15
Figure 13: Possible electronic transitions of σ , π and n electrons	16
Figure 14: Light transmission through a sample (8).....	17
Figure 15: Schematic of Cary 3E spectrophotometer	20
Figure 16: Melting curve	21
Figure 17: Light electric and magnetic fields	22
Figure 18: Light electric field vectors (A) unpolarized and (B) polarized	22
Figure 19: Plane polarized relative to (A) right and (B) left circularly polarized light.....	23
Figure 20: Diagram of CD spectropolarimeter	25
Figure 21: CD spectra of different DNA conformations	26
Figure 22: van't Hoff plot.....	27
Figure 23: Melting temperature at 100mM salt concentration	28
Figure 24: Alpha plot.....	29
Figure 25: Eigenvector v.....	31
Figure 26: Heat Capacity	35
Figure 27: Schematic diagram of a typical power compensation DSC	37
Figure 28: Typical Raw DSC scans	38
Figure 29: Calorimetric enthalpy and entropy calculation plot	40
Figure 30: 10% TBE-Urea polyacrylamide gel	48
Figure 31: 5' end of the bop gene mRNA.....	49
Figure 32: Mfold DNA model predicted secondary structures at 37°C, 1M salt.....	50
Figure 33: Mfold DNA model predicted secondary structures at 25°C.....	51
Figure 34: Mfold DNA model predicted secondary structures at 37°C.....	51
Figure 35: Mfold DNA model predicted secondary structures at 50°C, 1M salt.....	52
Figure 36: Mfold RNA model predicted secondary structures.....	53
Figure 37: UNAFold RNA model predicted secondary structures at 50°C, 1M salt.....	53
Figure 38: Repetitive DNA model heating and cooling scans at 260nm in 1M [K ⁺].....	55
Figure 39: Values of 1/T _M plotted Vs the natural log of the [DNA] in 1M K ⁺	55
Figure 40: UV spectra of the DNA model of the stem-loop at 260 nm as a function of temperature at different salt concentrations.....	56
Figure 41: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 15 mM K ⁺ , (B) van't Hoff plot.....	58

Figure 42: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 50 mM K ⁺ , (B) van't Hoff plot	59
Figure 43: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 100 mM K ⁺ , (B) van't Hoff plot	60
Figure 44: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 250 mM K ⁺ , (B) van't Hoff plot	61
Figure 45: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 500 mM K ⁺ , (B) van't Hoff plot	62
Figure 46: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 1000 mM K ⁺ , (B) van't Hoff plot	63
Figure 47: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 2000 mM K ⁺ , (B) van't Hoff plot	64
Figure 48: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 3000 mM K ⁺ , (B) van't Hoff plot	65
Figure 49: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 4000 mM K ⁺ , (B) van't Hoff plot	66
Figure 50: CD spectrum of DNA stem-loop in 150 mM K ⁺ at 5°C	68
Figure 51: Temperature dependent CD spectra of DNA stem-loop in 15 mM K ⁺	69
Figure 52: Temperature dependent CD spectra of DNA stem-loop in 50 mM K ⁺	69
Figure 53: Temperature dependent CD spectra of DNA stem-loop in 100 mM K ⁺	70
Figure 54: Temperature dependent CD spectra of DNA stem-loop in 250 mM K ⁺	70
Figure 55: Temperature dependent CD spectra of DNA stem-loop in 500 mM K ⁺	71
Figure 56: Temperature dependent CD spectra of DNA stem-loop in 1000 mM K ⁺	71
Figure 57: Temperature dependent CD spectra of DNA stem-loop in 2000 mM K ⁺	72
Figure 58: Temperature dependent CD spectra of DNA stem-loop in 3000 mM K ⁺	72
Figure 59: Temperature dependent CD spectra of DNA stem-loop in 4000 mM K ⁺	73
Figure 60: Salt dependent spectra of DNA stem-loop at 5°C	74
Figure 61: Salt dependent spectra of DNA stem-loop at 40°C	74
Figure 62: Salt dependent spectra of DNA stem-loop at 95°C	75
Figure 63: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 15 mM K ⁺ , (B) van't Hoff plot	76
Figure 64: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 50 mM K ⁺ , (B) van't Hoff plot	77
Figure 65: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 100 mM K ⁺ , (B) van't Hoff plot	78
Figure 66: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 250 mM K ⁺ , (B) van't Hoff plot	79
Figure 67: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 500 mM K ⁺ , (B) van't Hoff plot	80
Figure 68: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 1000 mM K ⁺ , (B) van't Hoff plot	81
Figure 69: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 2000 mM K ⁺ , (B) van't Hoff plot	82
Figure 70: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 3000 mM K ⁺ , (B) van't Hoff plot	83

Figure 71: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 4000 mM K ⁺ , (B) van't Hoff plot	84
Figure 72: Representative 3D plot of the CD spectra	86
Figure 73: (A) Spectral Eigenvector of CD melting of DNA in 15mM K ⁺ , (B) Kinetic Eigenvector	87
Figure 74: (A) Spectral Eigenvector of CD melting of DNA in 50mM K ⁺ , (B) Kinetic Eigenvector	88
Figure 75: (A) Spectral Eigenvector of CD melting of DNA in 100mM K ⁺ , (B) Kinetic Eigenvector	88
Figure 76: (A) Spectral Eigenvector of CD melting of DNA in 250mM K ⁺ , (B) Kinetic Eigenvector	89
Figure 77: (A) Spectral Eigenvector of CD melting of DNA in 500mM K ⁺ , (B) Kinetic Eigenvector	89
Figure 78: (A) Spectral Eigenvector of CD melting of DNA in 1000mM K ⁺ , (B) Kinetic Eigenvector	90
Figure 79: (A) Spectral Eigenvector of CD melting of DNA in 2000mM K ⁺ , (B) Kinetic Eigenvector	90
Figure 80: (A) Spectral Eigenvector of CD melting of DNA in 3000mM K ⁺ , (B) Kinetic Eigenvector	91
Figure 81: (A) Spectral Eigenvector of CD melting of DNA in 4000mM K ⁺ , (B) Kinetic Eigenvector	91
Figure 82: A representative residual plot (singular values magnitude as a function of component number) of CD melting of DNA in 50mM K ⁺	92
Figure 83: V coefficients for each significant component as a function of temperature of CD melting of DNA in 50mM K ⁺	92
Figure 84: Residual contour plots for matrices generated with one singular value of CD melting of DNA in 50mM K ⁺	93
Figure 85: Residual contour plots for matrices generated with two singular value of CD melting of DNA in 50mM K ⁺	94
Figure 86: Residual contour plots for matrices generated with three singular values of CD melting of DNA in 50mM K ⁺	94
Figure 87: Residual contour plots for matrices generated with four singular values of CD melting of DNA in 50mM K ⁺	95
Figure 88: Residual contour plots for matrices generated with five singular values of CD melting of DNA in 50mM K ⁺	95
Figure 89: A representative repetitive DSC thermal melting scans (buffer-buffer) in 100 mM K ⁺	97
Figure 90: A representative repetitive DSC thermal melting scans (sample-buffer) in 100 mM K ⁺	97
Figure 91: A representative molar heat capacity versus temperature plot in 100 mM K ⁺ with a polynomial-linear baseline treatment	98
Figure 92: A representative DSC thermal plot in 100 mM K ⁺	98
Figure 93: Baseline corrected DSC thermograms at various potassium ions concentration	99
Figure 94: Deconvolution of a representative DSC thermogram obtained at 1M K ⁺ into two transitions to calculate calorimetric enthalpy for each transition	100

Figure 95: Deconvolution of a representative DSC thermogram obtained at 1M K ⁺ into two transitions to calculate calorimetric entropy for each transition.....	100
Figure 96: T _M dependency on salt concentration.....	103
Figure 97: ΔH dependency on salt concentration	104
Figure 98: ΔS dependency on salt concentration.....	104
Figure 99: Comparison between the T _M values from different techniques	105
Figure 100: Repetitive RNA model heating and cooling scans at 260 nm in 1M [K ⁺] ...	106
Figure 101: UV spectra at 260 nm of the RNA stem-loop model at different temperature and salt concentrations.....	107
Figure 102: A representative van't Hoff plot of RNA model UV melting in 1M K ⁺	108
Figure 103: Comparison between the RNA and DNA T _M values	109
Figure 104: CD spectrum of RNA stem-loop in 1M K ⁺ at 5°C.	110
Figure 105: Temperature dependent CD spectra of RNA stem-loop in 50 mM K ⁺	110
Figure 106: Temperature dependent CD spectra of RNA stem-loop in 250 mM K ⁺	111
Figure 107: Temperature dependent CD spectra of RNA stem-loop in 1M K ⁺	111
Figure 108: Temperature dependent CD spectra of RNA stem-loop in 2M K ⁺	112
Figure 109: A representative van't Hoff plot of RNA model CD melting in 1M K ⁺	113
Figure 110: A representative residual plot (singular values magnitude as a function of component number) for CD melting of RNA in 1M K ⁺	114
Figure 111: V coefficients for each significant component as a function of temperature for CD melting of RNA in 1M K ⁺	115
Figure 112: Residual contour plots for matrices generated with one singular value for CD melting of RNA in 50mM K ⁺	116
Figure 113: Residual contour plots for matrices generated with two singular values for CD melting of RNA in 50mM K ⁺	116
Figure 114: Residual contour plots for matrices generated with three singular values for CD melting of RNA in 50mM K ⁺	117
Figure 115: Residual contour plots for matrices generated with four singular values for CD melting of RNA in 50mM K ⁺	117
Figure 116: Residual contour plots for matrices generated with five singular values for CD melting of RNA in 50mM K ⁺	118

List of Tables

Table 1: KPBS stock solution preparation (10X)	43
Table 2: Working Potassium Phosphate Buffer Solution (KPBS) preparation	44
Table 3: van't Hoff thermodynamic parameters as a function of $[K^+]$	67
Table 4: van't Hoff thermodynamic parameters as a function of $[K^+]$	85
Table 5: Transition I thermodynamic parameters	101
Table 6: Transition II thermodynamic parameters	102
Table 7: DSC overall thermodynamic parameters	102
Table 8: van't Hoff thermodynamic parameters as a function of $[K^+]$	108
Table 9: van't Hoff thermodynamic parameters as a function of $[K^+]$	113

Equations and Equilibria

Equation 1:	c = λν.....	13
Equation 2:	Photon energy e = hv = hc/λ	14
Equation 3:	T = I / I ₀	16
Equation 4:	- dI ₀ / I ₀ = dA/A	17
Equation 5:	- ∫ _{I₀} I _f dI ₀ / I ₀ = (a/A) ∫ ₀ ^b dn.....	18
Equation 6:	ln I ₀ / I _F = (a/A)b	18
Equation 7:	log I _F / I ₀ = (a/(2.303)A)n	18
Equation 8:	log I _F / I ₀ = ((ab)/(2.303 V))n	18
Equation 9:	log I _F / I ₀ = ((ab)/(2.303 V))(6.02 x 10 ²³ c/1000).....	18
Equation 10:	A = log I _F / I ₀ = εbc	18
Equation 11:	Beer's law A = εbc.....	18
Equation 12:	ΔA = A _L -A _R	23
Equation 13:	θ _λ = arc tan ((minor ellipse axis)/(major ellipse axis)).....	23
Equation 14:	θ _λ = 2.303((ε _L - ε _R) 180)/4π deg.....	24
Equation 15:	θ _λ = 3300 (Δε) mdeg M ⁻¹ cm ⁻¹	24
Equation 16:	K = [Folded] / [Unfolded].....	29
Equation 17:	K = α / 1- α.....	30
Equation 18:	A _{mn} = U _{mm} S _{mn} V ^T _{nm}	31
Equation 19:	C(X _i) = ΣX _{j,i} X _{j+1,l}	34
Equation 20:	Ax = USxV ^T	34
Equation 21:	heat flow = q / t.....	36
Equation 22:	heating rate = ΔT / t	36
Equation 23:	q / ΔT = Cp.....	36
Equation 24:	dH = dU +pdV +Vdp	36
Equation 25:	dU = q - pdV	36
Equation 26:	dH = q	36
Equation 27:	dH = C _p ΔT	36
Equation 28:	ΔH = ∫ _{T₁} ^{T₂} C _p ΔT.....	36
Equation 29:	dS = δq/T.....	36
Equation 30:	dS = (C _p dT)/T	36
Equation 31:	ΔS= ∫ _{T₁} ^{T₂} (C _p /T)dT.....	37
Equation 32:	ΔG° = ΔH° - TΔS°	37
Equation 33:	K _{folding} = [folded] / [unfolded] = 1.....	54

Abstract

Bacteriorhodopsin (bR) is a highly expressed transmembrane protein that acts as a light-driven proton pump converting light energy into a proton gradient. The extraordinary levels of expression achieved (15-30 mg per liter of culture) are a result of very efficient biogenesis that originates from molecular information encoded in the *bacterio-opsin* gene (*bop*) (1). DNA sequence analysis and predictive folding algorithms suggest that the first twenty-five bases of the *bop* gene mRNA can form a secondary structural element (a “stem-loop”). Using biophysical methods, the goal was to determine if the stem-loop structure exists in solution conditions that mimic the *in vivo* biological environment (42°C and 4M KCl) (2-4). The salt-dependent thermal stability of a DNA model of the stem-loop sequence was obtained using temperature-controlled UV-absorption and Circular Dichroism (CD) spectroscopy and Differential Scanning Calorimetry (DSC). The combinatorial analysis indicates that an energetically favorable stem-loop structure forms in all solution conditions examined. The stem-loop structure is stabilized at very high salt concentrations, a requirement for a role in *bop* gene expressivity ($\Delta G = -1.13$ kcal/mol, $\Delta H = -16.78$ kcal/mol, $\Delta S = -49.67$ cal/mol/K). While the T_M values evaluated by all three physical techniques were comparable, comparison of the calorimetric and van't Hoff enthalpies indicate that the folding mechanism is not a two-state process. SVD analysis of the UV and CD spectral transitions confirm that formation of the stem-loop structure proceeds through a single significant folding intermediate. Similar analysis of the RNA sequence demonstrates that an RNA stem-loop structure can form and is significantly more stable than that of the DNA structure at all salt concentrations evaluated.

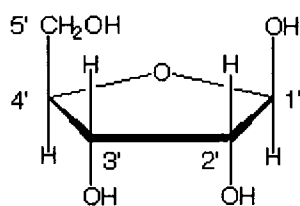
1 Introduction

Living matter is composed of various biological macromolecules. These macromolecules include proteins, lipids, carbohydrates and nucleic acids. The central dogma of molecular biology begins with nucleic acids.

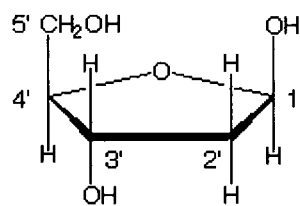
1.1 Nucleic Acids (DNA and RNA)

Nucleic acids are responsible for the genetic information storage and processing in living organisms. Deoxyribonucleic acid (DNA) serves as the cell's hereditary information, while ribonucleic acid (RNA) is involved in converting that information into functional products, such as proteins (5, 6). Nucleic acids are long polymers. The synthesis of such macromolecules is a challenge to the cell and could involve millions of reactions. Instead, the cell uses a modular approach for building these macromolecules by joining prefabricated monomers called nucleotides. These nucleotides can be coupled through a phosphodiester linkage to form a polynucleotide. A nucleotide is composed of three units: a nitrogenous base, a five-carbon sugar, and a phosphate group.

1. Nitrogenous bases: the bases of nucleotides are derivatives of pyrimidine or purine. Naturally occurring pyrimidines are cytosine and thymine in DNA, whereas cytosine and uracil occur in RNA. Naturally occurring purines in both RNA and DNA are adenine and guanine.
2. Five-carbon sugars (pentoses): RNA contains D-ribose, while DNA contains 2-deoxy-D-ribose (Figure 1). The apparent minor difference of a hydroxyl group at the 2' position of the ribose sugar has extensive effects on the secondary structures of RNA and DNA.



Ribose

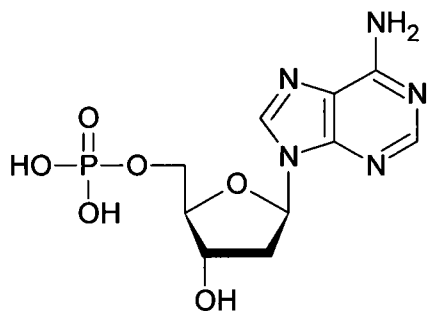


2-Deoxyribose

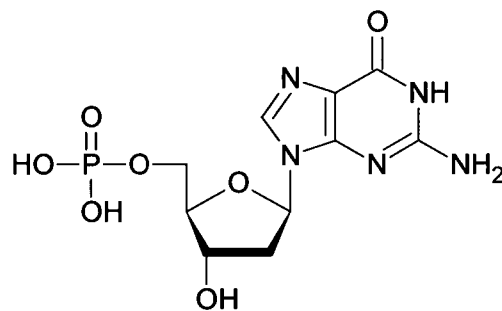
Figure 1: Five-carbon sugars in DNA (2-deoxy-D-ribose) and RNA (D-ribose)

3. Phosphate group: The third component of a nucleotide is a phosphate group that results when phosphoric acid is esterified to the sugar 5' hydroxyl group to yield a phosphate ester.

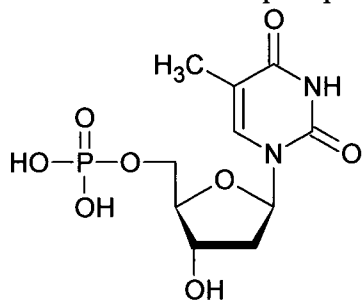
A nucleotide is created upon the formation of a glycosidic bond between the nitrogenous base and the sugar connecting the sugar C-1' to N-1 of a pyrimidine or to N-9 of a purine as well as a phosphodiester bond between the phosphate group and the sugar hydroxyl group at C-5'. Nucleotides can exist in the mono-, di-, or tri-phosphorylated forms. The structures of the common nucleotides in RNA and DNA are presented in (Figure 2).



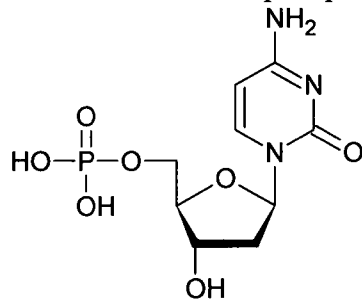
Adenosine 5' monophosphate



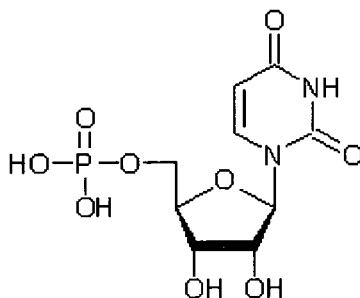
Guanosine 5' monophosphate



Thymidine 5' monophosphate



Cytidine 5' monophosphate



Uridine 5' monophosphate

Figure 2: The structure of common nucleotides in RNA and DNA

1.1.1 Primary Structure of Nucleic Acids

A single strand of polynucleotide chain is formed through the phosphodiester linkages between the 3' carbon of one monomer to the 5' carbon of the next to form the sugar-phosphate backbone (Figure 3) (7). The only significant variation in the chemical structure of the polynucleotide arises from the base sequence. These bases extend from the backbone and give the polynucleotide its unique identity. A common method to describe primary sequence is to write the bases in the 5' to 3' direction. To distinguish between RNA and DNA sequences, DNA sequences are typically preceded by the letter "d" to denote deoxy. The primary structure of DNA and RNA is the sequence itself.

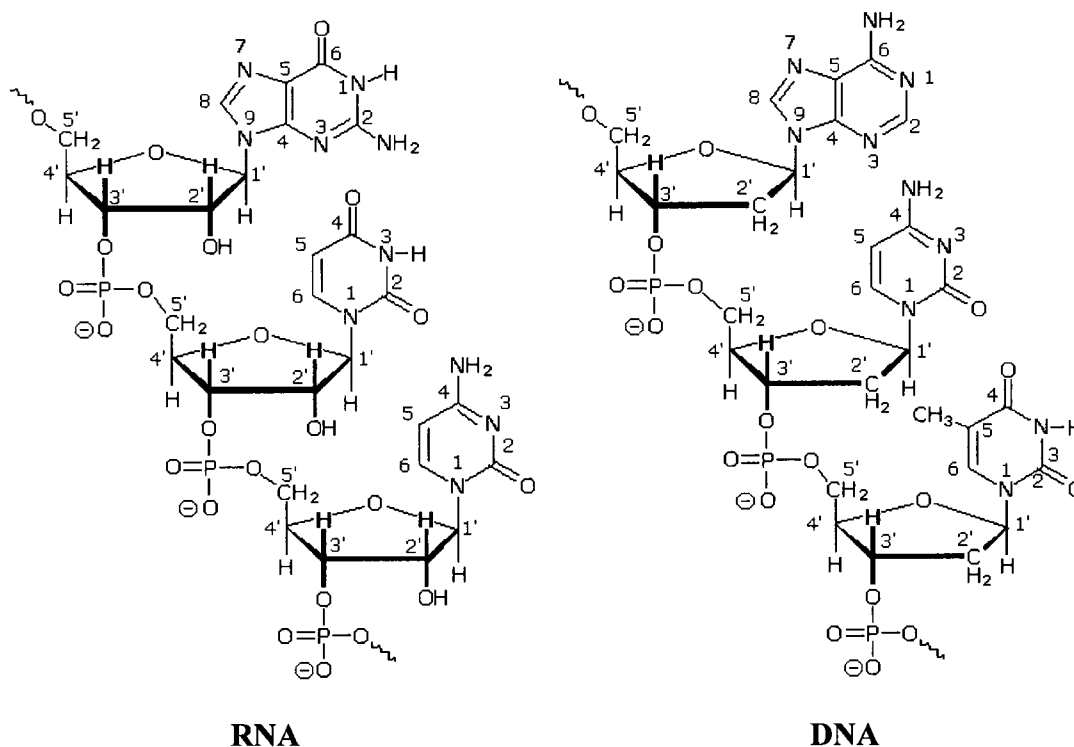


Figure 3: Nucleic acids primary structure (7)

1.1.2 Secondary Structure of Nucleic Acids

A higher order of structure is produced upon the formation of hydrogen bonds between two complementary base pairs. The most common base pairing is Watson-Crick base pairs between purines and pyrimidines, where the adenine base forms hydrogen bonds with the thymine base and the guanine base with the cytosine base. Actually, DNA can adopt various forms of secondary structures (52) such as B-DNA, Z-DNA and H-DNA. The most common form is the canonical B-form DNA structure described by Watson and Crick (53). Figure 4 (7) represent a double helix formed by two anti-parallel DNA strands linked by hydrogen bonding between bases on opposite strands. In RNA, the secondary structures are intramolecular loops (Figure 5) formed when a given strand folds back upon itself by complementary Watson Crick base pairing. This creates a duplex formation with strands that also run antiparallel to one another. RNA also utilizes

uracil in place of thymine for base pairing. The most common forms of RNA secondary structure are the helices and the stem-loop structures (9). Other forms of RNA secondary structure include internal loops, bulges and junctions (10).

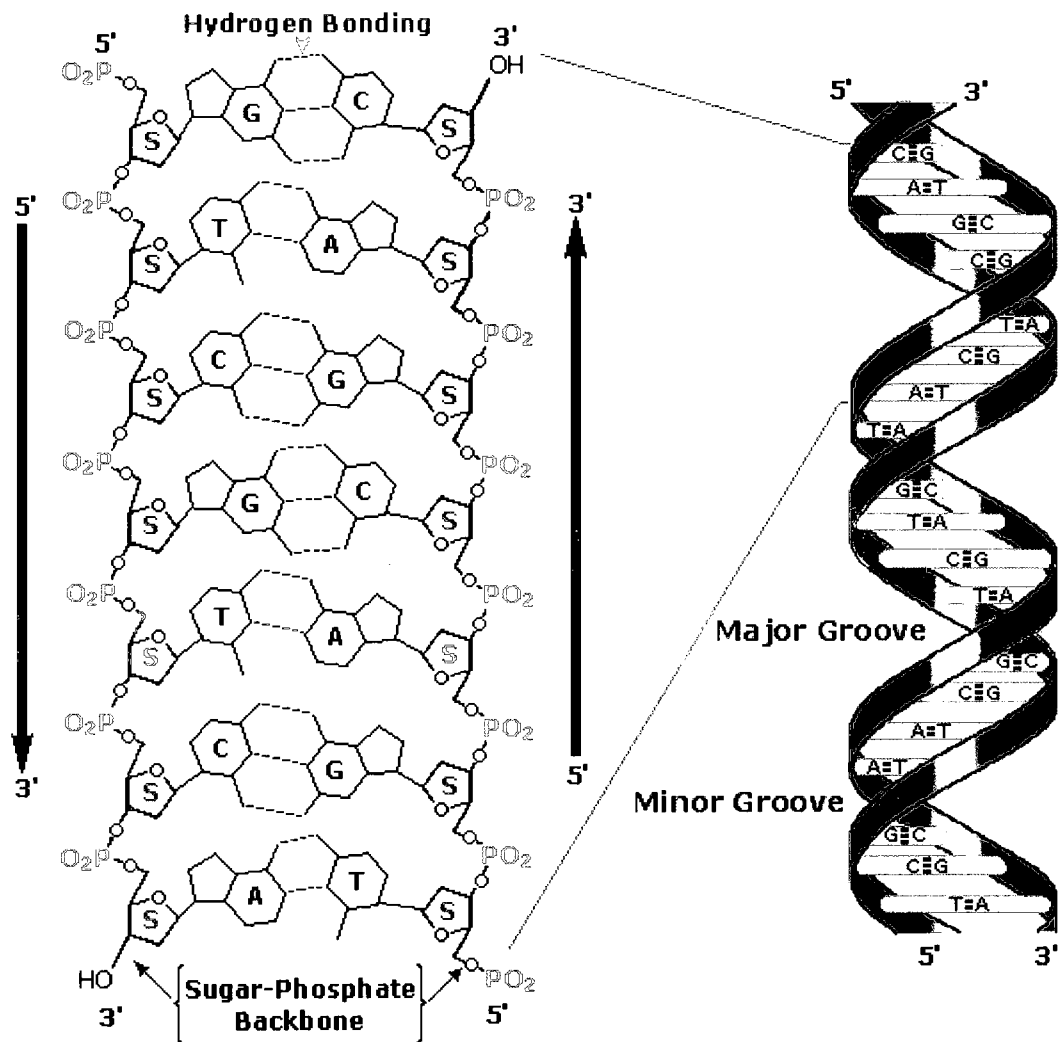


Figure 4: DNA secondary structure (7)

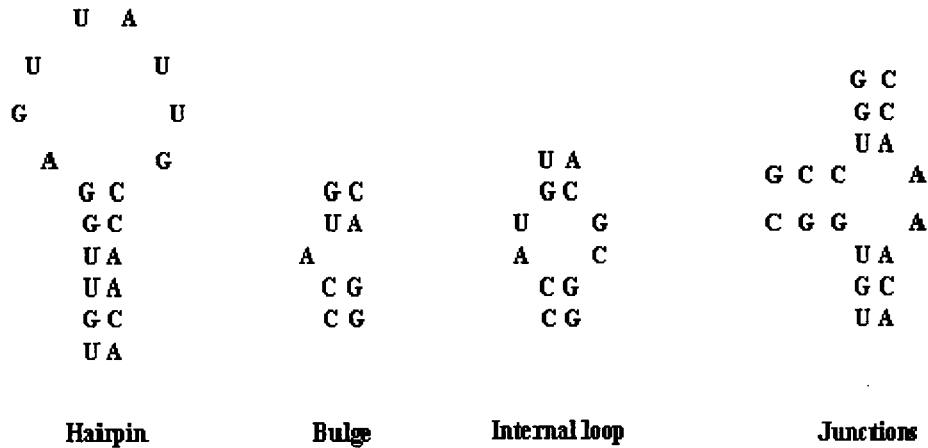


Figure 5: RNA secondary structures

1.1.3 Tertiary Structure of Nucleic Acids

There is an additional level of folding and supercoiling, where the nucleic acid strands folds into larger structures made up of a three-dimensional scaffold called the tertiary structure.

1.1.4 Factors contributing to the stability of nucleic acids

Nucleic acid structures are stabilized by non-covalent interactions between the bases via hydrogen bond formation. The phosphate groups are arranged on the exterior of the molecule facing the water environment. These phosphate groups can interact with ions in solution to form a gradient of condensed counterions. The hydrophobic and van der Waals interactions as well as base pair stacking interactions contribute significantly to the stability of nucleic acids as well.

1.2 The Central Dogma of Molecular Biology

DNA contains the complete genetic information that defines the structure and function of an organism. Proteins are formed using the genetic code of the DNA. The conservation of the genetic information is hugely complex. The flow of information is represented by 3 major stages: replication of DNA, transcription of DNA to RNA and translation to protein. This dogma forms the backbone of molecular biology. (Figure 6)

1. Replication: DNA replicates the coded genetic information. Each strand works as a template on which a new complementary strand is assembled producing two identical DNA molecules.
2. Transcription: The transcription process is initiated upon the binding of a promoter protein to the gene initiating sequence. The DNA double strand unwinds and one of the strands serves as a template for RNA synthesis. The DNA is transcribed into an alternate molecular nucleic acid form, known as messenger RNA (mRNA) since it serves as the information messenger from the DNA to the protein synthesis machinery, known as ribosomes. In eukaryotic cells, the DNA segments are transcribed into mRNA. The mRNA contains protein coding regions called exons, and the noncoding segments called introns. The mRNA is processed by splicing and ultimately migrates from the nucleus to the cytoplasm.
3. Translation: The messenger RNA carries coded information to the ribosome. The ribosome translates the RNA sequence into a primary sequence of amino acids as the protein is synthesized.

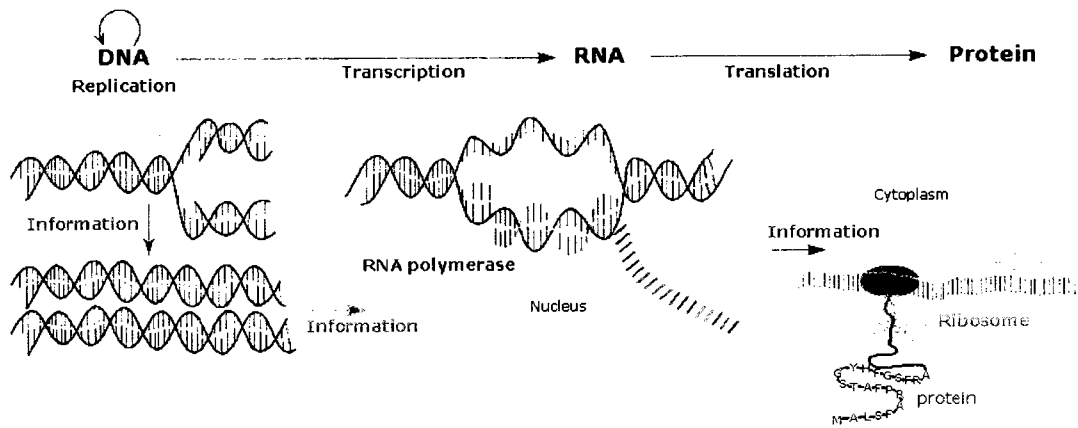


Figure 6: The Central Dogma of Molecular Biology (7)

1.3 *Halobacterium salinarum*

The replication, transcription and translation activities are conserved in all biological forms of life (Figure 7). The organism we study is *Halobacterium salinarum*. As a member of the archaea domain (11), *H. salinarum* shares many traits with both prokaryotes and eukaryotes allowing the use of this organism as a model to study protein expression.

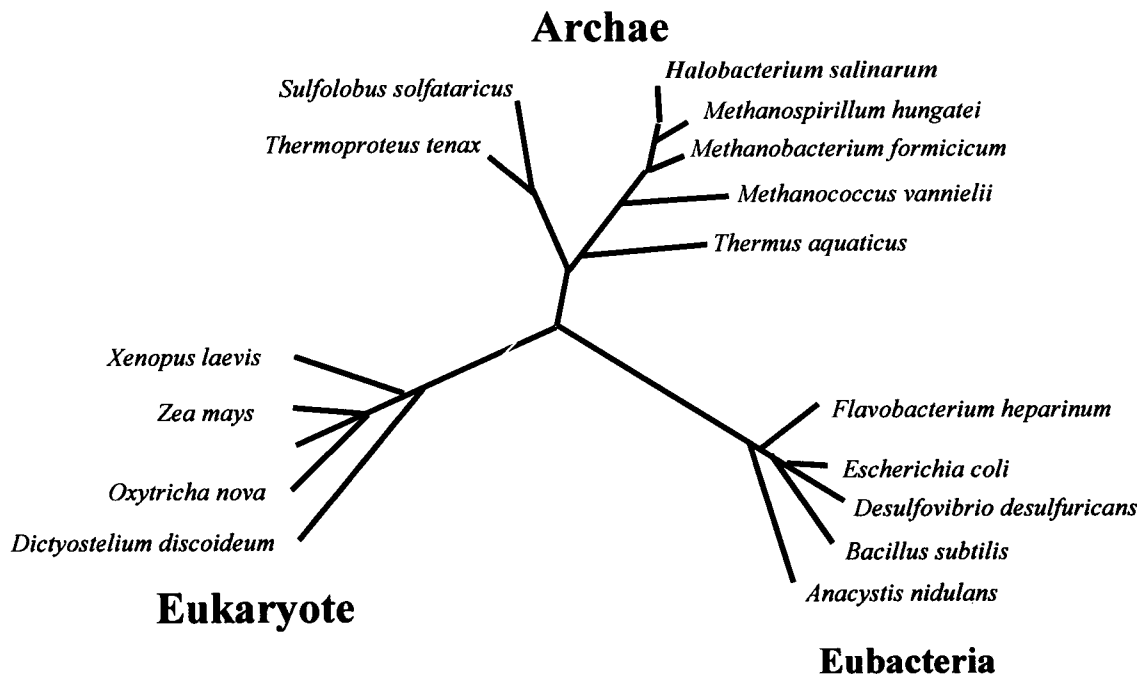


Figure 7: The three domains of life forms

Halobacterium Salinarum is an Archaeon that thrives in high salt media. It grows optimally at a NaCl concentration of 4.3M and at an internal salt concentration $\approx 5M K^+$ (12, 61). It carries out oxidative metabolism if oxygen and metabolic energy sources are available. However, in oxygen and nutrient-deficient conditions, purple patches appear on the cell membrane (58). These purple patches allow *H. Salinarum* to survive by capturing light energy “light-driven proton transport”. The main component of these purple patches is a photosynthetic light receptor known as Bacteriorhodopsin (bR). *H. salinarum* expresses bR at extraordinary levels. bR can occupy up to 80 percent of *H. salinarum* cell membrane and can be purified from cell culture yielding 15-30 milligrams per liter (13, 15).

1.4 Bacteriorhodopsin (bR)

Bacteriorhodopsin (bR) is a 26 kDa transmembrane protein that acts as a light-driven proton pump, converting light energy into a proton gradient. The proton gradient is used to support energy requiring metabolism essential for the viability of the host organism, *H. salinarum*. This is done by generating ATP via F1F0 ATPase. bR consists of a single polypeptide (bacterio-opsin, Bop) of 248 amino acids arranged in seven transbilayer alpha-helices (heptahelical transmembrane 7-TM). The helices enclose a retinal chromophore that is covalently bound by a protonated Schiff-base linkage to residue Lys216 in the seventh transmembrane helix of bacterio-opsin (14). bR is a 1:1 assembly of the chromophore retinal and the 248 amino acid bacterio-opsin protein. The high level of bR expression, along with the chromophoric properties of retinal, produce *H. salinarum* cultures with an intense purple color leading to an efficient visual assay of expression (15) (Figure 8). When bR is highly expressed, the heptahelical

transmembrane aggregates forming a homotrimer. Homotrimers aggregate and expand across the membrane to form the purple membrane (16) (Figure 9).

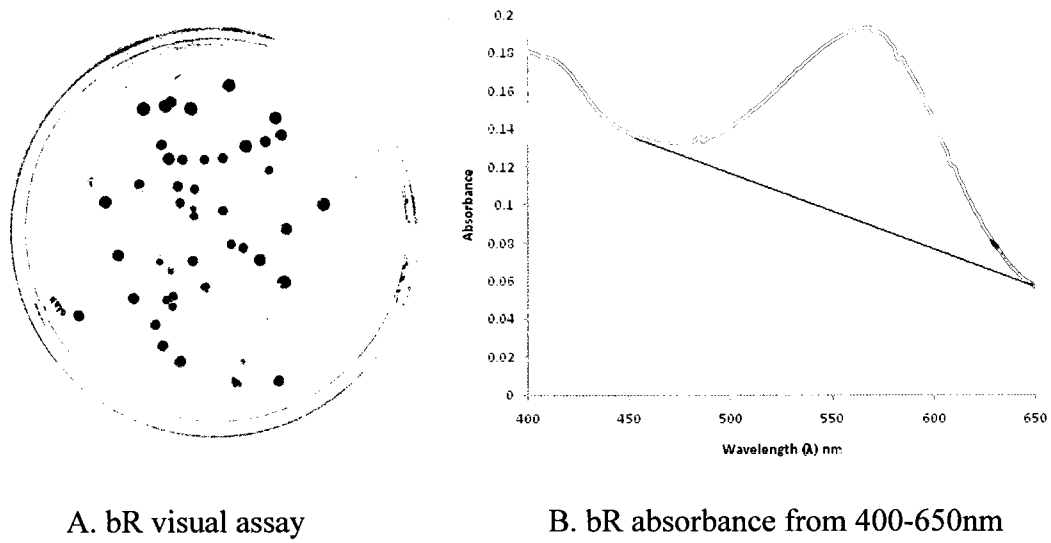


Figure 8: Bacteriorhodopsin (bR) expression

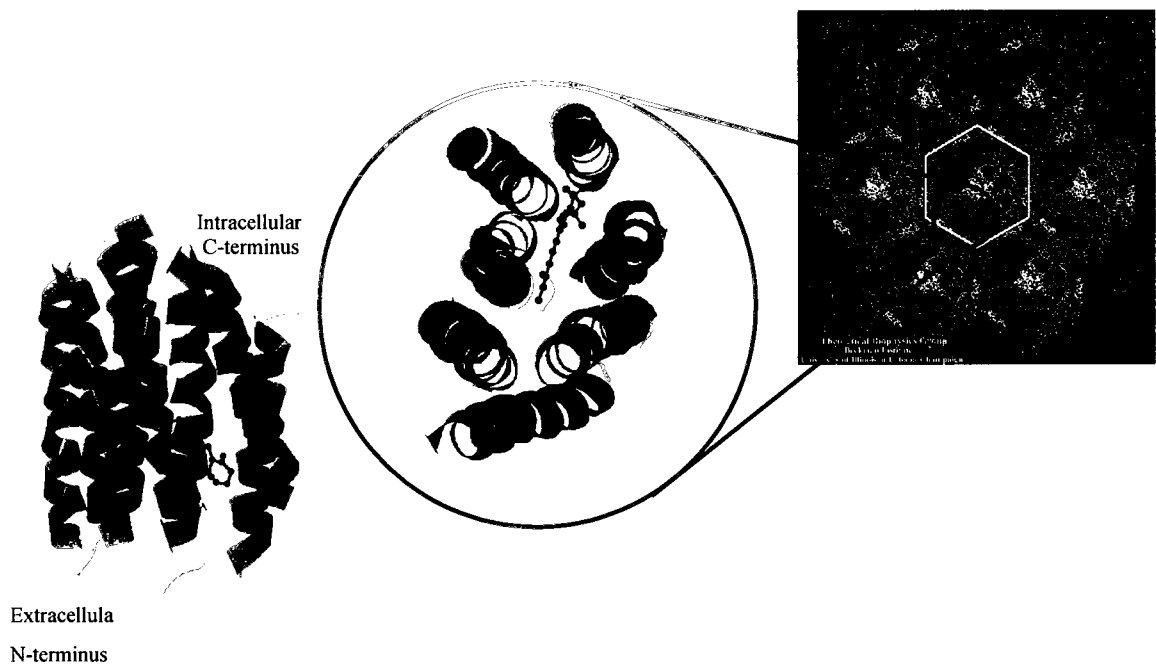


Figure 9: Homotrimers aggregate and expand across the membrane forming the purple membrane

1.5 The bacterio-opsin gene (*bop*)

This extraordinary level of membrane protein production is unique and is likely due to an efficient *bacterio-opsin* gene (*bop*). The *bop* gene has been cloned, sequenced and most of the *cis*- and *trans*-acting elements have been identified (1, 17, 18).

1.5.1 Nucleic acid sequences can contribute to gene expression

Nucleic acid sequences can contribute to gene expression in two ways:

1. *Trans*-acting elements:

These determinants contribute to *bop* gene expression but are not part of the *bop* gene nucleic acid sequence. Intense light and/or a decrease in oxygen have a direct effect on the amount of *bop* messenger RNA (mRNA) that is produced. As mRNA levels are raised, the levels of bR are increased. Light and oxygen are *trans*-acting determinants of expression.

2. *Cis*-acting elements:

These are factors that are wholly contained within the *bop* gene nucleic acid sequence. The *bop* gene mRNA starts just two nucleotides upstream of the initiating methionine codon, the first codon translated in the *bop* gene coding region (59). This essentially “leaderless” architecture is rare, but has been observed in all three domains of life. As a consequence of being leaderless, all *cis* acting determinants in the *bop* mRNA are located downstream of the initiating codon. All the information for efficient transcription and translation exist in the *bop* gene nucleotide sequence.

1.5.2 Secondary structural element (Stem-loop)

The extraordinary levels of expression are a result of very efficient biogenesis that originates from optimized molecular information encoded in the *bacterio-opsin* gene (*bop*). DNA sequence analysis (1) and predictive folding algorithms suggest that the first twenty-five bases of the *bop* gene mRNA can form a secondary structural element (a “stem-loop”) that contribute to bR biosynthesis.

1.6 Research Purpose

The goal was to determine if the putative stem-loop structure exists in solution conditions that mimic the *in vivo* biological environment (42°C and 4M KCl). Furthermore, the determination of the thermodynamic parameters of the stem-loop structure will yield insight into a folding pathway. These folding pathways can be instrumental to the regulation of gene expression.

2 Theory

2.1 Spectroscopy

2.1.1 Light

Light is electromagnetic radiation that may be considered in terms of two perpendicular vectors: electric and magnetic. These vectors oscillate as waves characterized by wavelength λ and frequency ν (Figure 10). The wavelength λ is defined as the distance between two neighboring peaks, and may be designated in meters, centimeters or nanometers. While, frequency ν is the number of wave cycles that travel through a given point per second. It is designated in cycles per second, or hertz (Hz). These terms can also be described as follow (38, 39):

Equation 1:
$$c = \lambda\nu$$

Where c is the velocity of light in a vacuum

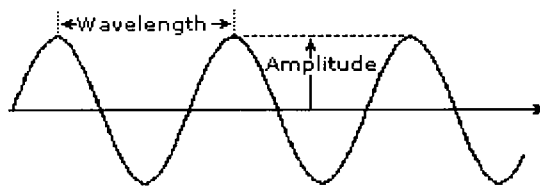


Figure 10: The wavelength of electromagnetic radiation

2.1.2 Electromagnetic radiation spectrum

The electromagnetic radiation spectrum covers a broad range of wavelengths ranging from very short wavelength such as gamma and x-rays to very long ones such as microwaves and broadcast radio waves (Figure 11). The visible wavelengths cover a

range from approximately 380 to 780 nm (40). The electromagnetic radiation is composed of photons. The energy of each photon is dependent on its (λ) and (ν) and can be described as follow:

Equation 2: Photon energy $e = h\nu = hc/\lambda$

where (h) is Plank's constant.

Thus, the shorter the wavelength, the higher the energy and vice versa.

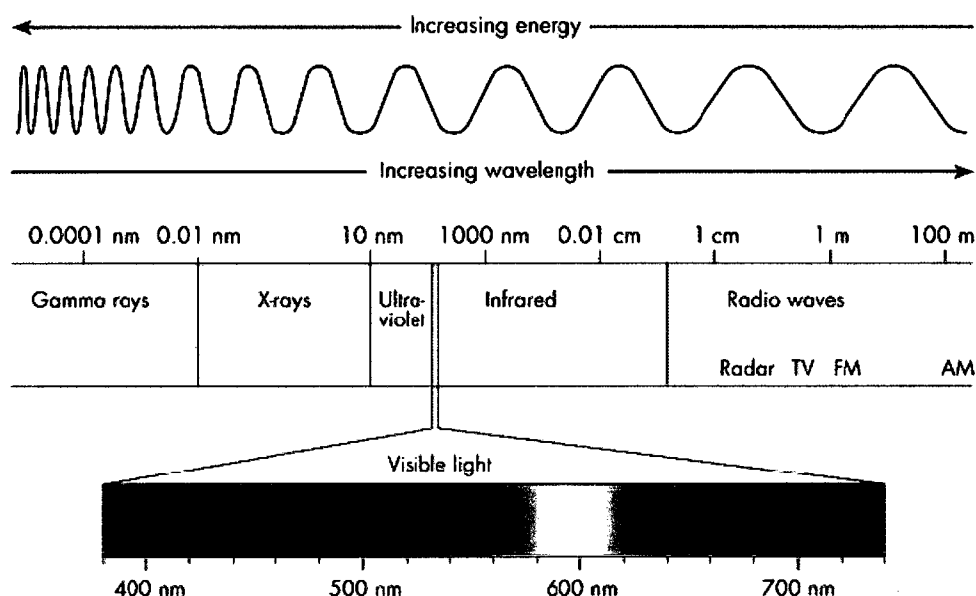


Figure 11: The electromagnetic spectrum

2.1.3 Ultraviolet Absorption

Different molecules absorb radiation of different wavelengths. An absorption spectrum will show a number of absorption bands corresponding to the excitation of the outer electrons (38-40). When a molecule absorbs energy, electrons are promoted from their ground state to an excited state. Atoms and molecules possess defined energy levels. The electronic energy levels are widely separated and usually only the absorption of high energy photons can excite a molecule from one level to another. The vibrational and

rotational energy levels are much closer to each other than the electronic energy levels. Therefore photons of lower energy are sufficient to bring about vibrational and/or rotational changes (Figure 12). Upon the absorption of UV light, the transition can be from any vibrational or rotational energy level to any vibrational or rotational energy level in any higher electronic energy level (20).

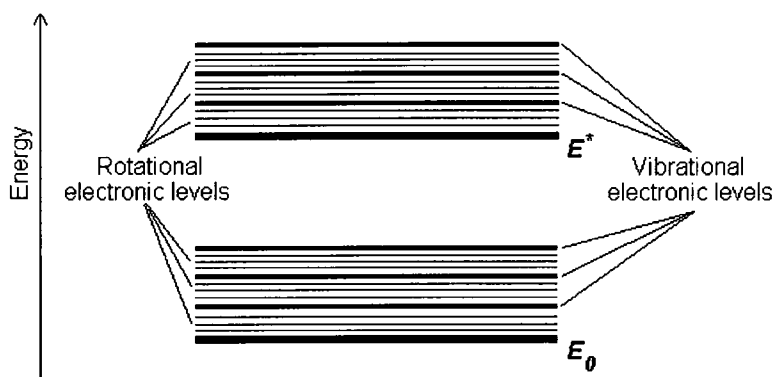


Figure 12: Energy levels of a molecule

A molecule will exhibit absorption in the visible or ultraviolet region when radiation causes an electronic transition within its structure. The energy supplied by the light will promote electrons from their ground state orbital to a higher energy, excited state orbital or antibonding orbital. Potentially, three types of ground state orbitals and two types of antibonding orbitals may be involved in the transition (Figure 13).

$\sigma \longrightarrow \sigma^*$ Transitions:

An electron in a bonding σ orbital is excited to the corresponding antibonding orbital σ^* which requires large energy. This type of transition is not typical in the UV-Vis spectra.

$n \longrightarrow \sigma^*$ Transitions:

This type of transition needs less energy than the previously described transition state and can be initiated with light in the range of 150-250 nm region.

$n \longrightarrow \pi^*$ and $\pi \longrightarrow \pi^*$ Transitions:

Most of the absorption spectroscopy of organic compounds is based on these two types of transitions as the absorption peaks for them fall in the UV-Vis region of the spectra (200-700 nm).

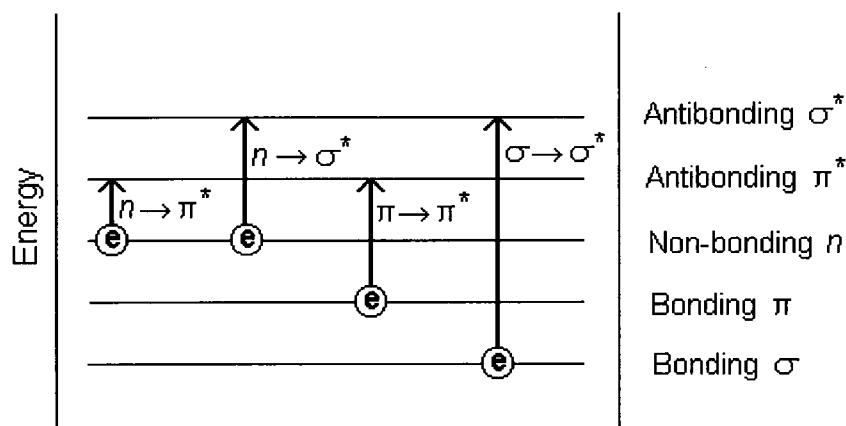


Figure 13: Possible electronic transitions of σ , π and n electrons

2.1.3.1 The Beer-Lambert Law

The Beer-Lambert Law is the linear relationship between the absorbance (A) and the concentration of a chromophore. The law is only true for monochromatic light, which is light of a single wavelength. It is necessary that the physical or chemical state of the substance does not change with concentration (8, 41).

Experimental measurements are usually made in terms of transmittance (T), which can be defined as:

Equation 3:
$$T = I / I_0$$

Where I_0 is the intensity of the incident light and I is the intensity of the transmitted light after passing through the sample. Transmittance (T) is usually expressed as a percentage of transmittance (%T).

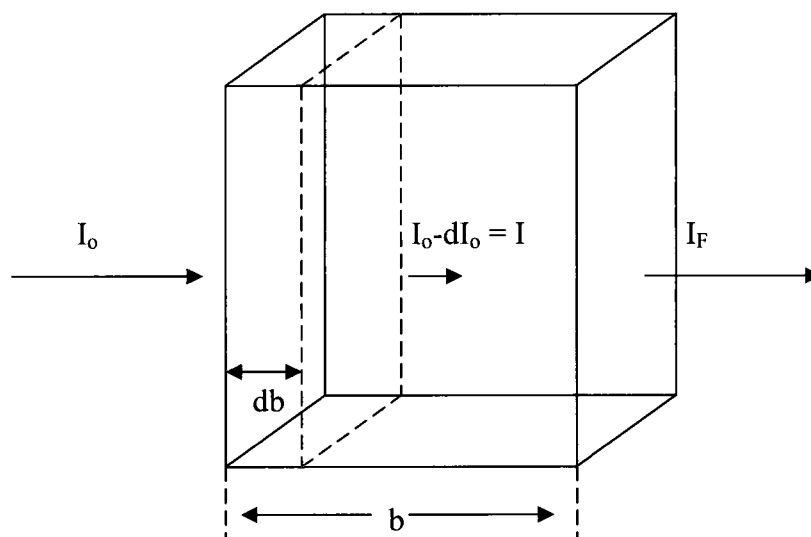


Figure 14: Light transmission through a sample (8)

The Beer-Lambert Law can be derived from Skoog and colleagues discussion based on an early paper (Strong, F.C. *Anal. Chem.*, **1952**, *24*, 338) (8). Consider a beam of monochromatic light passing through a sample which absorbs light. As indicated in Figure 14, I_0 is the light intensity prior to entry into the sample cell and I_F is the light intensity after passing through the sample. I is the light intensity at a given point in the cell, where the light has crossed a distance db through the sample. Thus, the original intensity I_0 is diminished by dI_0 . The fractional change in the intensity can be written as dI_0/I_0 . Additionally, the area outlined in red denotes the cross sectional area (A) of the vessel. When light is absorbed, dA is the area of detained beam. The ratio of dA/A signifies the amount of area that can detain the beam as it relates to the total area. The fractional change in intensity can be related to the fractional area of the detained beam in the following equation:

Equation 4:
$$- dI_0 / I_0 = dA/A$$

We can substitute, $dA=adn$, where the proportionality constant is a and dn is the number of absorbing particle. Integration produces:

Equation 5:
$$-\int_{I_0}^{I_F} \frac{dI}{I} = \frac{a}{A} \int_0^b dn$$

Equation 6:
$$\ln I_0 / I_F = (a/A)b$$

Convert to base 10 log units to produce:

Equation 7:
$$\log I_F / I_0 = (a/(2.303)A)n$$

Knowing that A is equivalent to volume (V) divided by a length (b), we can substitute for A as follows:

Equation 8:
$$\log I_F / I_0 = ((ab)/(2.303 V))n$$

We can substitute n with $6.02 \times 10^{23}c/1000$, where c is the molar concentration:

Equation 9:
$$\log I_F / I_0 = ((ab)/(2.303 V))(6.02 \times 10^{23}c/1000)$$

Finally, the terms $((a)/(2.303 V))(6.02 \times 10^{23}/1000)$ can be combined to become the molar extinction coefficient ϵ and related to the absorption (A) of a chromophore:

Equation 10:
$$A = \log I_F / I_0 = \epsilon bc$$

Equation 11: Beer's law
$$A = \epsilon bc$$

where ϵ is a wavelength-dependent constant called molar absorptivity coefficient with units of $M^{-1} \text{ cm}^{-1}$

The linearity of the Beer's law is limited by chemical and instrumental factors such as: light scattering due to the presence of particulates in the sample, fluorescence or phosphorescence from the sample and stray light.

2.1.3.2 UV-Vis Spectrometer

A diagram of the components of a Cary 3E spectrophotometer with a temperature controlled sample compartment similar to the one used conducting this research is presented in (Figure 15) (8). For visible or ultraviolet light, either a tungsten or deuterium lamp is used as a light source, respectively. The light beam is directed into a beam splitter using a series of mirrors. Upon the splitting of the light beam, it is directed through the sample and reference cells using another set of mirrors. The cells are housed in a multicell temperature controlled block in the sample compartment which is purged with nitrogen to prevent condensation on the sample cells. The temperature can be monitored using temperature probes within the block or in the sample cells. The block temperature is controlled by a water bath that is filled with 10% ethylene glycol. After passing through the sample and reference cells the light beams impinge on two mirrors. The mirrors direct the light beam into a double chopper, then to the detector. The instrument is interfaced with a PC, which contains the Cary WinUV Software Version 2.0.

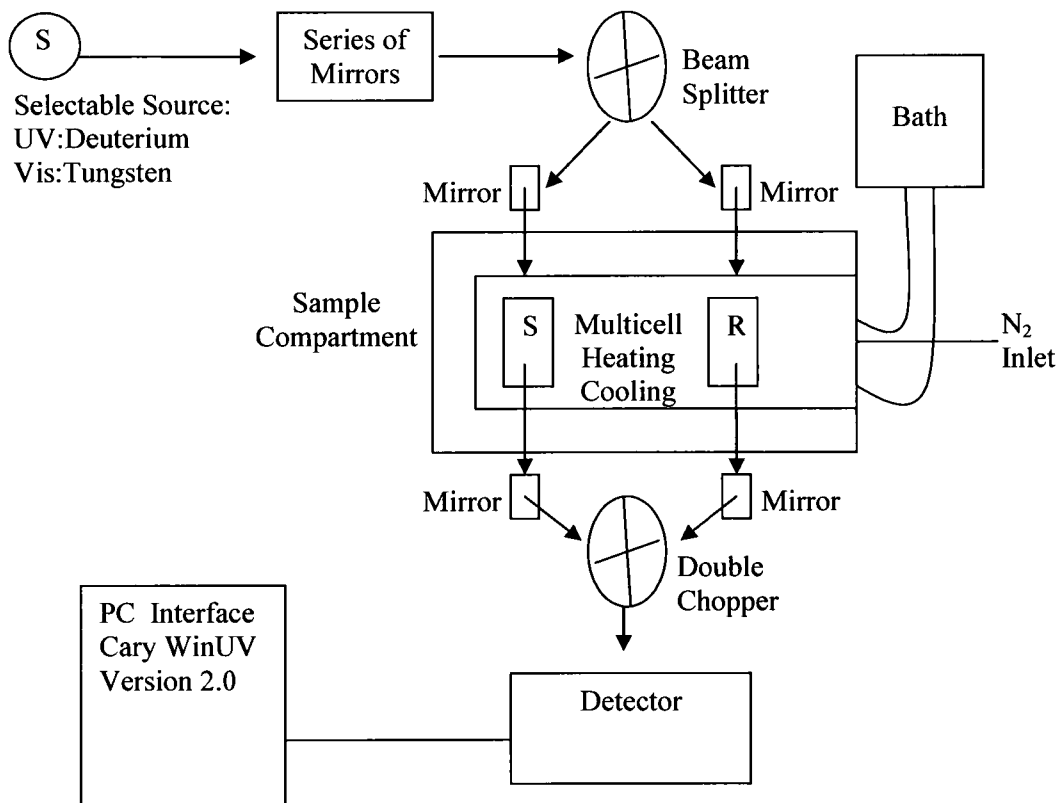


Figure 15: Schematic of Cary 3E spectrophotometer

2.1.4 Hyperchromicity

Hyperchromicity is the increase in the UV absorbance of a molecule as it undergoes conformational changes. The thermal melt of nucleic acid is a clear example of hyperchromicity as the UV absorption increases when the nucleic acid changes from the folded to unfolded state. The apparent increase in absorbance is attributed to the unstacking of bases and loss of Davydov splitting. The π -electrons of the bases stack to minimize the π -electron exposure to hydrophilic solvents. Since UV absorbance is a consequence of π electron transitions, stacked nitrogenous bases interact with each other

via π electrons to produce a decrease in base absorption. Unstacking releases the interactions and the UV absorbance increases.

2.1.5 Melting temperature T_M

The melting temperature, T_M , is defined as that temperature that 50% of the DNA is double stranded and 50% is single stranded. T_M reflects the intrinsic stability of nucleic acids. High GC content is associated with higher T_M values. T_M also reflects the extrinsic contributions from the solution, such as ionic strength. Ions suppress the electrostatic repulsion between the electronegative phosphate groups (54, 55).

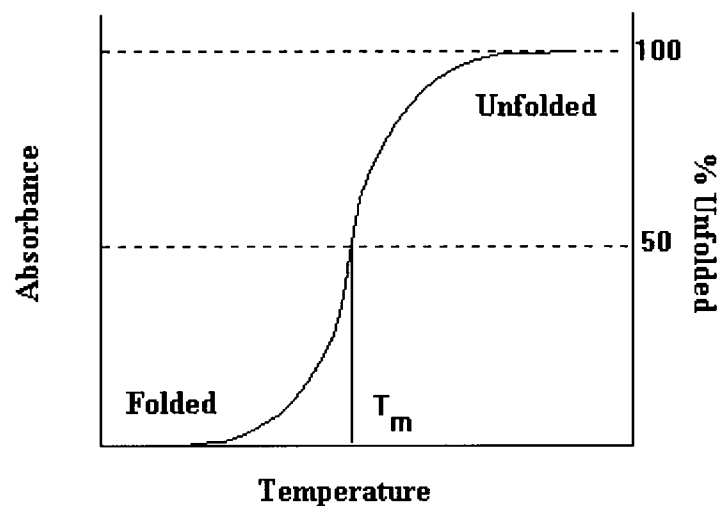


Figure 16: Melting curve

2.1.6 Circular Dichroism (CD) Spectroscopy

2.1.6.1 Polarization of Light

The polarization of light is a property that describes the orientation of the oscillation of the light's electric field (Figure 17). If the orientations of the electric fields vector oscillation are not correlated, the light is said to be unpolarized. Consequently, if there is correlation, the light is said to be polarized (8, 22) (Figure 18).

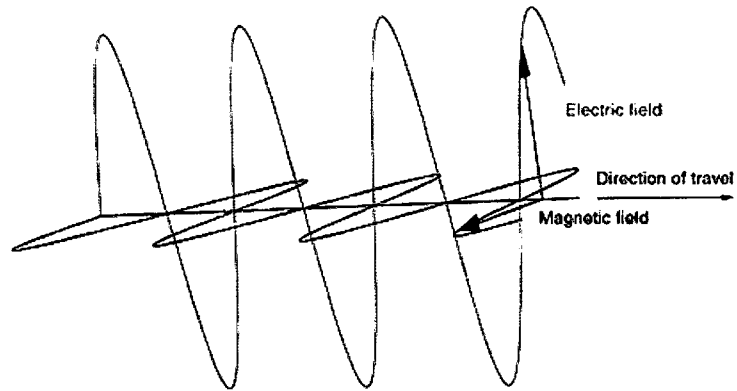


Figure 17: Light electric and magnetic fields

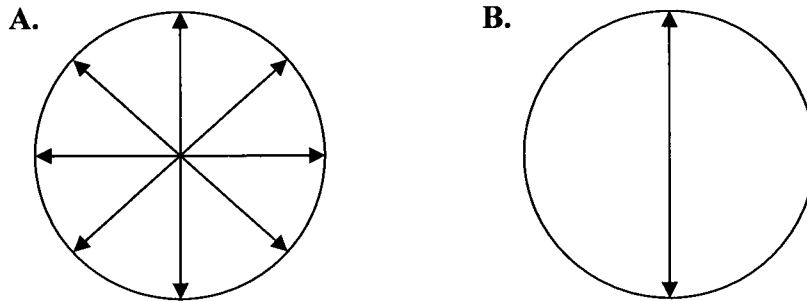


Figure 18: Light electric field vectors (A) unpolarized and (B) polarized

When the polarization is perpendicular to the wave's direction of travel, the electric field may be oriented in a single direction called linear polarization (42, 43). When the polarization rotates as the wave travels, it is called circular polarization. Circularly polarized light oscillations can rotate in a clockwise fashion to be right circularly polarized light or in a counterclockwise fashion, where it is considered left circularly polarized (Figure 19).

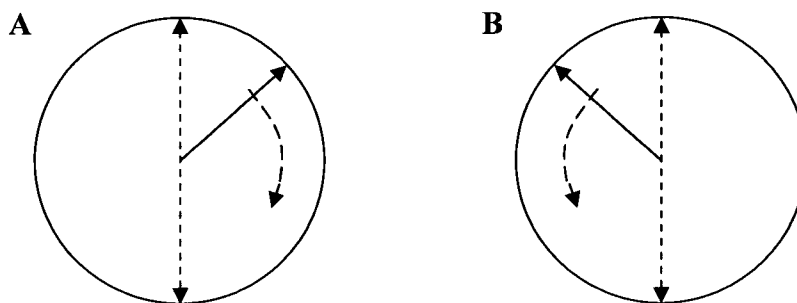


Figure 19: Plane polarized relative to (A) right and (B) left circularly polarized light

When the polarization vector rotates as the wave travels and the vectors are not of equal magnitude, the vectors can sum to form elliptically polarized light. Here the electric field vector sweeps out an elliptical helix (23, 24).

2.1.6.2 Circular Dichroism (CD)

Circular Dichroism is a form of spectroscopy based on the differential absorption of left- and right-handed circularly polarized light (8, 22, 23, 24) and can be written as

Equation 12:
$$\Delta A = A_L - A_R$$

Where A is the absorbance and L or R indicate the circularly polarized light components. If linearly polarized light passes through an optically active substance (i.e. an asymmetric chromophore) the sample would absorb and/or delay the two circularly polarized components at different rates. As a result of the absorption phenomenon, elliptically polarized lights emerge from the sample. Ellipticity (θ) can be determined from the difference of the absorption coefficients ϵ_L and ϵ_R and measured as the arc tan of the minor ellipse axis/major ellipse axis (25)

Equation 13:
$$\theta_\lambda = \text{arc tan} ((\text{minor ellipse axis})/(\text{major ellipse axis}))$$

The molar ellipticity is then given as:

Equation 14:
$$\theta_{\lambda} = 2.303((\epsilon_L - \epsilon_R) / 180) / 4\pi \text{ deg}$$

If the molar extinction coefficient for the left and right circularly polarized light are known, simplification leads to:

Equation 15:
$$\theta_{\lambda} = 3300 (\Delta\epsilon) \text{ mdeg M}^{-1}\text{cm}^{-1}.$$

The CD spectrum is defined by the wavelength dependent ellipticity.

2.1.6.3 CD Instrumentation

A diagram of the components of a CD Spectropolarimeter 62A DS from Aviv Associates with a temperature controlled sample compartment similar to the one used conducting this research is presented in (Figure 20) (8). A high voltage xenon lamp is used as the light source. The lamp produces a light beam that is directed into a polarizer. The polarizer produces left and right circularly polarized light which is directed through the sample cells. The cells are housed in a multicell temperature controlled sample compartment. The sample compartment and the instrument optics are under constant nitrogen purge to prevent ozone formation which can damage the instrument optics and reduce the efficiency of the lamp.

After passing through the sample cell the elliptically polarized transmitted beam is directed into the detector. The elliptically polarized beam has an optical rotation angle Φ and ellipticity θ that can be determined by the CD. The instrument is interfaced with a PC and a Peltier temperature controller to determine spectra as a function of temperature as well as a function of wavelength.

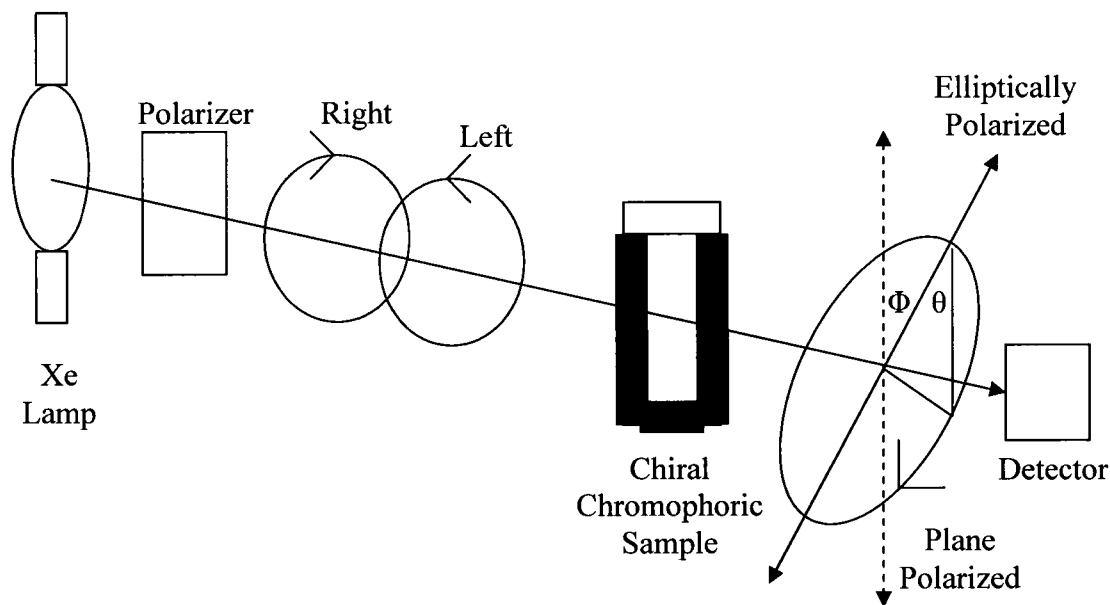


Figure 20: Diagram of CD spectropolarimeter

2.1.6.4 Stem-loop CD Spectrum

As stated in section 1.1, nucleic acids are polymers that are composed of nucleotides. Each nucleotide is composed of three units: a nitrogenous base, a five-carbon sugar, and a phosphate group. The phosphate groups have electronic transitions at wavelengths shorter than 170 nm. The pentose sugar has a low intensity absorption starting at 190nm. Therefore, the principle visible chromophores in nucleic acids are the nitrogenous bases as they possess electronic $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions (26). Although the bases have a plane of symmetry (optically inactive), the ribose and deoxyribose sugars are asymmetric and can induce CD in the absorption of the chromophoric bases. This explains the apparent low intensity of the induced CD spectra for the chromophoric but symmetric bases. In addition to the spectral differences between the mononucleotides, the bases themselves are directly involved in close interactions with one another. For

example, the CD spectrum of a dinucleotide is different from the sum of the CD spectra of the two mononucleotides (27). Therefore, the CD spectrum of a dinucleotide is the average of the two mononucleotides plus an additional term to account for base-base interactions. Also, as the number of nucleotides increase an additional term is introduced to account for the interactions between the nearest neighbors. Furthermore, the π -electrons of the bases makes them hydrophobic and they tend to stack to minimize the π -electron exposure to solvents. The base stacking can assume a helical structure inducing chirality and preferential handedness. Thus, circular dichroism is sensitive to the conformational state of the nucleic acid oligomer. The circular dichroism signal is the result of the summation of all the transitional dipole contributions from each base (interacting and non-interacting) existing in a conformation (Figure 21).

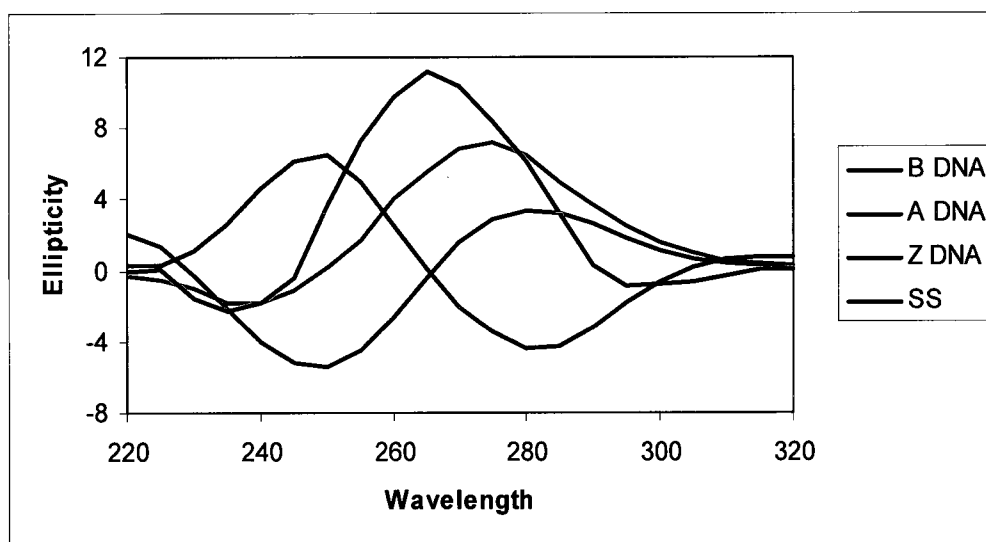


Figure 21: CD spectra of different DNA conformations

2.2 Spectroscopic Data Analysis

2.2.1 van't Hoff Analysis

The secondary structure of nucleic acids can be distinguished by CD (Figure 21) as changes in the CD spectra are indicative of conformational changes. Melting studies can be conducted by CD and the conformational changes can be monitored as a function of temperature. CD is also used to assess the thermal stability of nucleic acids in the presence of other variables such as salt concentration. The conformational changes can be monitored over the entire UV spectrum. Alternatively, a single wavelength can be chosen as a function of temperature to produce equilibrium melting curves (Figure 16).

For a reversible transition with no intermediates, the curves exhibit a sigmoidal shape. This is said to be a two state process, whereas the native folded nucleic acid reversibly converts to an unfolded state with no unfolding intermediates. There is no apparent hysteresis in the reverse melting curve (8). van't Hoff analysis of equilibria melting curves assumes a two state unfolding process and ΔH is temperature independent. A plot of $R \ln K$ versus $1/T$ (in Kelvin) is linear with a slope given by $-\Delta H$ and an intercept of ΔS . This plot is known as a van't Hoff plot (Figure 22) (56, 57).

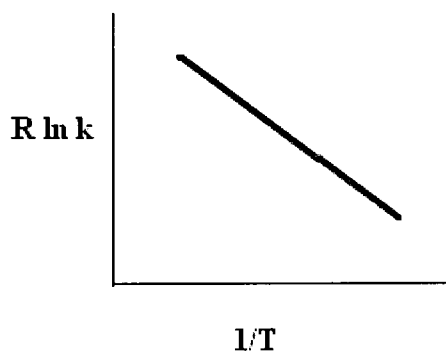


Figure 22: van't Hoff plot

2.2.2 Equilibrium Melting Curve

For optimal and successful data analysis, a proper monitoring wavelength must be selected at which conformational changes can be monitored. It is common to monitor UV melting of nucleic acids at 260 nm. Optimal wavelength can be determined experimentally by collecting the UV and CD data at low and high temperatures. The two spectra are then subtracted to yield a difference spectrum. The wavelength of absolute maximal difference is utilized for the melt (8).

A sigmoidal shaped curve can be drawn using the boundaries as an upper and lower baseline (Figure 16) (62). The temperature at which the distance from the upper baseline to the melting curve equals the distance from the curve to lower baseline is the melting temperature T_M . Alternately, the T_M value can be determined by plotting the first derivative of the spectroscopic absorbance as a function of temperature, the apex of the resulting peak is the T_M (Figure 23).

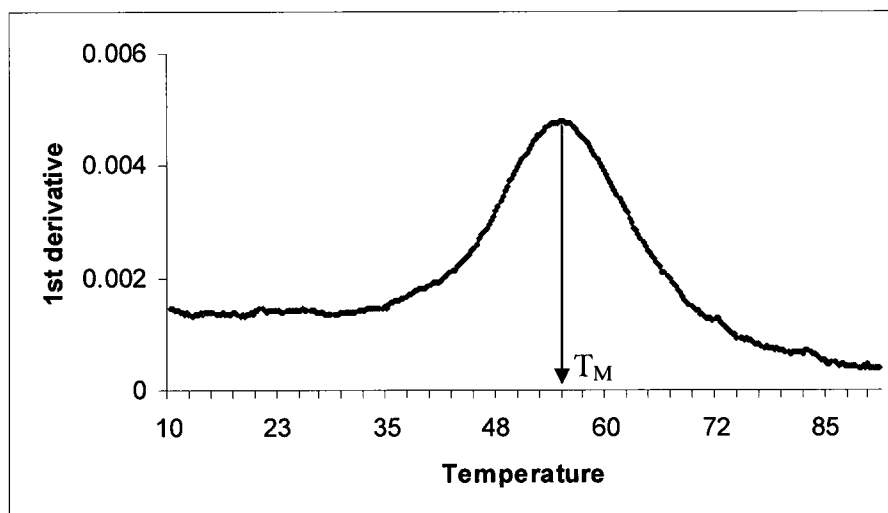


Figure 23: Melting temperature at 100mM salt concentration

Statistical analysis programs can also be used to establish the best sigmoidal fit (Boltzmann fit) of the data where the inflection point is the T_M or the best Gaussian fit of the data first derivative where the apex of the peak is the T_M .

2.2.3 α Plots

The melting curves can be represented in terms of (α), where α is the fraction of folded nucleic acid. The values for α are obtained by subtracting the lower baseline from the upper baseline of a melting curve to yield a value, UL. Then, the distance for the upper baseline to the curve is determined for all points in the melt and symbolized UC. Finally, UC is divided by UL to give α at all temperatures shown in (Figure 24). Note that at T_M , α is 0.5 (8, 28, 29).

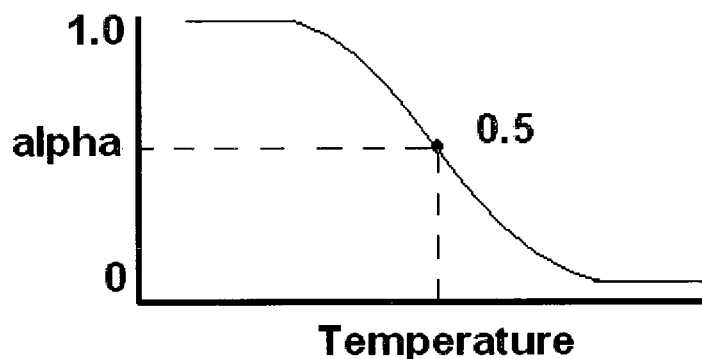


Figure 24: Alpha plot

Unlike typical UV absorption signals, the CD signal decreases in intensity as function of temperature. The UL value as described above is still utilized. However, the distance for the lower baseline to the curve is determined for all points in the melt and symbolized CL. Similarly, CL is divided by UL to give α at all temperature.

The equilibrium constant K can be written as

Equation 16:
$$K = [\text{Folded}] / [\text{Unfolded}]$$

Equation 17:
$$K = \alpha / 1 - \alpha$$

2.2.4 Singular Value Decomposition (SVD)

Singular value decomposition (SVD) (44) is a mathematical technique utilized to decompose a matrix into its component matrices, revealing some of the interesting properties of the original matrix. SVD applies the matrix linear algebra to split a vector space into lower-dimensional subspaces. It is widely used in statistics, where it is related to principal component analysis, and in signal processing and pattern recognition.

The following discussion by K. Baker cited by Antonacci (8) describes SVD mathematical process and rational. A vector must have magnitude and direction. Numerous numbers can be used to describe a vector in the form of a matrix. Matrix (A) is an (m times n) matrix, where m is a row and n is a column of the matrix. Two vectors are considered to be orthogonal if the inner dot product of the two vectors is 0 (90° relative to each other). A vector's length is defined as the square root of the sum of squares of each number in a matrix. Normal vectors have a length of 1. Orthonormal vectors are vectors that are both normal and orthogonal. The process of converting a set of column vector data into orthonormal vectors is known as the Gram Schmidt orthonormalization. The process involves normalizing one column vector followed by normalizing subsequent column vectors taking into account those which have already been normalized. A transpose matrix (A^T) is produced upon converting the rows of the original matrix A into columns of a transpose matrix A^T . An identity matrix contains 1 along its diagonal and zeros elsewhere. Any matrix multiplied by the corresponding identity matrix is equal to the original matrix. Matrix multiplication requires the number

of rows of the first matrix must equal the number of columns in the second matrix. A determinant is calculated for a 2 x 2 matrix, where the product of the two diagonals is subtracted. Expansion by row allows determinant calculation of larger matrices.

Eigenvectors (\mathbf{v}) are nonzero vectors that relate a square matrix \mathbf{A} to some quantity λ (eigenvalues). In other words, \mathbf{A} acts to stretch the vector \mathbf{v} for a distance λ without changing its direction. Determinants are utilized to solve the general formula is $\mathbf{A}\mathbf{v} = \lambda \mathbf{v}$, for their corresponding eigenvalues and eigenvectors (Figure 25).

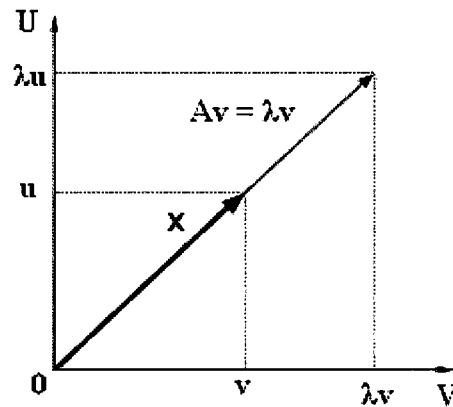


Figure 25: Eigenvector \mathbf{v}

2.2.4.1 Mathematical Process of SVD

SVD pertain to the mathematical decomposition of a matrix into its component matrices applying the matrix linear algebra to split a vector space into lower-dimensional subspaces. The expression is given as follows (8):

Equation 18:
$$\mathbf{A}_{mn} = \mathbf{U}_{mm}\mathbf{S}_{mn}\mathbf{V}_{nn}^T$$

where \mathbf{A} is the matrix undergoing SVD analysis. \mathbf{U} and \mathbf{V} are square orthogonal matrices, where the columns of \mathbf{U} denote the orthonormal eigenvectors of $\mathbf{A}\mathbf{A}^T$ and the columns of \mathbf{V} denote the orthonormal eigenvectors of $\mathbf{A}^T\mathbf{A}$. \mathbf{S} is the diagonal matrix containing the

square roots of the eigenvalues for \mathbf{U} and \mathbf{V} in numerical order. SVD analysis occurs in the following way:

Step 1. \mathbf{AA}^T to \mathbf{U}

Convert the original \mathbf{A} matrix to its transpose \mathbf{A}^T .

Multiply \mathbf{A} and \mathbf{A}^T to obtain \mathbf{AA}^T .

Determine eigenvalues and eigenvectors for \mathbf{AA}^T .

Order eigenvectors in a matrix, where each row corresponds to a given eigenvalue. The highest eigenvalue places its eigenvectors in the first row.

Eigenvector matrix is converted to an orthogonal matrix through Gram Schmidt orthonormalization to obtain \mathbf{U} .

Step 2. $\mathbf{A}^T\mathbf{A}$ to \mathbf{V}

Multiply \mathbf{A}^T and \mathbf{A} to obtain $\mathbf{A}^T\mathbf{A}$.

Determine eigenvalues and eigenvectors for $\mathbf{A}^T\mathbf{A}$.

Order eigenvectors as in step 1.

Eigenvector matrix is converted to an orthogonal matrix through Gram Schmidt orthonormalization to obtain \mathbf{V} .

\mathbf{V} is transposed to obtain \mathbf{V}^T .

Step 3. \mathbf{S}

Take the square roots of the non-zero eigenvalues for \mathbf{U} or \mathbf{V} .

Place on the diagonal from highest to lowest, starting in the right hand corner of the matrix. All other values of the matrix are zero.

The procedure can be expedited by the use of a PC containing MATLAB or Olis Globalworks software. MATLAB simply requires the command $[\mathbf{U},\mathbf{S},\mathbf{V}] = \text{svd}(\mathbf{A})$ to

perform the above calculation except V^T . This requires a separate command, $[Vt] = \text{transpose}(V)$.

2.2.4.2 Rationale of SVD

The rationale of SVD analysis (45 – 48) is to examine and reveal concealed information in a data set. Matrix A for spectroscopic data, contains n columns reflecting the experimental perturbation (salt, temperature and pH) and m rows corresponding to the measured variable at various wavelengths (absorbance, ellipticity) (8). Upon decomposition of matrix A, the U matrix contains the basis vectors describing the linear combinations of spectra. The S matrix reflects the singular values along its diagonal that are a weighing factor for each component. The magnitude of the singular value suggest the possible number of spectral species present. A plot of singular values versus the component number yields a curve that appears to undergo exponential decay. The singular values identified in the steep portion of that curve are related to the significant spectral species. Multiplication of $U \times S$ produces the matrix US. The US matrix is utilized to produce basis spectra for each component in the S matrix. The V matrix provides the amplitude vectors that corresponds to each spectral component. A plot of V versus the perturbation conditions (salt, temperature, pH), produces kinetic data that suggests certain equilibria. For example, if one component is diminishing, while a second component is increasing with temperature the equilibria may be interpreted as a folded to unfolded state transition with no intermediate. However, if a third component is present as a bell shaped curve in the plot, the data suggests an unfolding equilibrium involving an intermediate (8).

2.2.4.3 Complementary Analysis

Additional complementary analysis is required to determine the number of significant spectral species (8, 47, 48). The randomness in the **U** and **V** matrices is required to determine which components are significant. The first order autocorrelation function utilized is given as:

Equation 19:
$$C(X_i) = \sum X_{j,i} X_{j+1,i}$$

where $X_{j,i}$ is a row in column i . Therefore, the autocorrelation function monitors for fluctuations in a given column. Each column is autocorrelated to give a value between 1 and -1. Variation along a column in the form of random noise is assigned -1, while non-random values with slow row to row variations are assigned +1. $C(X_i) \geq 0.8$ has been previously utilized as a criteria for acceptance as a singular value.

A second analysis requires the construction of residual plots. Is one able to reproduce the data in **A** after removal of certain singular values? Theoretically, residual plots should appear random only after all significant singular values have been included in the analysis. **S** is modified to include the singular values of interest. The diagonal matrix is zeroed at all other singular values. The **S_x** matrix, containing x singular values, is substituted into equation [29] for **S**. The resultant **A_x** matrix is produced as shown below:

Equation 20:
$$A_x = US_xV^T$$

A_x is subtracted from the original **A** matrix to yield **R_x**, which is a residual matrix produced from x singular values. Contour surface plots of each **R_x** are assessed for randomness. Non-random behavior is qualified by ordered peaks in the plot. Random behavior is characterized by evenly scattered and dispersed peaks.

The complimentary procedures can also be performed in MATLAB. The autocorrelation requires the autocorrelation function to be written to a bin file. The residual plots require other commands that are easily viewed in the MATLAB help menu. SVD analysis can be carried out in OLIS, where the U and V matrices represent the spectral and kinetic eigenvectors respectively.

2.3 Differential Scanning Calorimetry (DSC)

2.3.1 Principles of DSC

Differential scanning calorimetry (DSC) (49, 60) is a technique for the measurement of the amount of energy required to establish a near zero temperature difference between a sample and a reference as both are subjected to a controlled heating or cooling regimen as a function of time (8). The amount of heat absorbed or released during a transition can be measured by measuring the difference in heat flow between the sample and reference cells. The heat capacity C_p is directly measured and related to the thermodynamic parameters that describe a macromolecule.

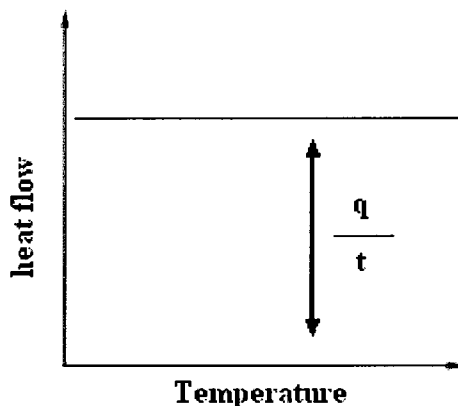


Figure 26: Heat Capacity

Equation 21: heat flow = q / t

Equation 22: heating rate = $\Delta T / t$

By dividing equation 21 by equation 22, we get the heat capacity

Equation 23: $q / \Delta T = C_p$

The following is known:

Equation 24: $dH = dU + pdV + Vdp$

where V is volume, p is pressure, H is enthalpy and U is internal energy of the system.

Equation 25: $dU = q - pdV$

At constant pressure, substituting 25 into 24:

Equation 26: $dH = q$

Substitution into 23 gives:

Equation 27: $dH = C_p \Delta T$

Integration produces:

Equation 28: $\Delta H = \int_{T_1}^{T_2} C_p \Delta T$

where $\int dH = \Delta H$, because enthalpy is a state function and path independent. Remember

that $\Delta H = dH_2 - dH_1$, where the subscripts indicate two different states of the system. If

intermediates are involved $\Delta H = \sum_i \Delta H_i$, where i is a specific state. Basically, in order to

obtain ΔH , one must account for all ΔH_i .

In a similar way, entropy of a system (S) can also be determined through heat capacity at constant pressure:

Equation 29: $dS = \delta q / T$

Substituting gives:

Equation 30: $dS = (C_p dT) / T$

Integration produces:

Equation 31:
$$\Delta S = \int_{T_1}^{T_2} (C_p/T) dT$$

Since Gibb's free energy is:

Equation 32:
$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$$

Results from equations 28 and 31 can be substituted into 32, to determine the free energy of the transition at a given temperature.

2.3.2 DSC Instrumentation

The Nano II power compensation Differential Scanning Calorimeter (DSC) (50) from Calorimetric Sciences Corporation was used to perform the experiments for this research. A schematic diagram describing the components of a typical power compensation DSC is provided (Figure 27). The reference and sample cells are enclosed in a thermal jacket and linked to a thermosensor and a power compensation heater. Both cells are pressurized under a sealed pressure handle to allow heating of aqueous solutions to 130°C (8).

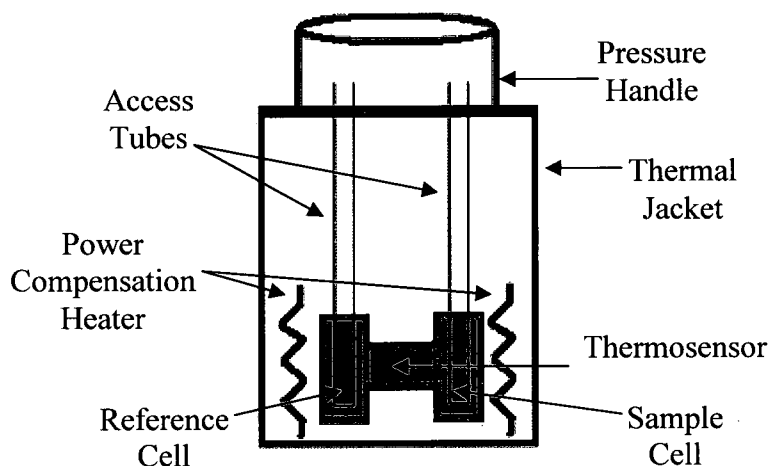


Figure 27: Schematic diagram of a typical power compensation DSC

Upon loading the sample and reference cells with sample and buffer respectively, the compartment is pressurized and allowed to equilibrate. The DSCRun version 2.1.1

software on the interfaced PC with the instrument allow us to program the parameters of the experiment. The temperature difference between the cells is monitored and measured by the thermosensor. The power compensation heaters are activated to nullify any differences. The data is recorded as μW applied to the sample as a function of temperature during heating/cooling scans as shown in (Figure 28) (8). A reversible process produces heating and cooling scans that are mirror image of each other.

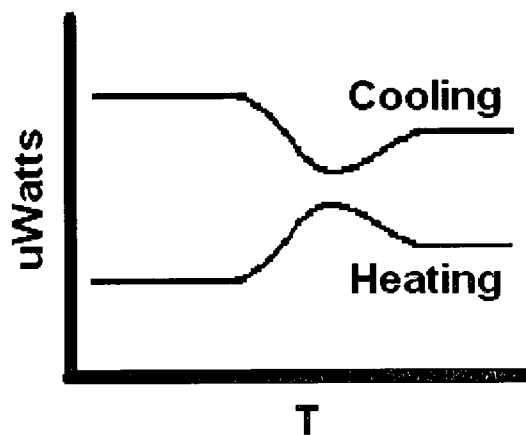


Figure 28: Typical Raw DSC scans

The data is subtracted from a previously run buffer-buffer DSC scan using CpCalc version 2.1 software (Calorimetric Sciences Corporation). Subsequently, the molar heat capacity is determined by inputting the concentration and molecular weight of the macromolecule.

As per the manufacturer recommendation, prior to any DSC experiment the instrument must undergo a thorough cleaning. The cleaning process include multiple heating scans with 4 M NaOH, and 50% aqueous formic acid solution, each followed by thorough washes with deionized water. Prior to sample introduction, a baseline scan using water should be performed to determine if additional cleaning is required or if the

instrument requires a balance and residual scan. The balance scan is a feature that allows one to compensate for instrumental difference in cells and their contribution to experiments. The balance scan is followed by a residual scan, which allows for the baseline to be adjusted for specific heating/cooling rates.

2.3.3 Data Analysis

Different baseline treatments are available in CpCalc: linear, polynomial and linear-polynomial (8). The molar heat capacity versus temperature plot is treated with one of the three possible baseline treatments to yield ΔC_P of zero and allow the integration of the peak area. As per equation 28, the area under the curve is the calorimetric enthalpy denoted as ΔH_{CAL} and the apex of the peak is the T_M for the transition (60). Replotting the data as C_P/T versus T , produces a similar curve. According to equation 31, the area under this curve is the calorimetric entropy given as ΔS_{CAL} . The free energy of the transition, ΔG at any given temperature can be calculated using equation 32. Figure 29 shows the calorimetric enthalpy and entropy plots.

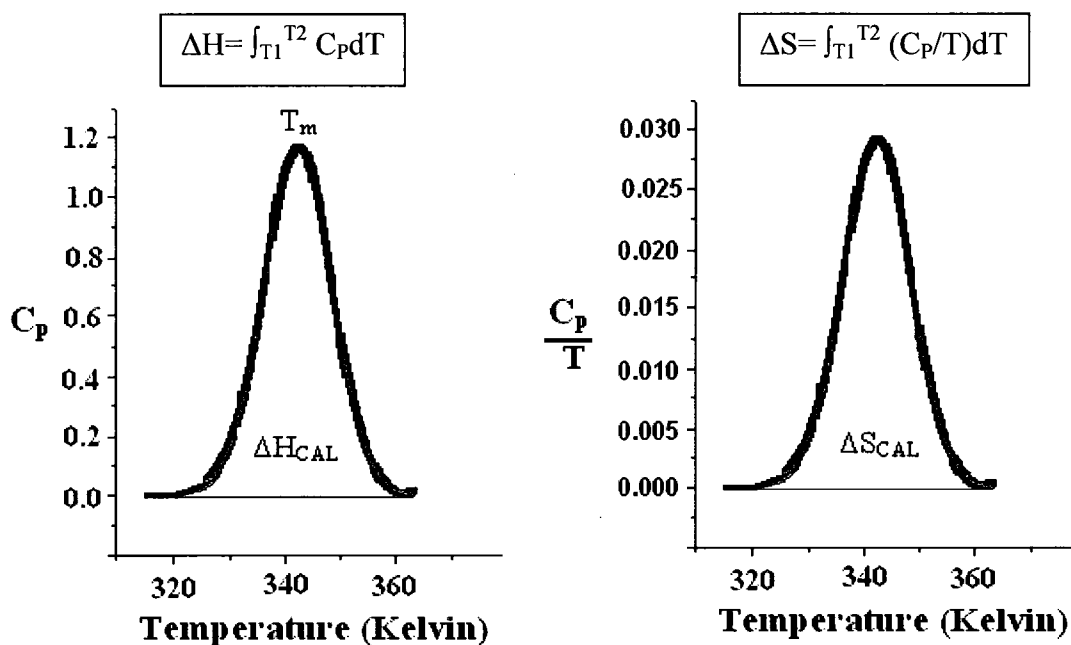


Figure 29: Calorimetric enthalpy and entropy calculation plot

Comparison of van't Hoff plots and calorimetric enthalpies can provide valuable information about the possibility of the presence of intermediates in the unfolding process.

If $\Delta H_{VH} = \Delta H_{CAL}$, the process is two state process with no intermediates.

If $\Delta H_{VH} < \Delta H_{CAL}$, the transition involves intermediates.

If $\Delta H_{VH} > \Delta H_{CAL}$, the process involves aggregation.

2.4 Predictive Structural Analysis

The secondary structure of nucleic acid molecules plays an important functional role. The prediction of secondary structures of nucleic acids is an important initial step in the process of studying and understanding their function. Predictions of secondary structure initially relied on sequence homology alignments (32, 33) and were usually performed on single sequences rather than families of homologous sequences. Now,

nucleic acids secondary structures can be predicted with great accuracy using *in silico* methods. Many bioinformatics applications use the concept of secondary structure in the analysis of nucleic acids. In addition, the availability of large families of homologous sequences has revolutionized secondary structure prediction algorithms (34). The technique utilized for this research is Mfold (35). It is a secondary structure prediction algorithm that is based on Nearest Neighbor Thermodynamic Models (NNTM) and Minimum Free Energies (MFE) (36).

One focus of this study is to predict *in silico*, the structural elements in the 5' end of the *bacterio-opsin (bop)* gene, present in the Archaeal organism, *Halobacterium salinarum*. This gene is predicted to have a stem-loop structure that may contribute to the highly expressed Bacteriorhodopsin (bR) protein (1). Then, these *in silico* predictions can be verified by the *in vitro* thermodynamic analysis of the putative stem loop. We will perform studies to determine if the stem loop exists, its form and what predicted structures can interconvert.

The hypersaline organism *H. salinarium* grows optimally at 4 M NaCl (37). It is known that monovalent cations stabilize DNA structures up to 1.0 M. One of the limitations we encountered using the *in silico* method is that all known prediction algorithms do not account for salt concentrations above 1.0 M.

3 Materials and Methods

3.1 Materials

3.1.1 DNA oligo

HPLC purified 25mer stem-loop DNA sequence (5' to 3')

GCATGTTGGAGTTATTGCCAACAGC (MW = 7697) was purchased from Sigma-Genosys. The sample was dialyzed (3,500 MWCO Pierce Slide-A-Lyzer Dialysis Cassettes) versus deionized H₂O. It should be noted that the DNA concentration was determined using UV spectroscopy at 260 nm prior to the start of each K⁺ concentration after equilibration at 95°C to prevent secondary structures formations and the associated hypochromicity shift. Stock solutions of each K⁺ concentration were prepared from 10X standard potassium phosphate buffer (100 mM mono- and dibasic forms with 1 mM K₂EDTA), deionized H₂O, and the appropriate volume of 5M KCl. All solutions were at pH 7.0 and were filtered through a 0.45 μm filter and then degassed by sonication.

3.1.2 RNA oligo

Similarly, HPLC-purified stem-loop 25mer RNA sequence (5'-GCAUGUUGGAGUUAUUGCCAACAGC-3') (MW = 7998) was purchased from Biosynthesis and treated in the same manner as the DNA sample except for the use of nuclease-free water instead of the deionized water.

3.1.3 Preparation of the Stem-loop

Nucleic acid samples were reconstituted in phosphate buffer at different K⁺ concentrations ranging from 15mM to 4M. Each sample was then heated to 95°C for two

minutes. The heating block was turned off and the sample allowed to slowly equilibrate to room temperature. The sample was removed and incubated at 5°C for 48 hr. At the completion of the incubation period, the sample was dialyzed exhaustively versus phosphate buffer at the identical K⁺ concentration for 24hr. The sample was then divided into aliquots and used separately in UV spectroscopy, circular dichroism and differential scanning calorimetry experiments. Samples were degassed by gentle sonication for 10 min prior to each experiment.

3.1.4 Buffers

Potassium Phosphate Buffer Solution (KPBS) 10X was prepared and utilized as a stock solution to prepare subsequent working buffer concentrations to make the sample solutions.

Table 1: KPBS stock solution preparation (10X)

Reagent	Molecular Weight	Moles	Grams
Dibasic K ₂ HPO ₄	174.18	0.058	10.102
Monobasic KH ₂ PO ₄	136.09	0.042	5.716
K ₂ EDTA:2H ₂ O	404.5	0.001	0.405

1. Add above reagents to a volumetric flask and fill to 1 L with distilled deionized (dd)H₂O.
2. Mix and filter through 0.45 μm filter (store at 4°C).

KPBS working solution preparation:

1. Mix 100 mL of KPBS stock solution (10X) with distilled deionized H₂O in 1L volumetric flask (approximately contains 15.8 mM K⁺)

2. Mix and bring the solution to pH 7.0 with KOH (1M). Note the volume added (3 mM K⁺ typically added)
3. Add the additional K⁺ in the form of 5 M KCl solution to obtain the desired working buffer concentration according to Table 1
4. Fill to 1L with distilled deionized (dd)H₂O.
5. Filter through 0.45 μm filter and store at 4°C.

Table 2: Working Potassium Phosphate Buffer Solution (KPBS) preparation

Working buffer concentration (M)	Additional required K⁺ (mM)	5 M KCl stock (mL)
0.015	0	0
0.050	34.2	6.8
0.100	84.2	17
0.250	234.2	46.8
0.500	484.2	96.8
1.0	984.2	197
2.0	1984.2	396.8
3.0	2984.2	597
4.0	3984.2	796.8

3.1.5 Cleaning Solutions

The DSC was subjected to a cleaning protocol with the following solvents: 4 M NaOH, 50% formic acid diluted from an 88% solution and THF. Prior to any DSC sample scans, the DSC cells were heated in the presence of 4M NaOH at a rate of

2°C/min from 25°C to 75°C. The cells were rinsed with 1 L deionized water and the procedure repeated with 50 % Formic Acid. The cells were vigorously rinsed (2-3 L) once again. If the water-water scan provided an aberrant scan, then the instrument underwent a baseline scan with water to adjust for instrumental differences. It should be noted that it was sufficient to rinse the cells with 4 L of water after each DNA sample. Intermediate water-water scans provided no aberrant scans.

All cuvettes were washed with 1 N HCl, deionized water and methanol, followed by drying with N₂ (zero grade).

3.2 Methods

3.2.1 Ultraviolet-visible spectroscopy (UV/Vis)

UV melting studies were carried out on Cary-3E Spectrophotometer equipped with temperature controlled sample compartment, multicell changer and N₂ purged sample chamber. All cuvettes were cleaned as previously described (Section 3.1.5). DNA and RNA melts were monitored at 260 nm and thermally scanned from 5°C to 95°C. Total of six thermal scans alternating between heating and cooling scans ramped between 5°C and 95°C at a rate of 0.25°C min⁻¹. The data points were collected every 0.25°C versus its respective working KPBS buffer concentration as a blank. The resulting data was exported to Origin 6.1, where it was subjected to van't Hoff analysis assuming two state transitions.

3.2.2 Circular Dichroism Spectropolarimetry (CD)

Circular Dichroism spectra were recorded with an Aviv Associates CD Spectropolarimeter Model 62-ADS equipped with Peltier heating/cooling device and

nitrogen purging capabilities. An identical macro for all samples was written in DOS to facilitate data collection every 1°C, beginning at 5°C and ending at 99°C. After 0.5 min equilibration at each sample temperature, the instrument collected spectral data in the 320-220 nm range, with a scan step of 1nm, and an averaging time of 1s and a bandwidth of 1.5 nm. Prior to each melt, the sample compartment was allowed to equilibrate to 5°C for 30 minutes. Samples stored at 5°C were transferred to stoppered cuvettes and placed in the instrument for 30 minutes to reach thermal equilibrium prior to melt initiation. The data, after subtraction of the buffer CD over the same spectral range and subsequent smoothing, were exported to Origin 6.1 for conversion into molar ellipticity data. All cuvettes were cleaned as previously described (Section 3.1.5).

3.2.3 Differential Scanning Calorimeter (DSC)

All calorimetric experiments were carried out with 6100 N-DSCII Nano-Differential Scanning Calorimeter from Calorimetric Sciences Corporation. The DSC is interfaced with a PC containing DSCRun version 2.1.1 and CpCalc version 2.1 software (Calorimetric Sciences Corporation). The DSCRun software allows for user interface with the DSC to modulate experimental parameters and run experiments. The CpCalc software was utilized to subtract buffer-buffer scans from buffer-sample scans. Linear-polynomial baselines were drawn for each scan, and molar heat capacities generated with the molecular weight of the oligomer and the concentration (milligrams per milliliter) used during each scan as input. The molar heat capacity data were exported to Origin version 6.1, where it was deconvoluted into composite transitions using a Gaussian multipeak fitting function. The area of each peak yields the transition enthalpy, and the peak maximum yields the transition melting temperature. Since $\Delta S^\circ = \int ((\Delta C_p/T)/dT)$, the

calorimetric Gaussian fits can be replotted in Origin 6.1 as $\Delta C_p/T$ versus T to yield a curve, the area of which is ΔS° . The enthalpy and the entropy determinations allowed for the calculation of Gibb's free energy through the following equation: $\Delta G = \Delta H - T\Delta S$ at $T = 37^\circ\text{C}$. The DSC was subjected to the cleaning protocol previously described in section 3.1.5.

The reference cell contained the degassed dialysis buffer at its respective K^+ concentration. The sample cell contained the DNA or RNA stem-loop sample in buffer identical to that in the reference cell. The DNA and RNA stem-loop sample concentrations ranged from 7.8 – 3.6 mg/mL. Each DNA/RNA sample was thermally scanned from 10°C to 90°C at $0.25^\circ\text{C} / \text{minute}$ over 3 forward and 3 reverse scans to ensure reversibility as an indication of thermal equilibrium.

3.2.4 Singular-Value Decomposition (SVD).

Singular-value decomposition was performed using Matlab 6.5.0.1 (Mathworks) and Olis Global works. SVD is a tool for enumerating the number of species required to account for the spectral changes that accompany a transition. CD spectra were obtained at 95 different temperatures ($5\text{-}99^\circ\text{C}$ at 1°C increments) and collected in the wavelength range of 220-320 nm, at 1 nm intervals. Therefore, the initial 101×95 data matrix A for SVD analysis was organized such that the rows consisted of data collected at 1 nm increments of wavelength (220-320 nm) and the columns at 1°C increments ($5\text{-}95^\circ\text{C}$) for each set of data collected at a particular concentration of K^+ .

4 Results

4.1 Stem-loop Purity

10% TBE-Urea polyacrylamide gel was used to test the purity of the stem-loop sequence before and after each experiment. The presence of any failure sequences or degradation product is detectable by this method. Figure 30 shows a representative gel with the DNA ladder in lane 1 containing 18 and 31 mer. Lane 2 contained 2 ng of the stem-loop sequence in 15mM buffer before the start of the biophysical analysis experiments. Lane 3 contained 2 ng of the stem-loop sequence in 4M buffer after the completion of the biophysical analysis experiments. Our oligomer is of the appropriate size and does not contain failure sequence oligomers or any degradation sequences. Failure sequences are commonly observed in automated synthesis products prior to purification or post insufficient purification while degradation sequences are commonly observed after repeated heating/cooling cycles or as a result of mishandling.

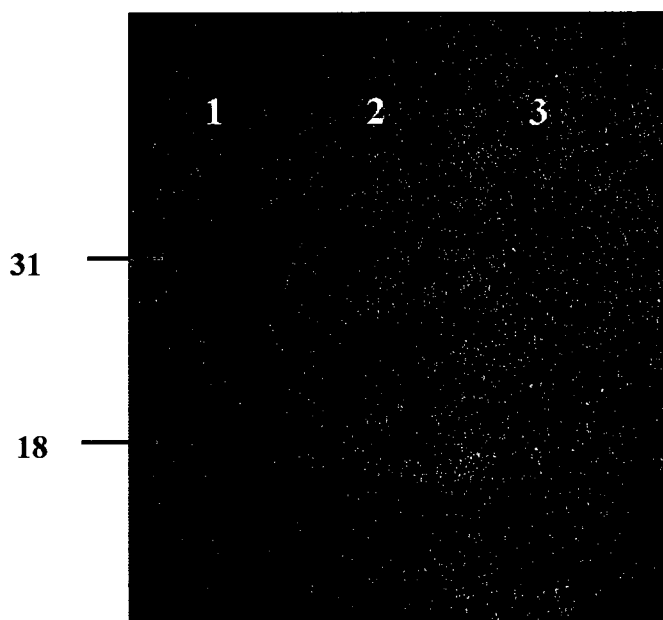


Figure 30: 10% TBE-Urea polyacrylamide gel

4.2 Predictive Structural Analysis

4.2.1 Nucleic acid sequence analysis

Nucleic acid sequence homology alignments and analysis suggest that the *bop* gene mRNA is leaderless with no 5' untranslated nucleotides. The first twenty-five bases of the *bop* gene mRNA can form a secondary structural element “stem-loop” that is essential for gene expression (Figure 31). Therefore, any *cis*-acting determinants of *bop* gene expression that exist in the 5' region are within the coding region.

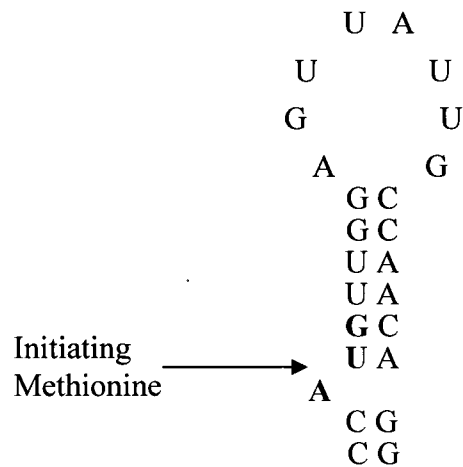


Figure 31: 5' end of the *bop* gene mRNA

4.2.2 *In silico* sequence analysis

Mfold (35) was used as an *in silico* approach to predict the secondary structure element expected to be present in the first twenty-five bases of the *bop* gene mRNA. Predictions were made using both DNA and RNA model sequences at various temperatures and salt concentrations. The melting temperature and the thermodynamic values for the predicted structures were used as a measure for the structural stability.

4.2.2.1 DNA model

Structure B depicted in Figure 32 is the stem loop predicted by sequence analysis. Figure 32 also shows that there is a probability of an alternative secondary structure (structure A) that is thermodynamically more stable. This stability is evident in the presence of NaCl concentrations varying between 15mM to 1.0 M (Figure 33 and 34) and temperatures up to 37°C.

Please note that all known DNA prediction algorithms do not account for salt concentrations above 1.0 M. At the maximum allowable salt concentration (1M) and at temperatures above 37°C, *Mfold* predict only one structure to be present (Structure A) and the disappearance of structure B (Figure 35).

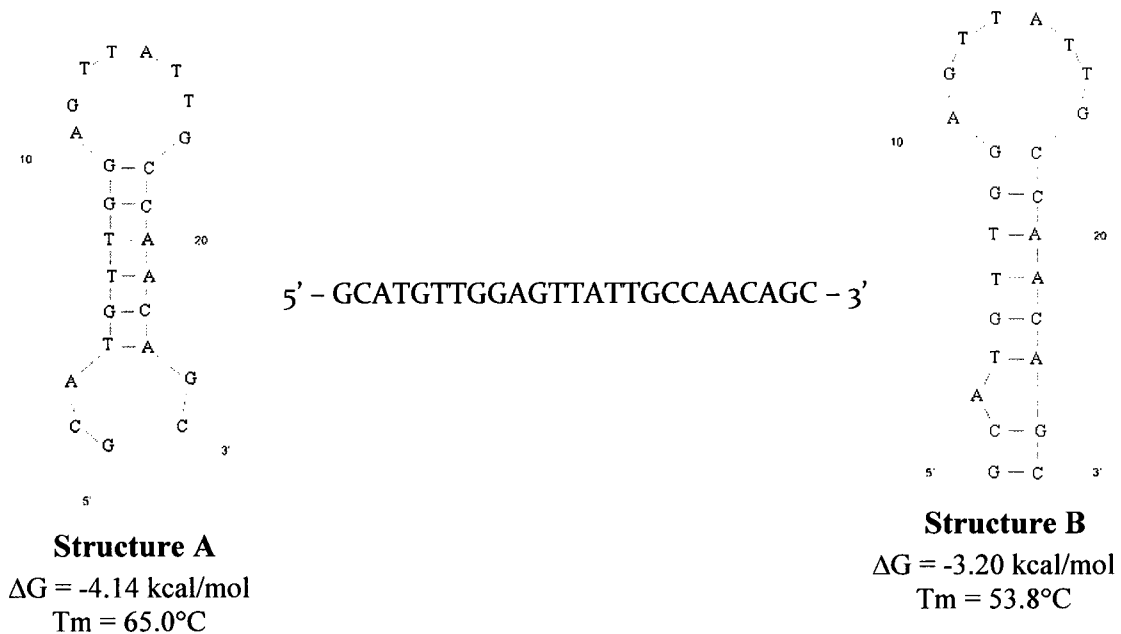


Figure 32: Mfold DNA model predicted secondary structures at 37°C, 1M salt

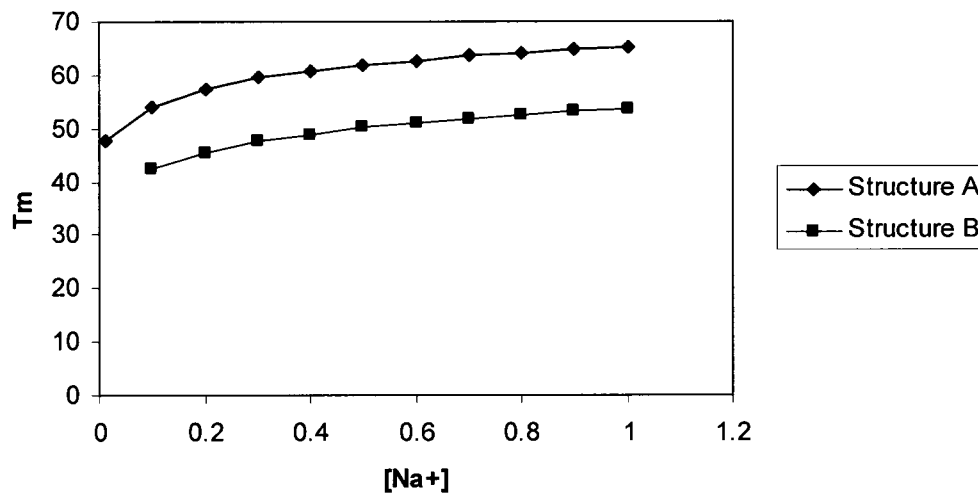


Figure 33: Mfold DNA model predicted secondary structures at 25°C

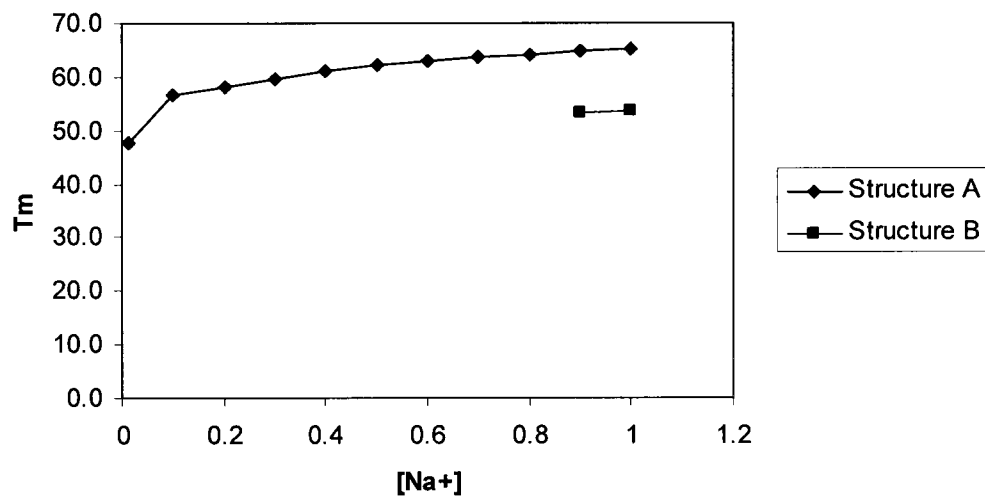


Figure 34: Mfold DNA model predicted secondary structures at 37°C

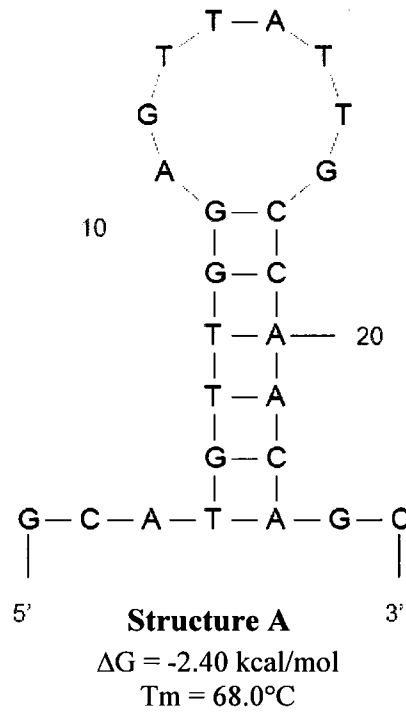
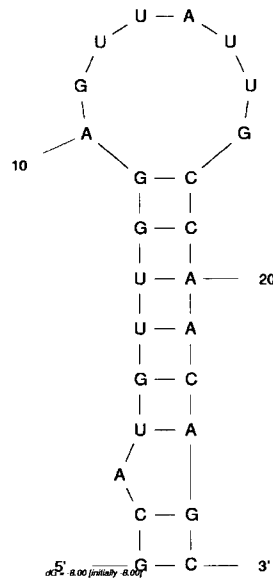


Figure 35: Mfold DNA model predicted secondary structures at 50°C, 1M salt

4.2.2.2 RNA model

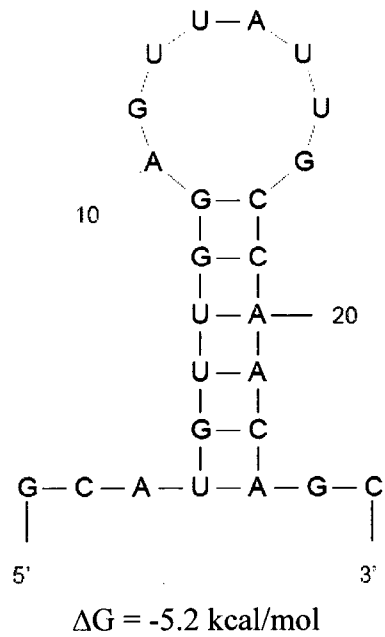
Results from *Mfold* suggest one structure is predicted for the RNA model at 1.0 M NaCl and 37°C. (Figure 36). Please note that all known RNA prediction algorithms do not support or account for salt corrections. Another RNA prediction server (UNAFold) allows for temperature corrections. It suggests that at a NaCl concentration of 1.0M and temperatures above 37°C, an alternate structure is present (Figure 37). This structure is similar to structure A of the DNA model (Figure 35).



$$\Delta G = -7.60 \text{ kcal/mol}$$

$$T_M = 70.8^\circ\text{C}$$

Figure 36: Mfold RNA model predicted secondary structures



$$\Delta G = -5.2 \text{ kcal/mol}$$

Figure 37: UNAFold RNA model predicted secondary structures at 50°C, 1M salt

4.3 Biophysical analysis of the DNA model sequence

4.3.1 UV Spectroscopy

4.3.1.1 Reversibility

Repetitive UV thermal melts at 260 nm were performed on the DNA stem-loop as depicted in (Figure 38). Clearly, the repetitive scans illustrate the folding to unfolding transition is a reversible process involving a hyperchromic shift. A total of six thermal scans alternating between heating and cooling scans ramped between 5°C and 95°C were collected. At the T_M , K_{folding} can be represented as:

Equation 33:
$$K_{\text{folding}} = [\text{folded}] / [\text{unfolded}] = 1$$

It is understood that an equilibrium process must be reversible. This requires the heating and cooling scans of a melting curve overlay each other. Many times researchers observe that the heating/cooling curves do not overlay each other resulting from hysteresis.

Hysteresis may be observed if the rate of sample heating is too high or if irreversible alterations have occurred to the sample.

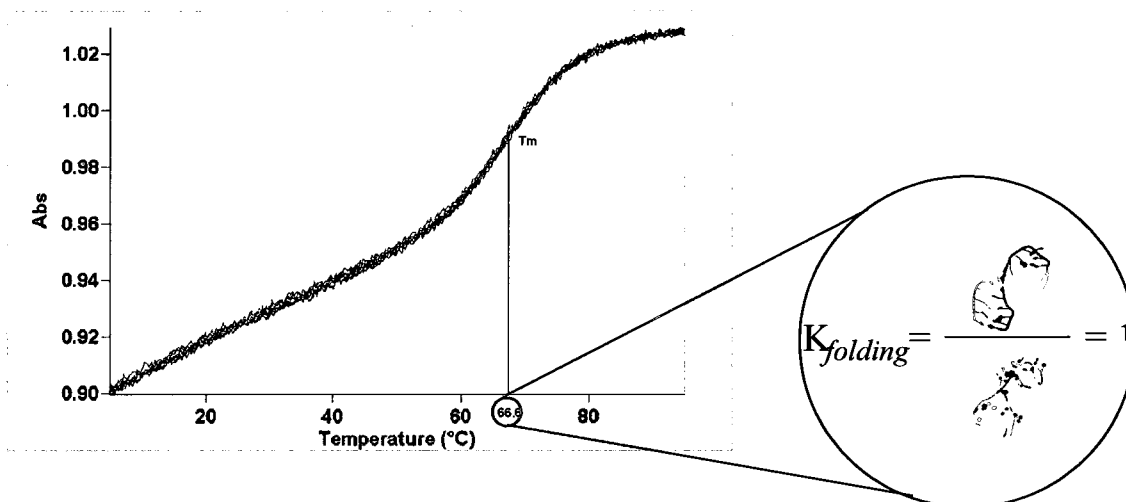


Figure 38: Repetitive DNA model heating and cooling scans at 260nm in 1M [K⁺]

4.3.1.2 Unimolecular

UV thermal melts were performed at different potassium ion concentrations to facilitate a plot of $1/T_M$ versus \ln [strand concentration] shown in Figure 39.

The analysis of the calculated T_M values from the UV thermal melt revealed that the T_M is independent of the DNA concentration as evidenced by a slope on the order of 10^{-6} ; therefore, the reaction is unimolecular. The above mentioned two criteria (reversibility and unimolecularity) are essential before performing any type of thermodynamic analysis of conformational change.

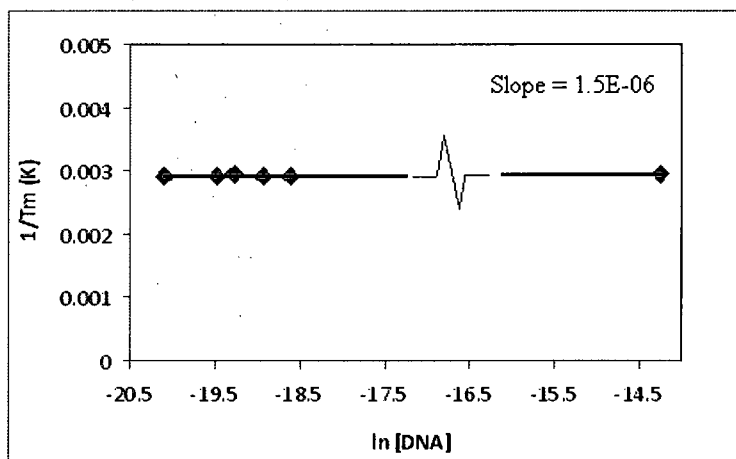


Figure 39: Values of $1/T_M$ plotted Vs the natural log of the [DNA] in 1M K⁺

4.3.1.3 Salt dependent UV thermal melt of DNA stem-loop model

Salt dependent UV spectra of the DNA stem-loop sequences as a function of temperature collected at 260 nm are presented in Figures 40. Total of six thermal scans alternating between heating and cooling scans ramped between 5°C and 95°C were collected at each potassium ions concentrations as follow 15, 50, 100, 250, 500, 1000, 2000, 3000 and 4000 mM K⁺. The data points were collected every 0.25°C in its respective working KPBS buffer concentration as a blank. These melting curves show an apparent transition midpoint (T_M) that is directly proportional to the salt concentration. Differences in T_M are a direct link to differences in secondary structure stability.

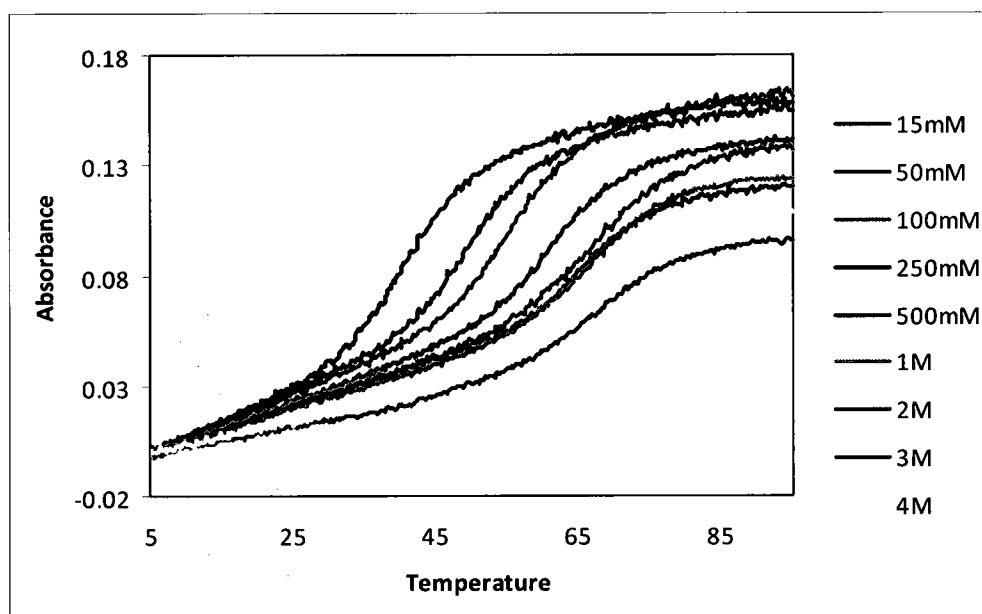


Figure 40: UV spectra of the DNA model of the stem-loop at 260 nm as a function of temperature at different salt concentrations.

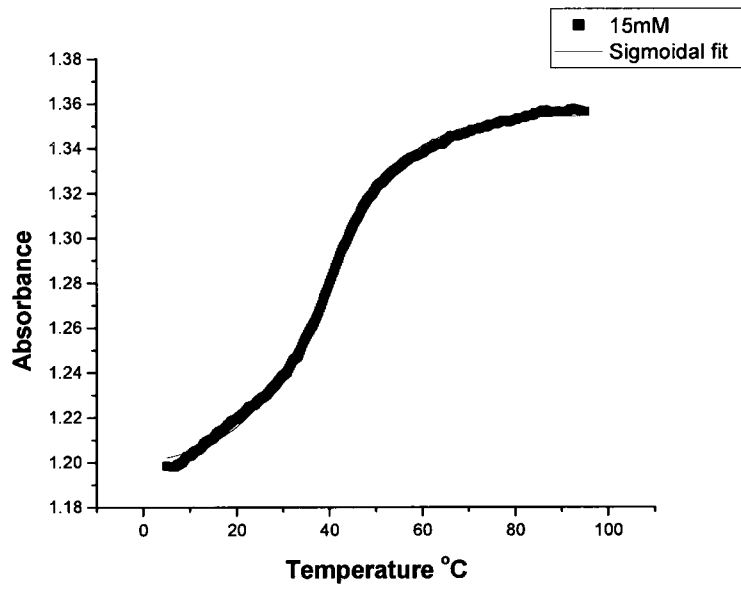
Close inspection of the melting curves reveal a hyperchromic transition which is an indication of loss of base pairing.

4.3.1.4 van't Hoff Analysis

The collected data was subjected to van't Hoff analysis at various salt concentrations. The data was converted in terms of α (fraction folded DNA) and utilized to calculate K_F . van't Hoff plot ($R\ln K$ Vs $1/T$) was constructed yielding a slope and an intercept that are essential for van't Hoff enthalpy and entropy determination respectively.

Figures 41a – 49a provide the sigmoidal fit for the UV melts while Figures 41b -49b provide van't Hoff plot for the same melts. Table 3 presents the van't Hoff data obtained at various salt concentrations. T_M and ΔH_{VH} values were determined from UV absorbance data at 260 nm. All data is compiled from 6 scans.

A.



B.

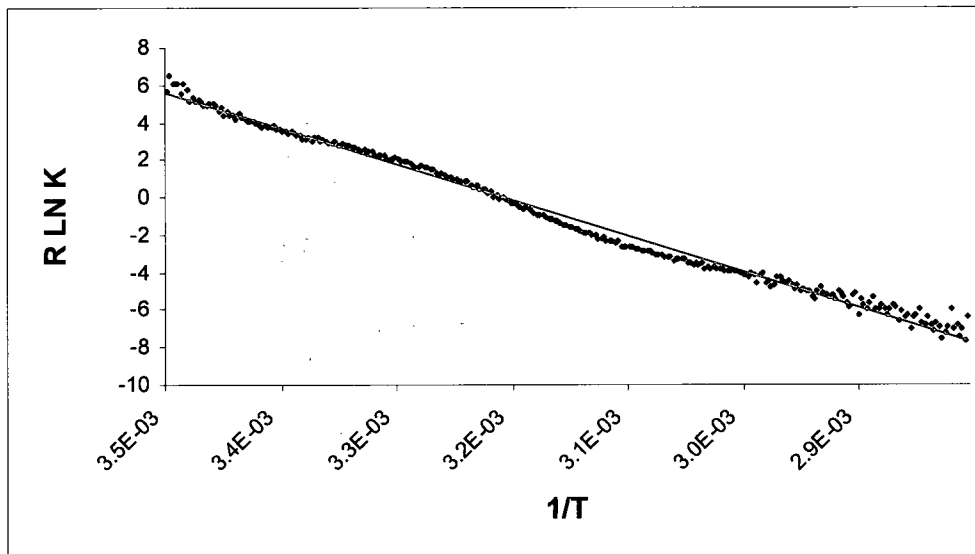
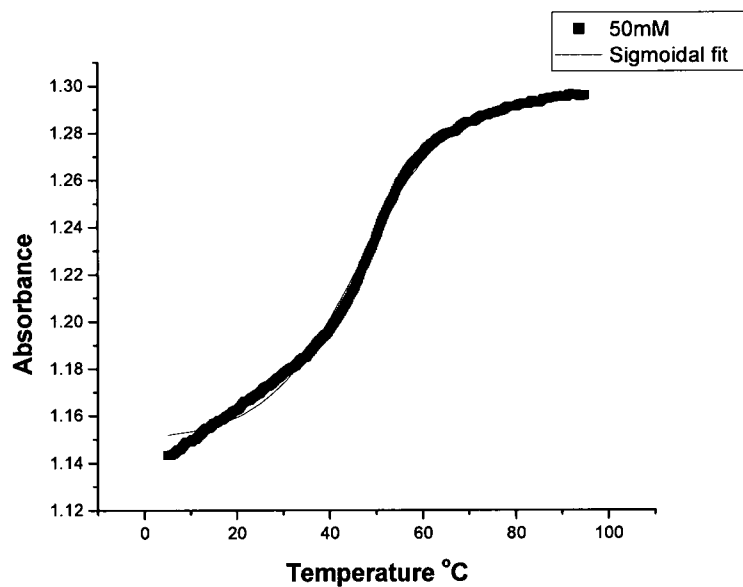


Figure 41: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 15 mM K⁺, (B) van't Hoff plot

A.



B.

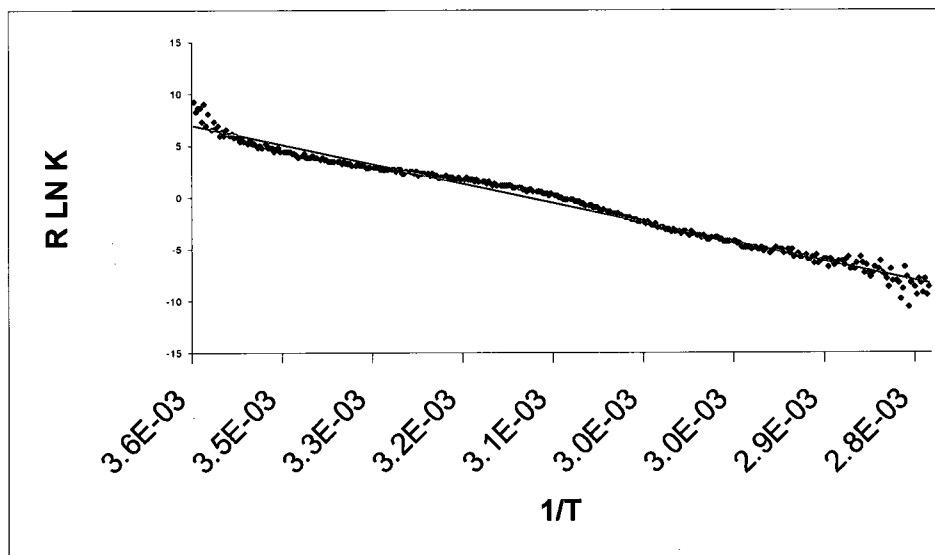
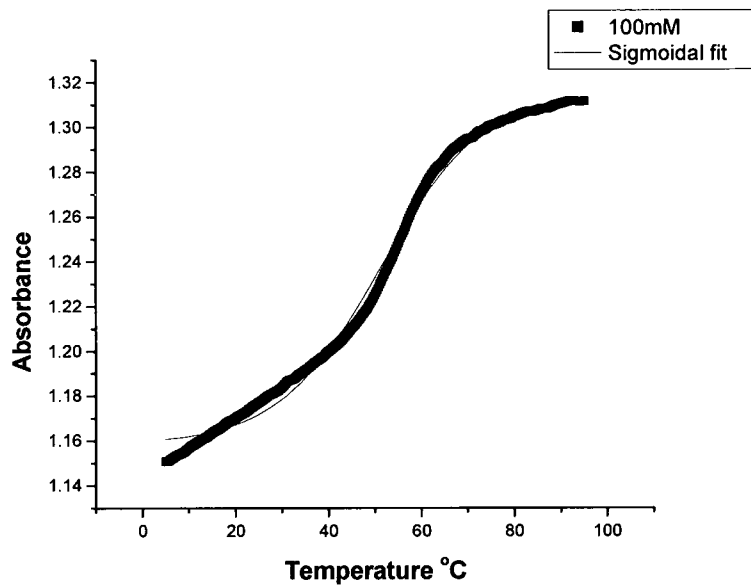


Figure 42: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 50 mM K⁺, (B) van't Hoff plot

A.



B.

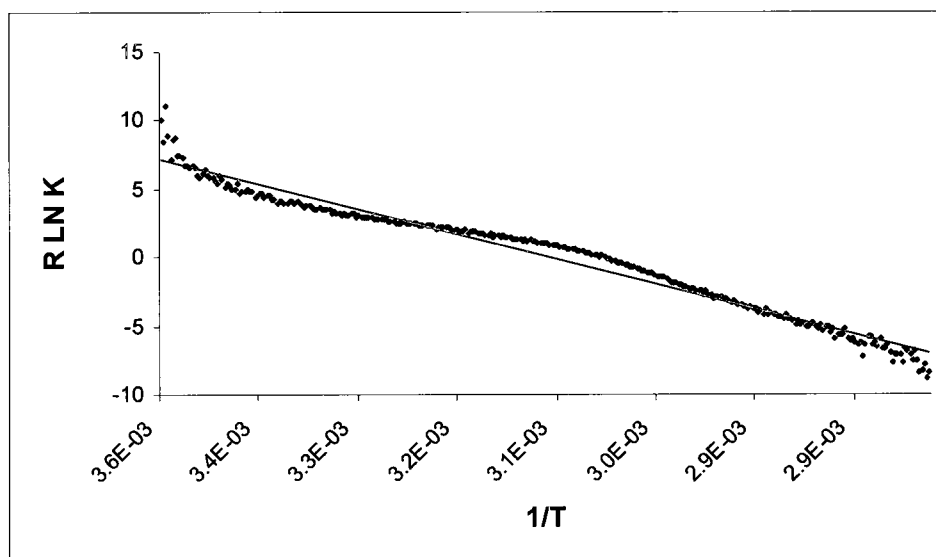
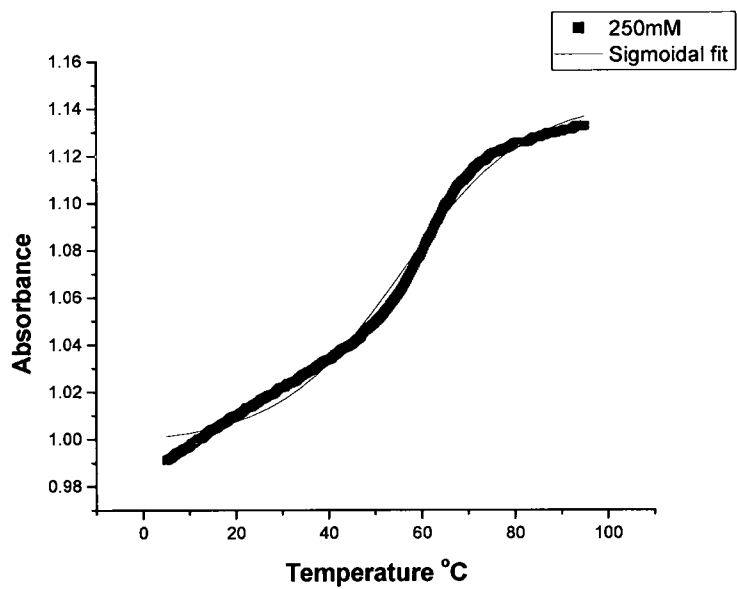


Figure 43: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 100 mM K⁺, (B) van't Hoff plot

A.



B.

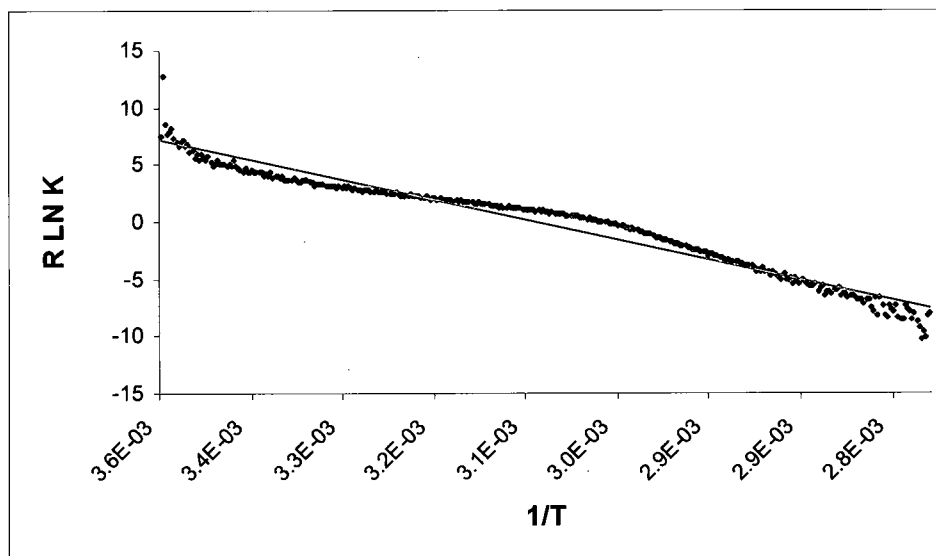
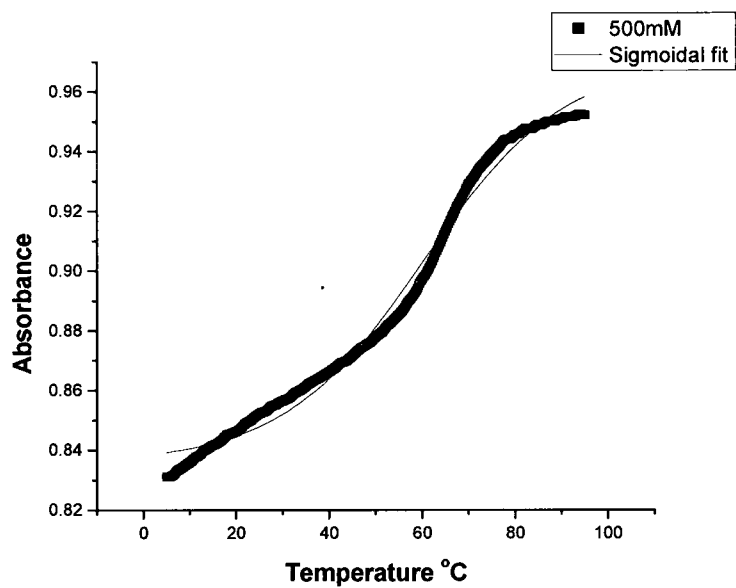


Figure 44: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 250 mM K⁺, (B) van't Hoff plot

A.



B.

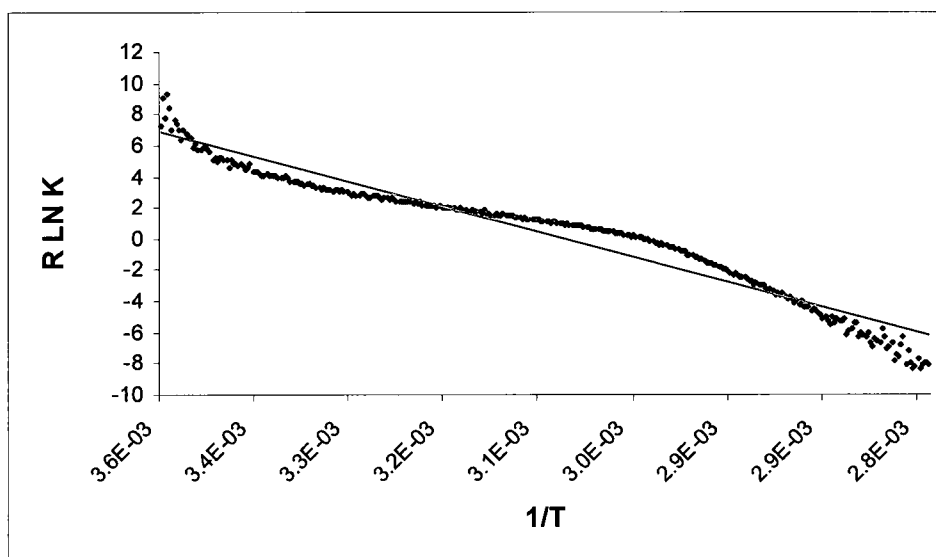
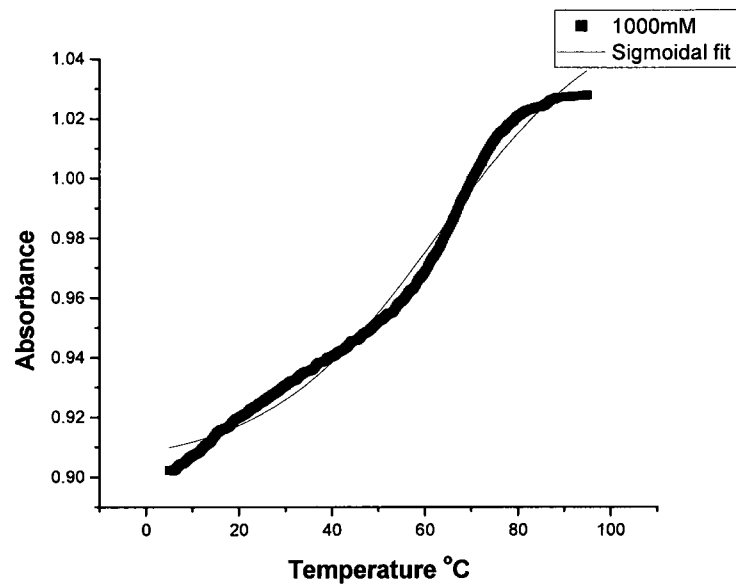


Figure 45: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 500 mM K⁺, (B) van't Hoff plot

A.



B.

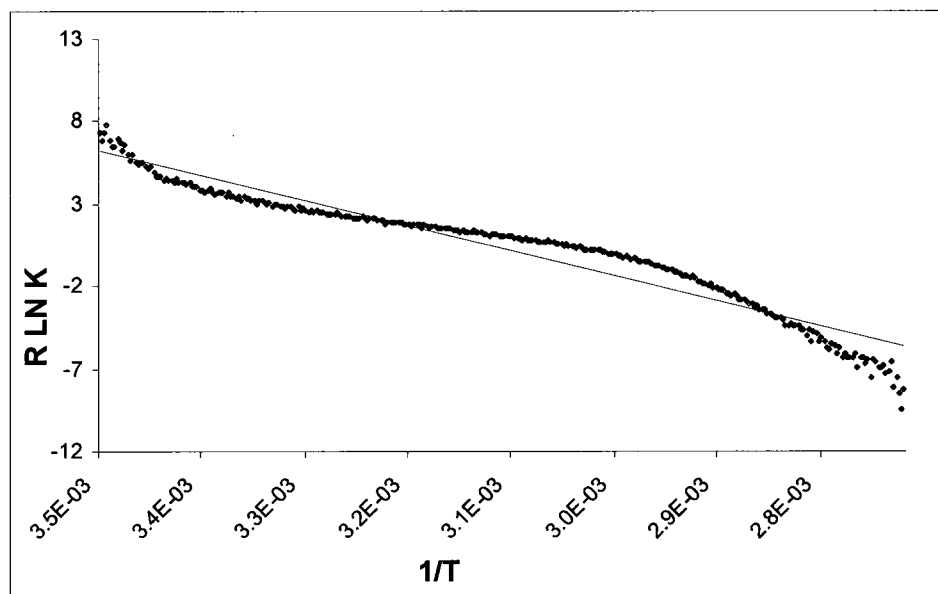
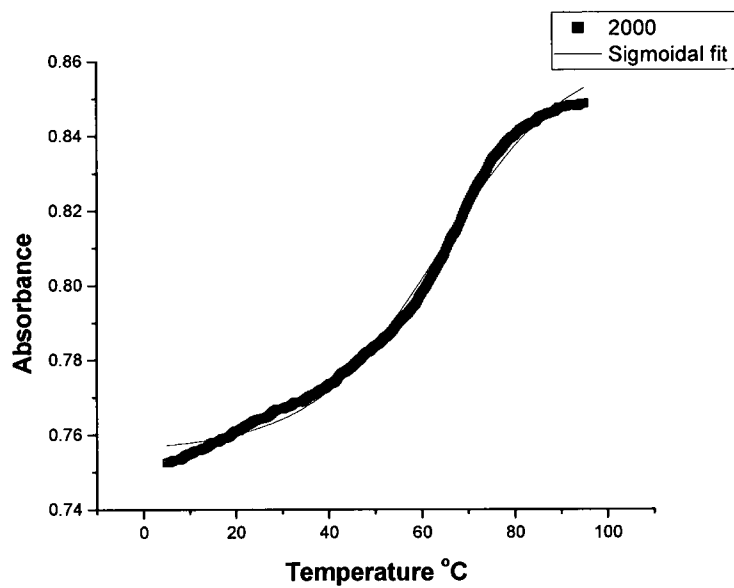


Figure 46: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 1000 mM K⁺, (B) van't Hoff plot

A.



B.

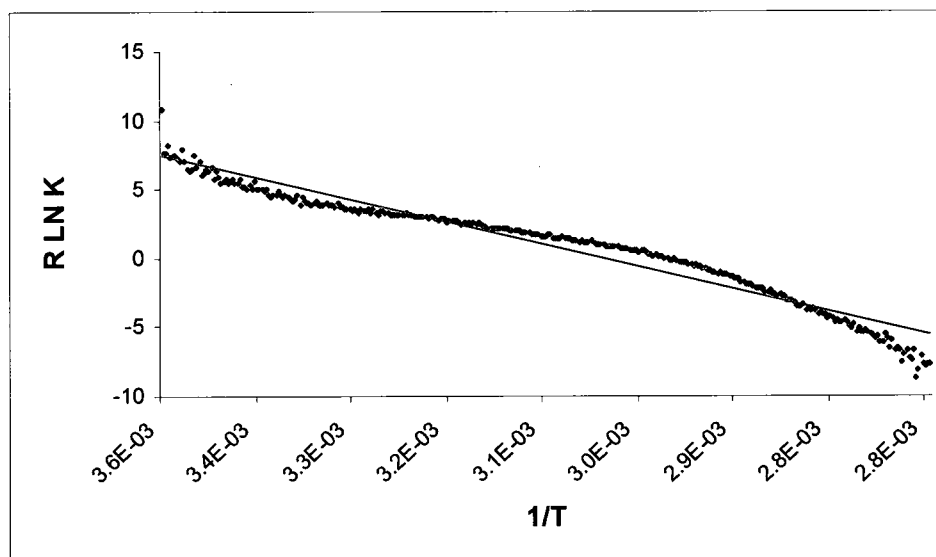
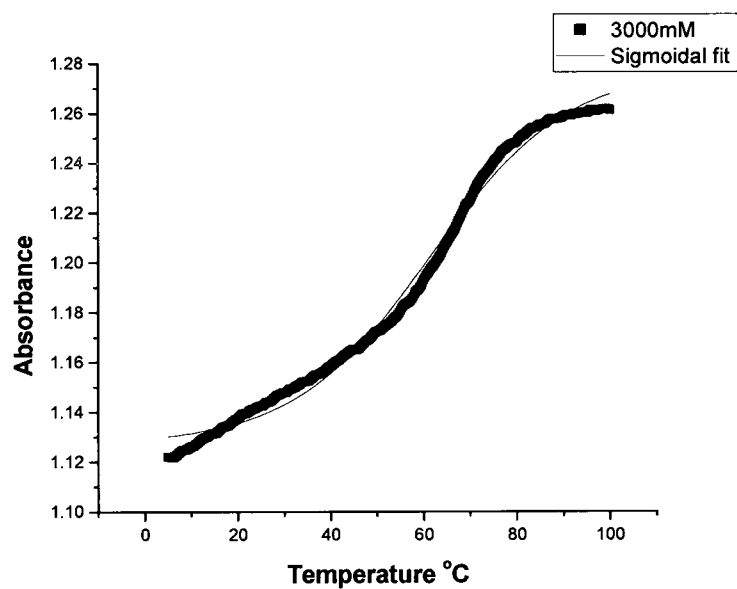


Figure 47: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 2000 mM K⁺, (B) van't Hoff plot

A.



B.

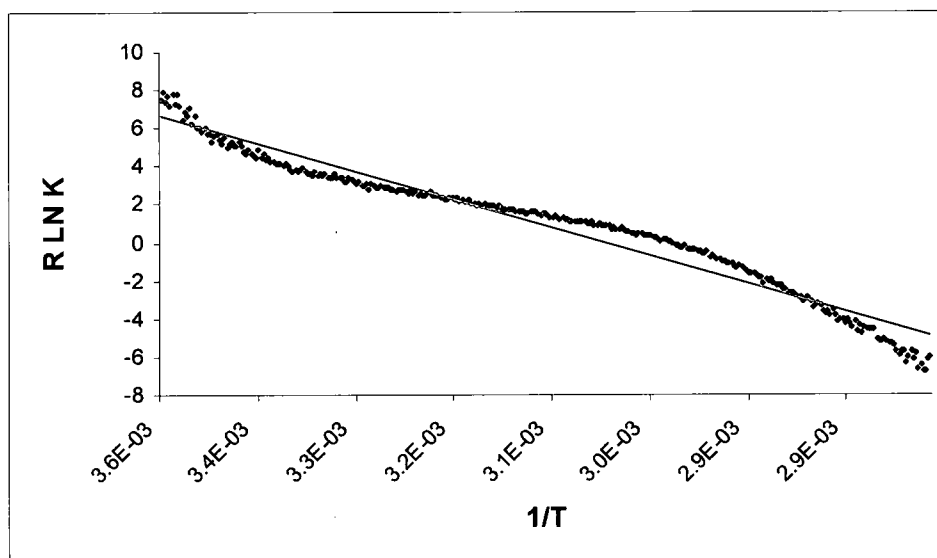
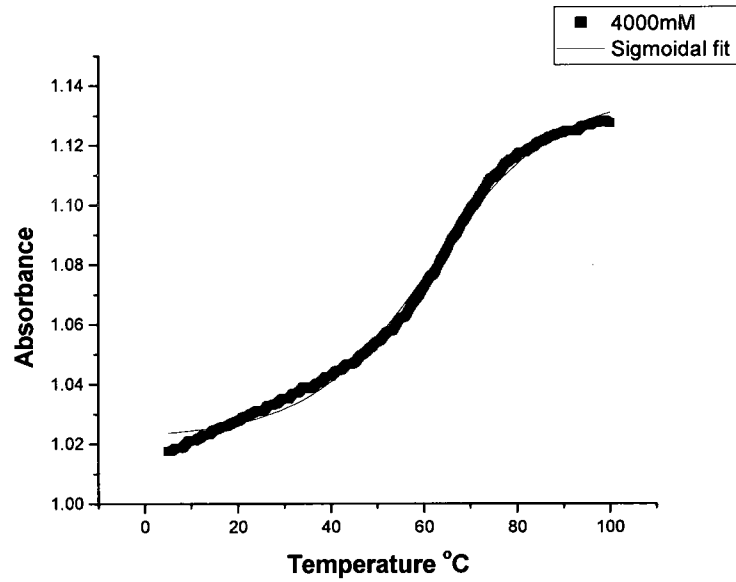


Figure 48: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 3000 mM K⁺, (B) van't Hoff plot

A.



B.

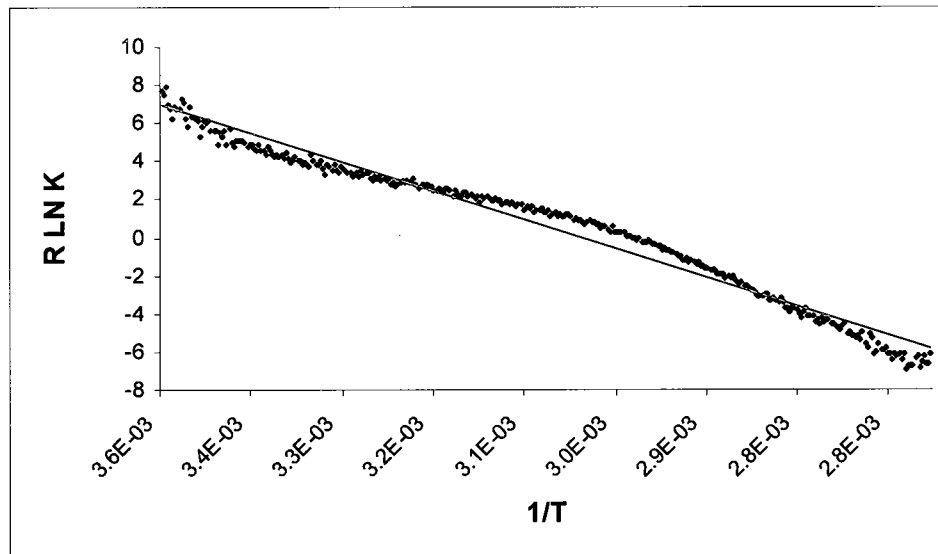


Figure 49: (A) Sigmoidal fit of the UV melting of the stem-loop DNA model in 4000 mM K⁺, (B) van't Hoff plot

Careful review of the aforementioned sigmoidal fit curves and van't Hoff plots (Figures 41 – 49) reveals several issues. The increased stability and the inability to obtain a sufficient upper baseline at the higher salt concentrations, makes it quite difficult to establish an apparent T_M values for these transitions. The assignment of inappropriate T_M values for these transitions may lead to miscalculated van't Hoff thermodynamic data. Systemic deviation from a simple sigmoidal shape curve was observed at most of the sigmoidal fit curves at all tested salt concentrations. Lastly, as van't Hoff analysis assumes a two state transition, the results highlight clear deviation from that assumption. This is evident by the nonlinear van't Hoff plot obtained at all salt concentrations.

Table 3: van't Hoff thermodynamic parameters as a function of $[K^+]$

$[K^+]$	T_M (°C)	ΔH_{VH} (kcal mol ⁻¹)	ΔS_{VH} (kcal K ⁻¹ mol ⁻¹)	$\Delta G_{VH 42^\circ C}$ (kcal mol ⁻¹)
15 mM	39.7 ± 0.3	-19.8 ± 0.5	-63.6 ± 1.8	0.20 ± 0.11
50 mM	48.9 ± 0.5	-18.7 ± 0.2	-59.6 ± 0.8	0.09 ± 0.02
100 mM	54.2 ± 0.2	-17.9 ± 0.3	-56.8 ± 1.1	-0.04 ± 0.02
250 mM	60.7 ± 0.2	-16.7 ± 0.3	-52.5 ± 0.8	-0.13 ± 0.03
500 mM	64.8 ± 0.3	-15.2 ± 0.3	-47.5 ± 1.0	-0.21 ± 0.04
1000 mM	66.6 ± 0.3	-16.3 ± 0.4	-47.5 ± 1.2	-0.31 ± 0.02
2000mM	67.1 ± 0.2	-15.0 ± 0.3	-46.4 ± 0.9	-0.40 ± 0.02
3000mM	67.1 ± 0.3	-14.9 ± 0.4	-46.0 ± 1.2	-0.39 ± 0.03
4000mM	65.4 ± 0.3	-15.7 ± 0.2	-48.5 ± 0.6	-0.46 ± 0.02

4.3.2 Circular Dichroism Spectroscopy

4.3.2.1 Temperature dependent CD Spectra of DNA stem-loop model

Figure 50 provides the CD spectrum of the DNA stem-loop at 5°C in 15 mM K⁺. A peak is observed at 280 nm and trough at 250 nm. 250 nm was selected to monitor and analyze the data as more apparent spectral transition was detected at this wavelength. The temperature dependent spectra of the DNA stem-loop model are provided in Figures 51 – 59 at different potassium ions concentrations as follow 15, 50, 100, 250, 500, 1000, 2000, 3000 and 4000 mM K⁺, respectively.

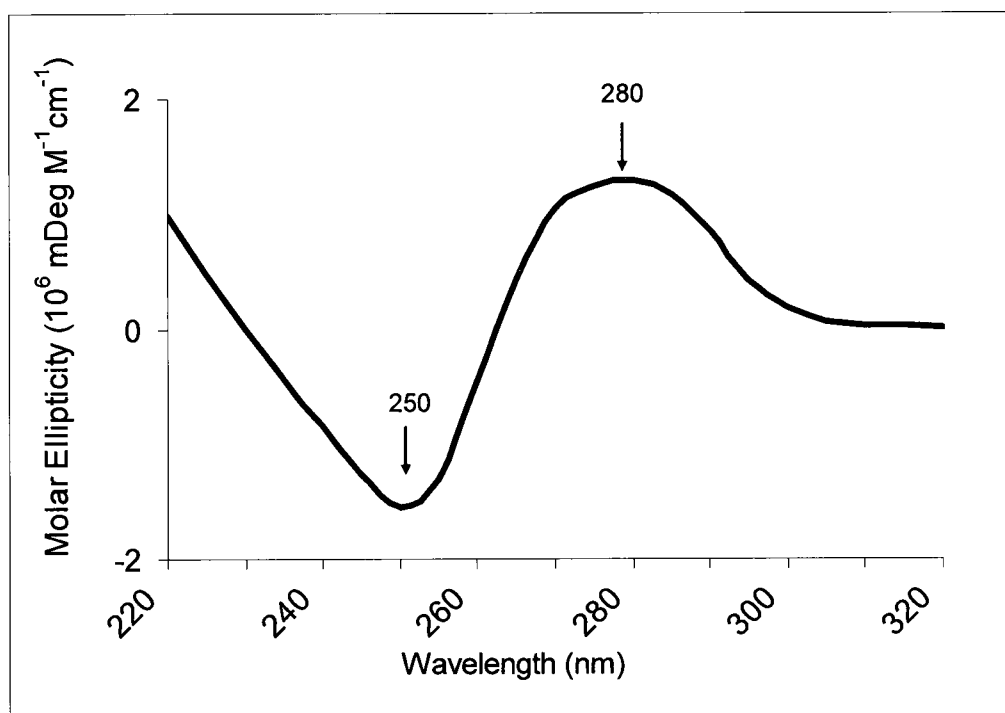


Figure 50: CD spectrum of DNA stem-loop in 150 mM K⁺ at 5°C.

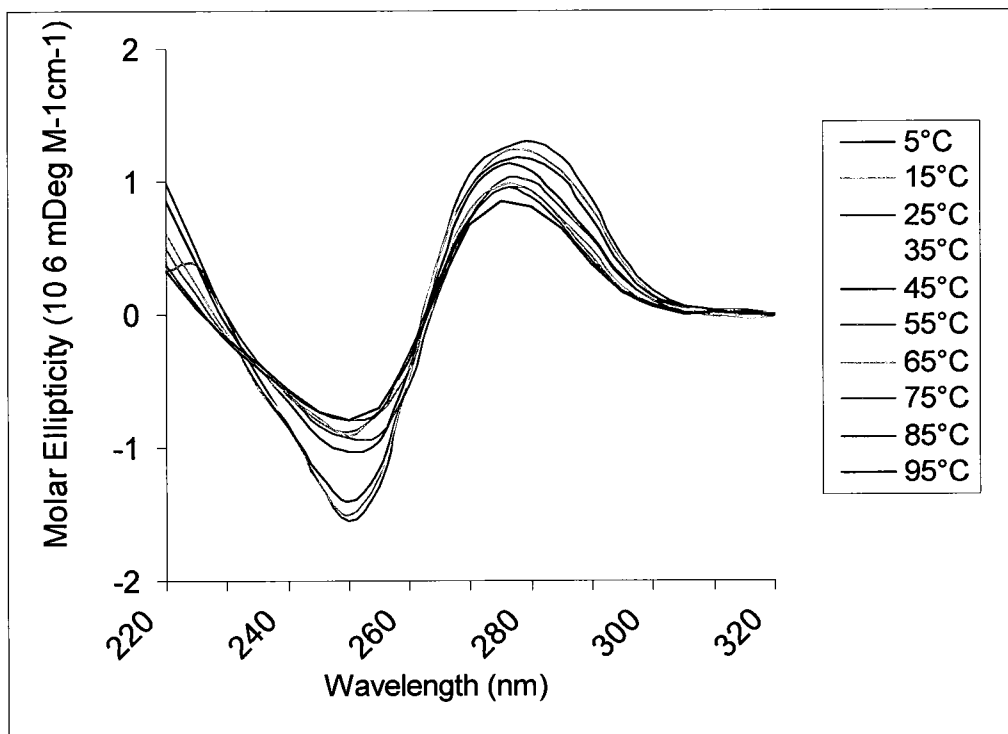


Figure 51: Temperature dependent CD spectra of DNA stem-loop in 15 mM K⁺

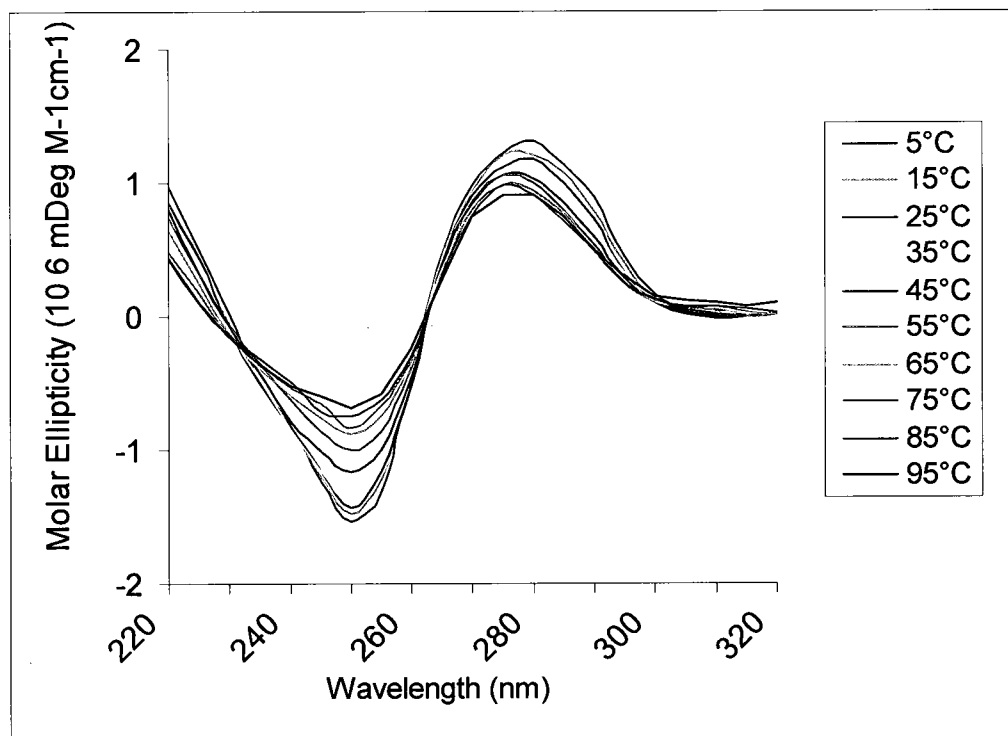


Figure 52: Temperature dependent CD spectra of DNA stem-loop in 50 mM K⁺

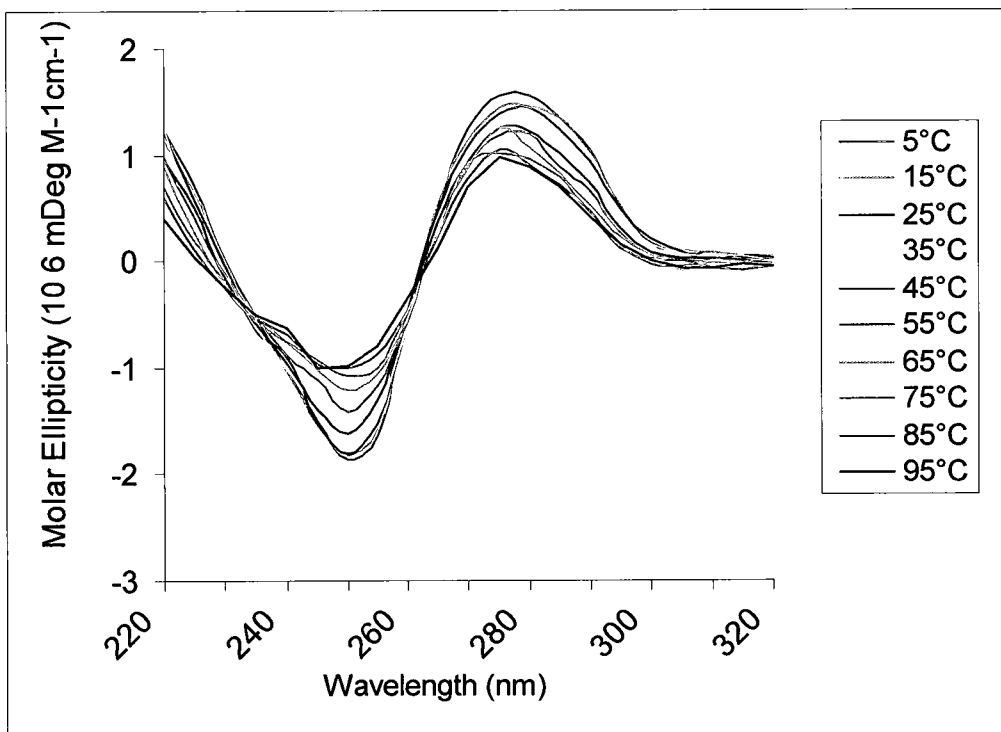


Figure 53: Temperature dependent CD spectra of DNA stem-loop in 100 mM K⁺

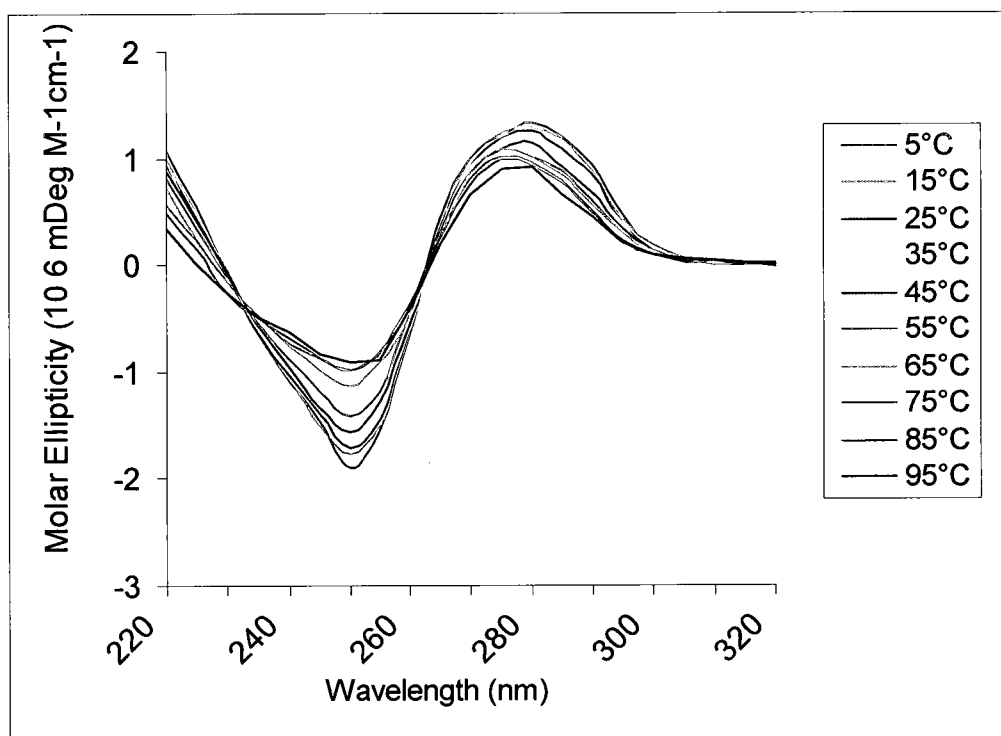


Figure 54: Temperature dependent CD spectra of DNA stem-loop in 250 mM K⁺

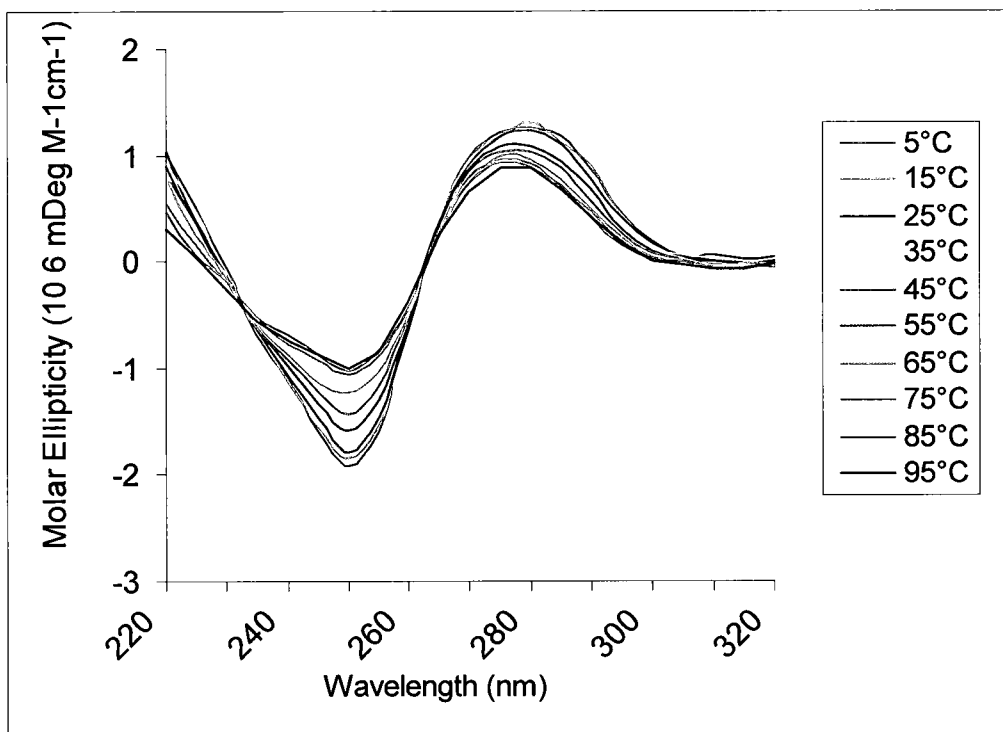


Figure 55: Temperature dependent CD spectra of DNA stem-loop in 500 mM K⁺

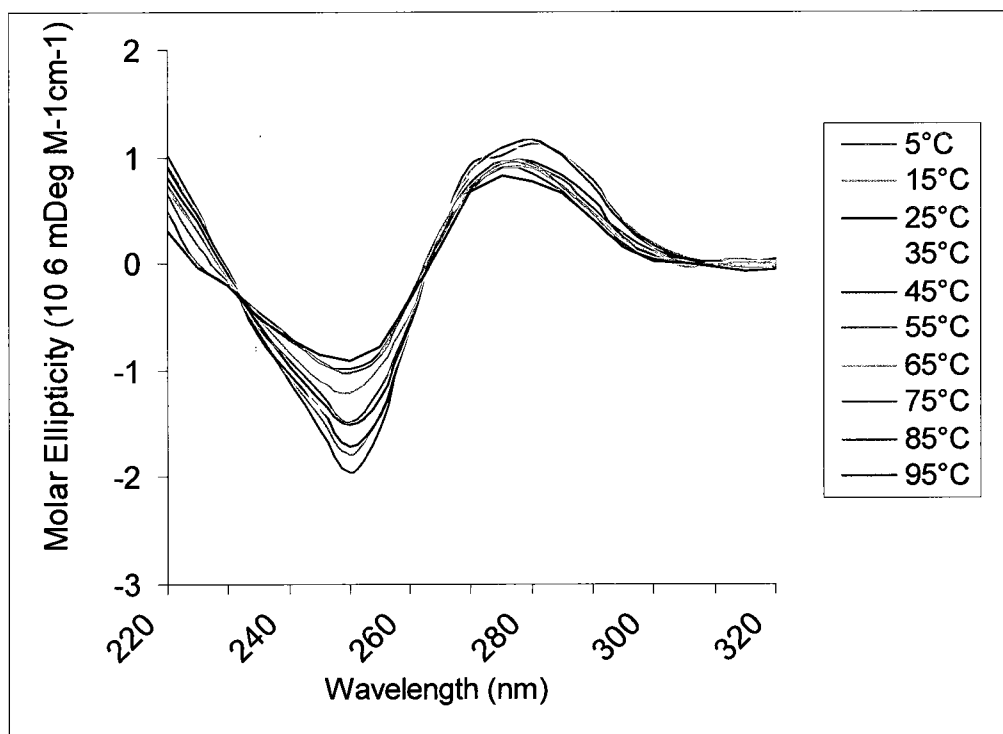


Figure 56: Temperature dependent CD spectra of DNA stem-loop in 1000 mM K⁺

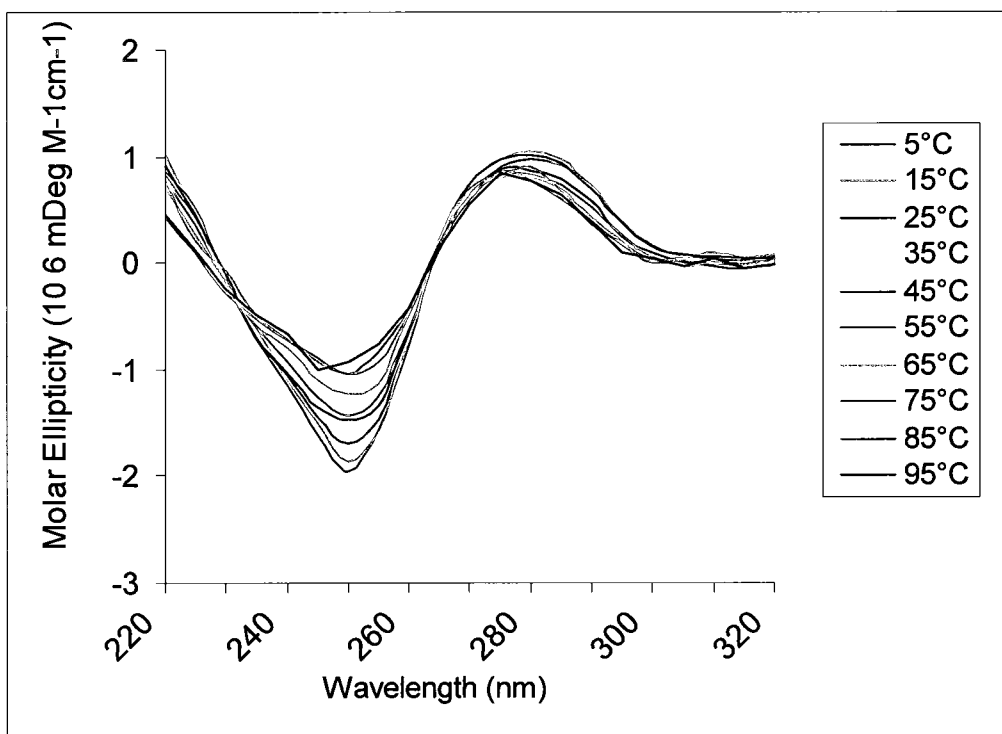


Figure 57: Temperature dependent CD spectra of DNA stem-loop in 2000 mM K⁺

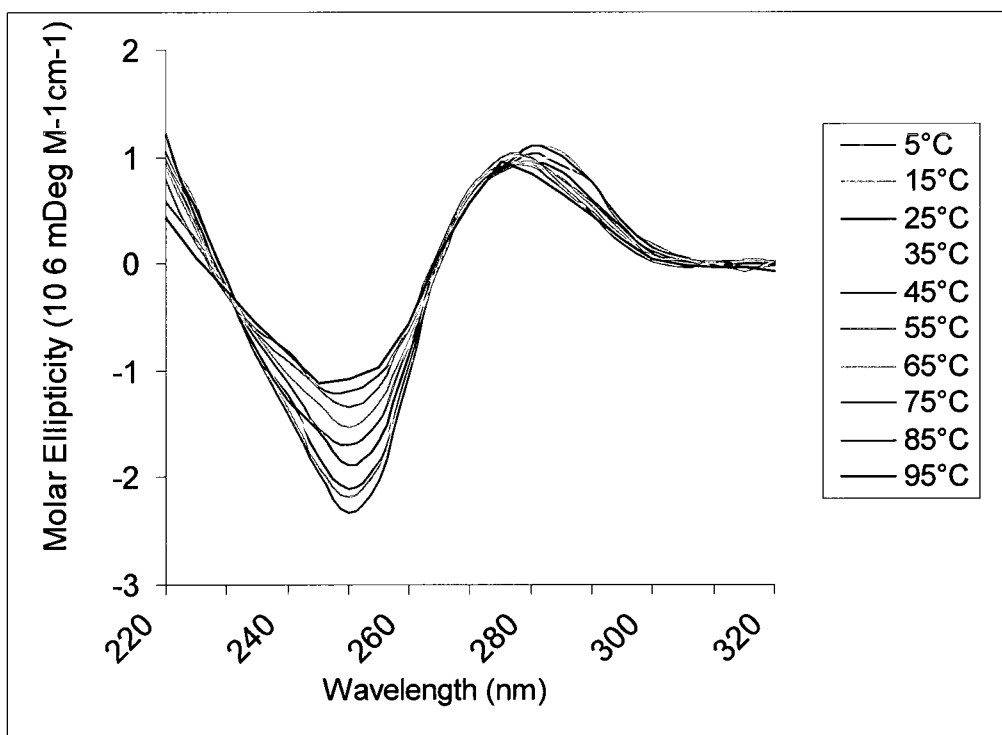


Figure 58: Temperature dependent CD spectra of DNA stem-loop in 3000 mM K⁺

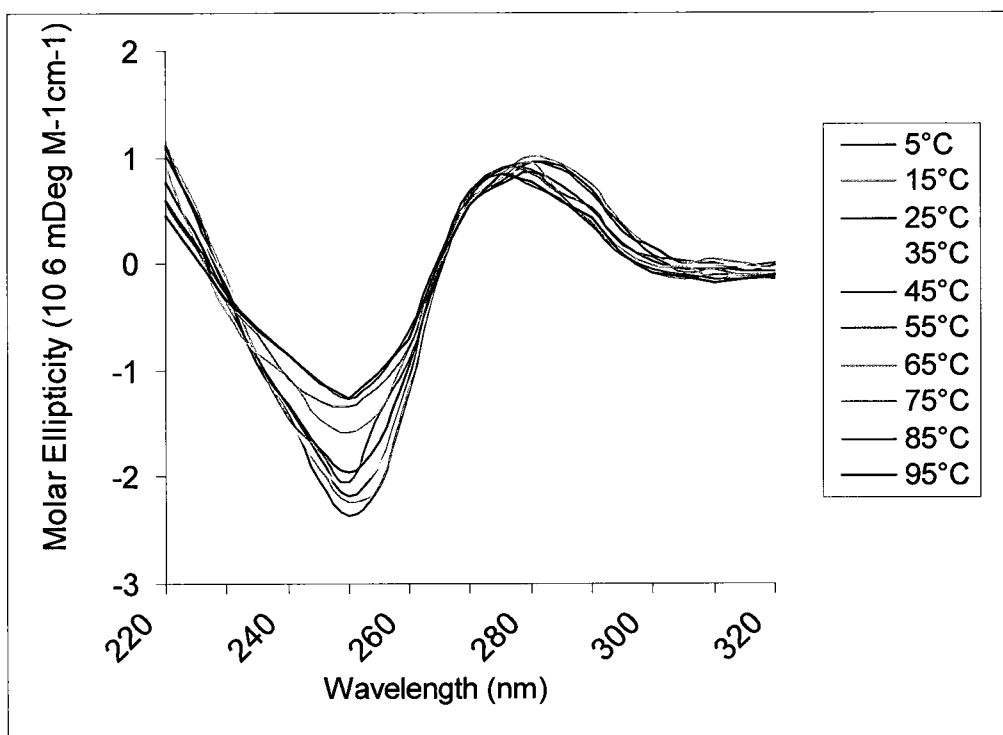


Figure 59: Temperature dependent CD spectra of DNA stem-loop in 4000 mM K⁺

An isosbestic point was observed throughout the aforementioned temperature dependent CD scans. An isosbestic point suggests a transition between two different molecular states or species. In this case, it could represent the transition from folded to less folded structure.

4.3.2.2 Salt dependent CD Spectra of DNA stem-loop model

The salt dependent spectra of the DNA stem-loop at various temperatures were extracted from the aforementioned data and presented in Figures 60 - 62.

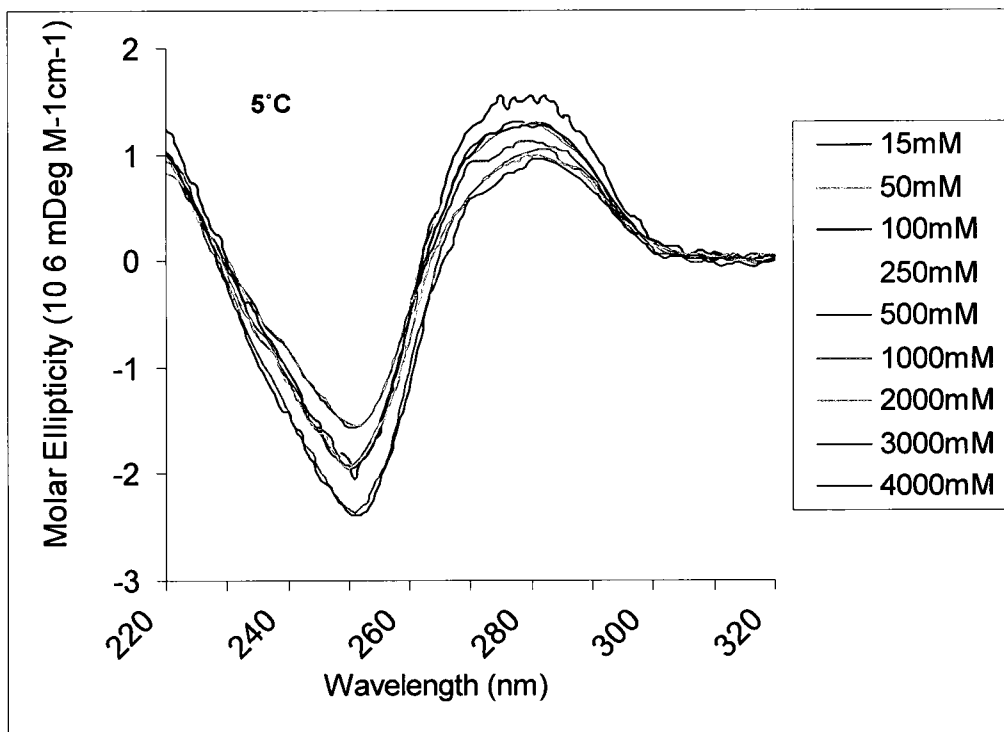


Figure 60: Salt dependent spectra of DNA stem-loop at 5°C

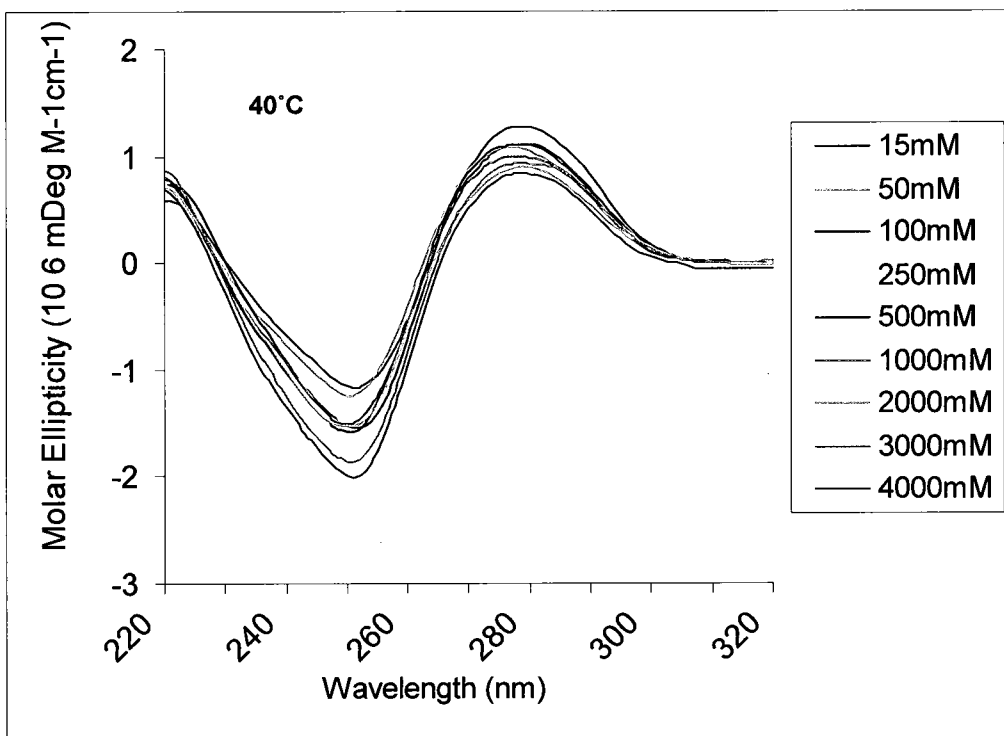


Figure 61: Salt dependent spectra of DNA stem-loop at 40°C

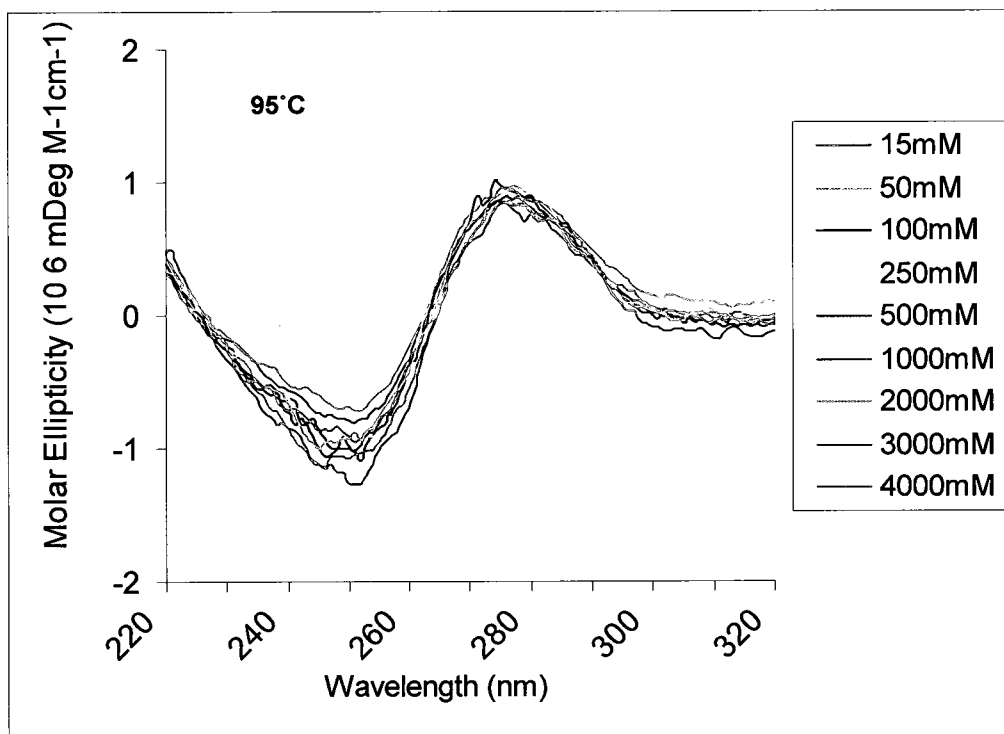


Figure 62: Salt dependent spectra of DNA stem-loop at 95°C

From the aforementioned salt dependent spectra of the DNA stem-loop model depicted in (Figures 60-62) it is clear that the spectra are salt dependent. Furthermore, it is important to mention that no clear isosbestic points were observed. This indicates that there is no transition between states or species as the salt concentration increases and the observed changes are related to alteration of conformation.

4.3.2.3 van't Hoff Analysis

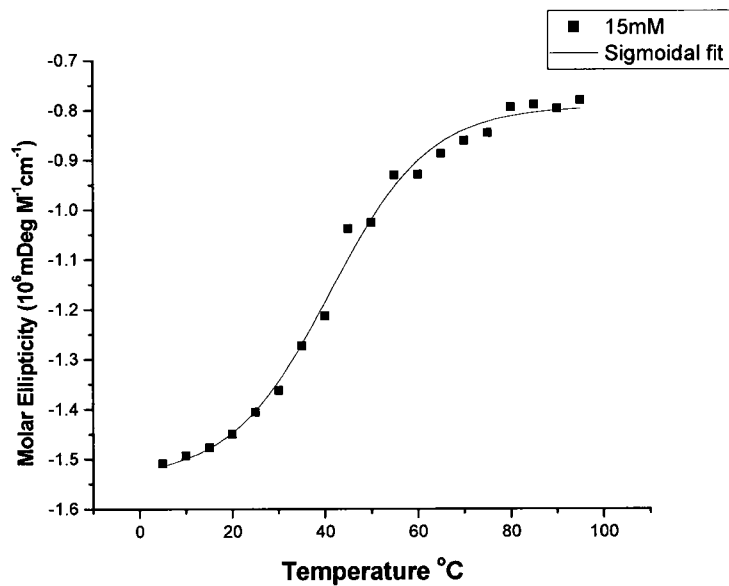
The collected data was subjected to van't Hoff analysis at various salt concentrations. The data was converted in terms of α (fraction folded DNA) and utilized to calculate K_F . The van't Hoff plot ($R\ln K$ vs $1/T$) was constructed yielding a slope and an intercept essential for van't Hoff enthalpy and entropy determination respectively.

Figures 63a – 71a provide the sigmoidal fit melting curve for the CD melts from Figures

51 - 59 while Figures 63b - 71b provide van't Hoff plot for the same melting curves.

Table 4 presents the van't Hoff data obtained data obtained for the various salt concentrations.

A.



B.

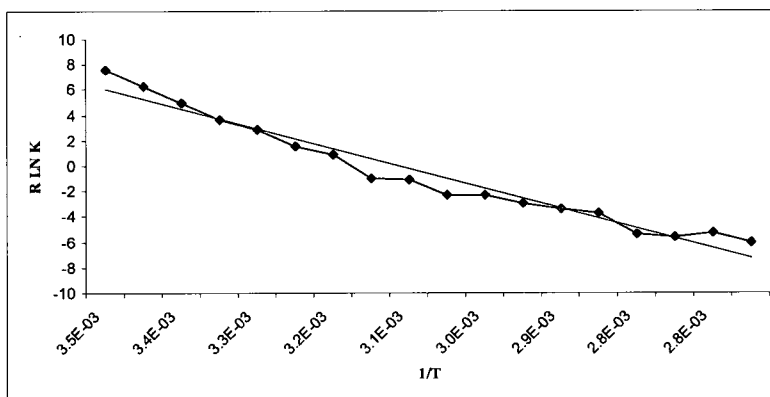
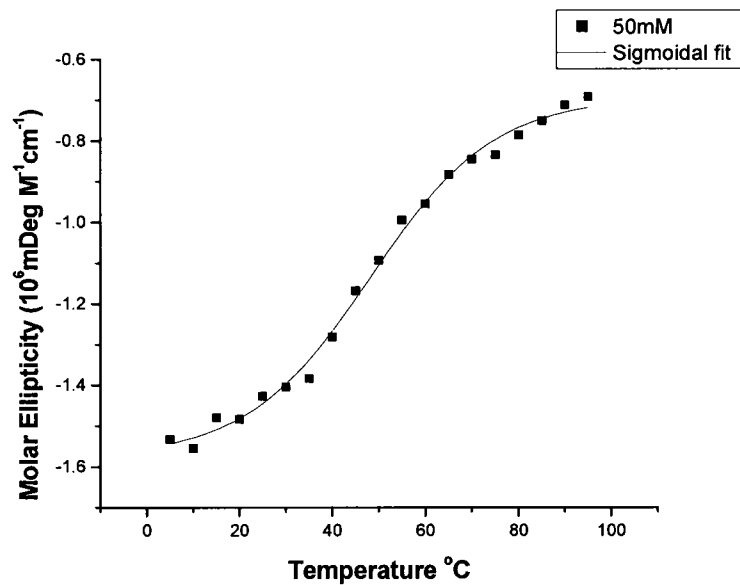


Figure 63: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 15 mM K^+ , (B) van't Hoff plot

A.



B.

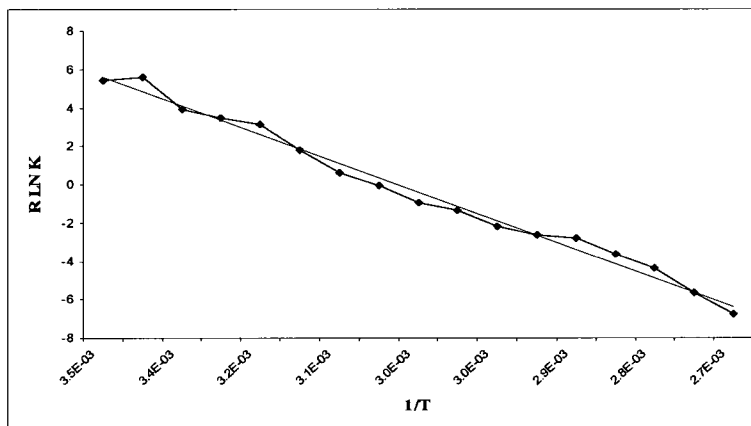
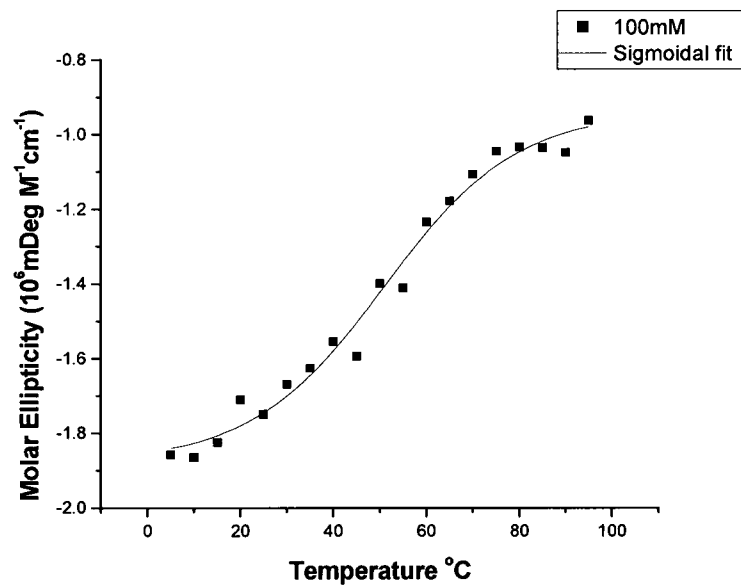


Figure 64: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 50 mM K^+ , (B) van't Hoff plot

A.



B.

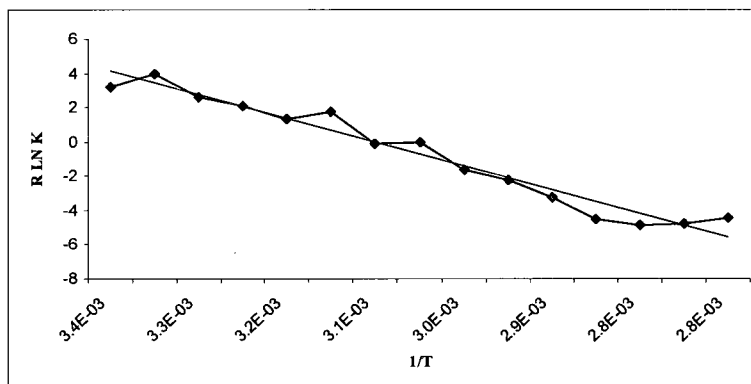
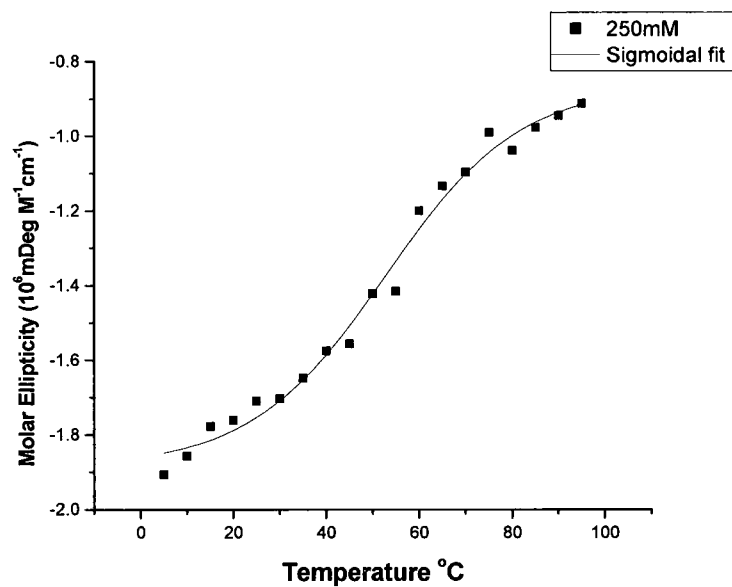


Figure 65: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 100 mM K⁺, (B) van't Hoff plot

A.



B.

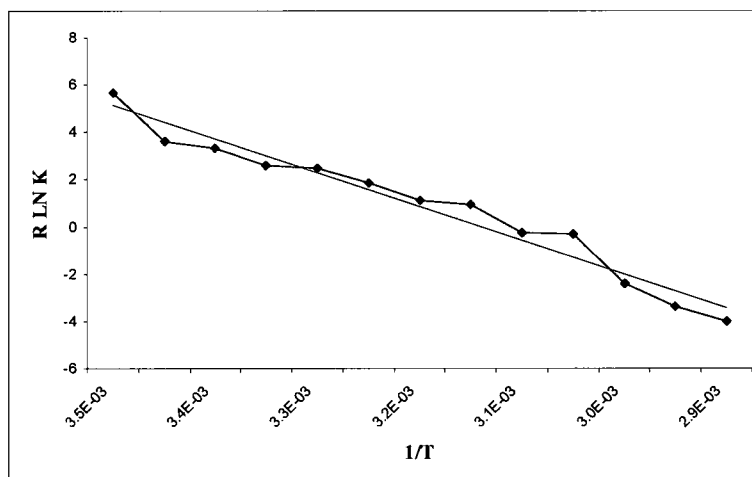
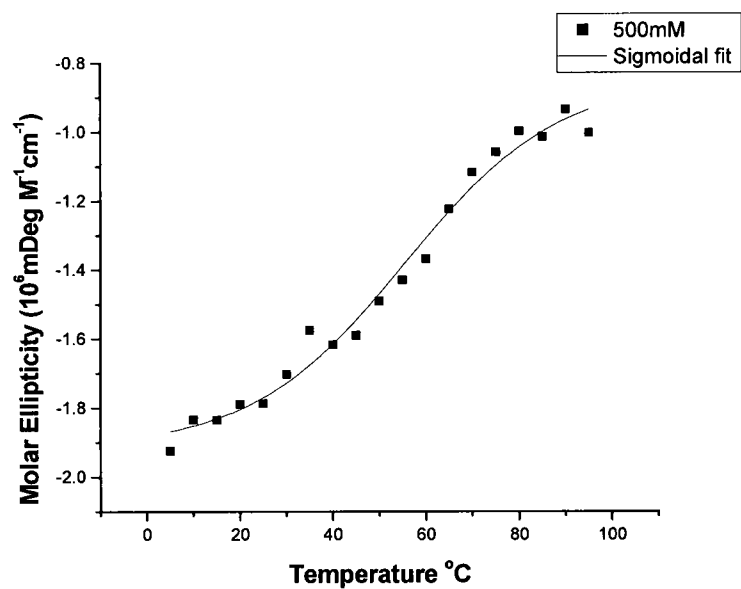


Figure 66: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 250 mM K^+ , (B) van't Hoff plot

A.



B.

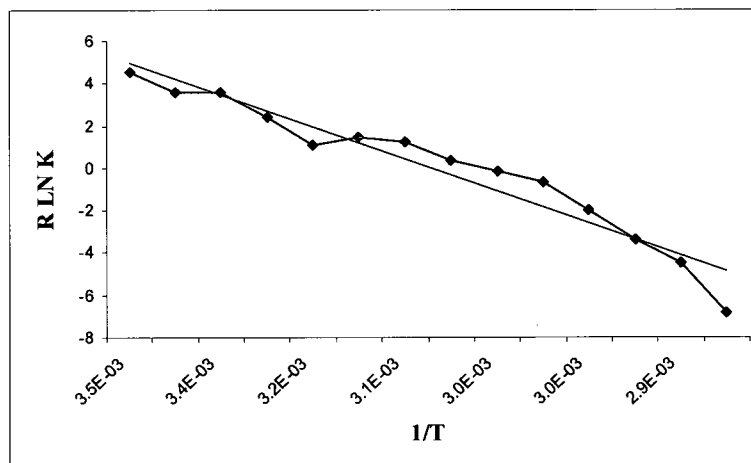
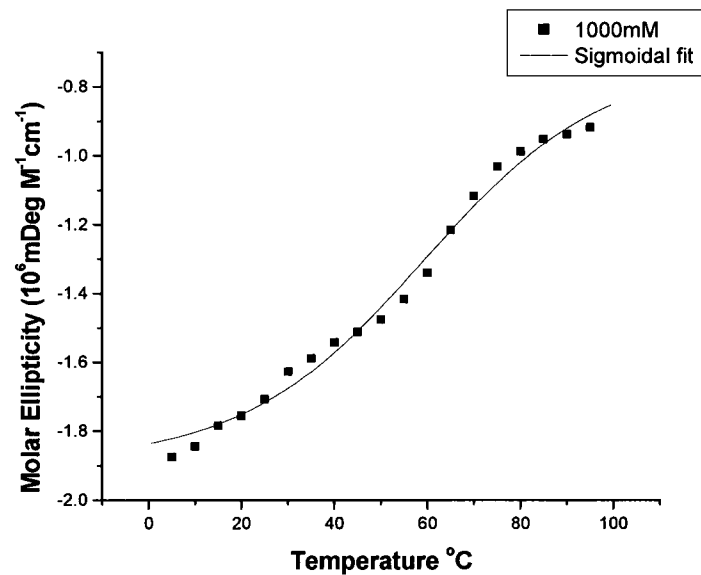


Figure 67: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 500 mM K^+ , (B) van't Hoff plot

A.



B.

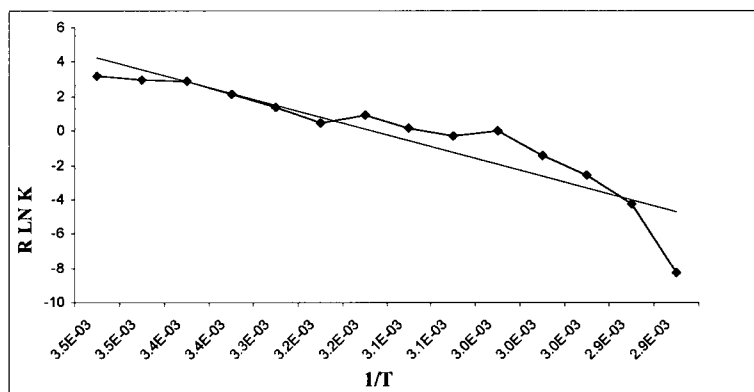
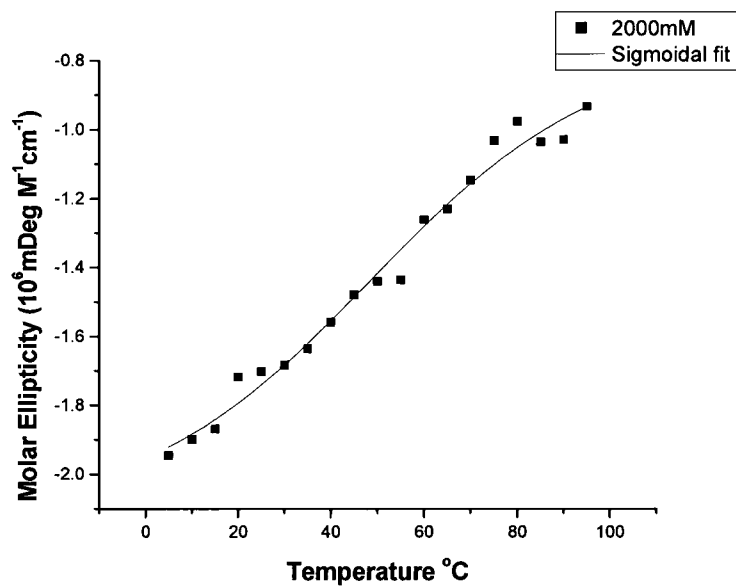


Figure 68: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 1000 mM K^+ , (B) van't Hoff plot

A.



B.

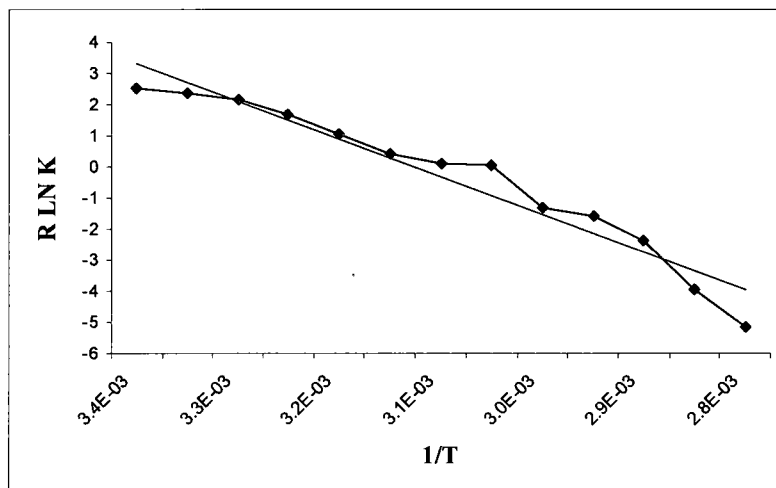
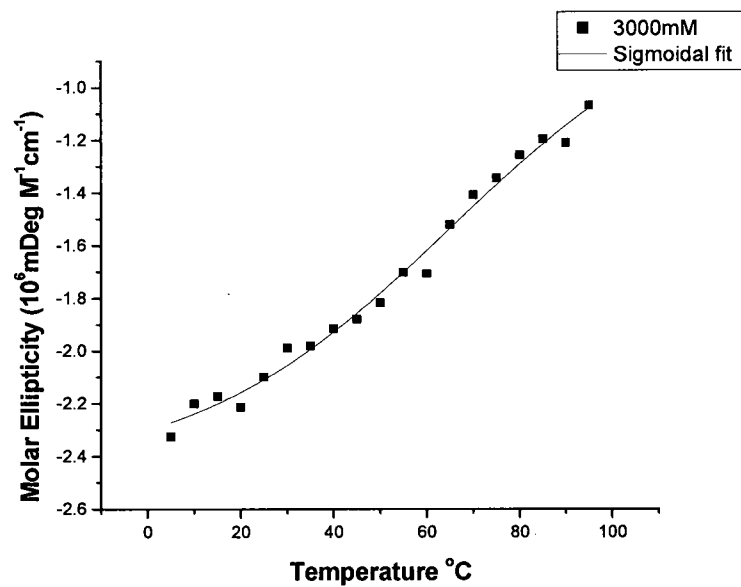


Figure 69: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 2000 mM K^+ , (B) van't Hoff plot

A.



B.

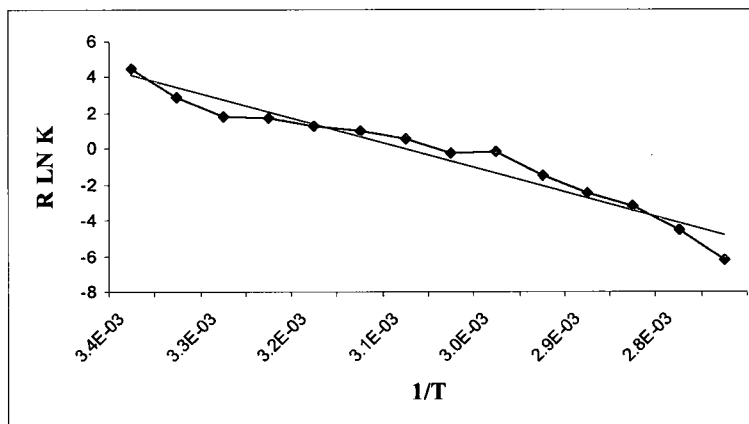
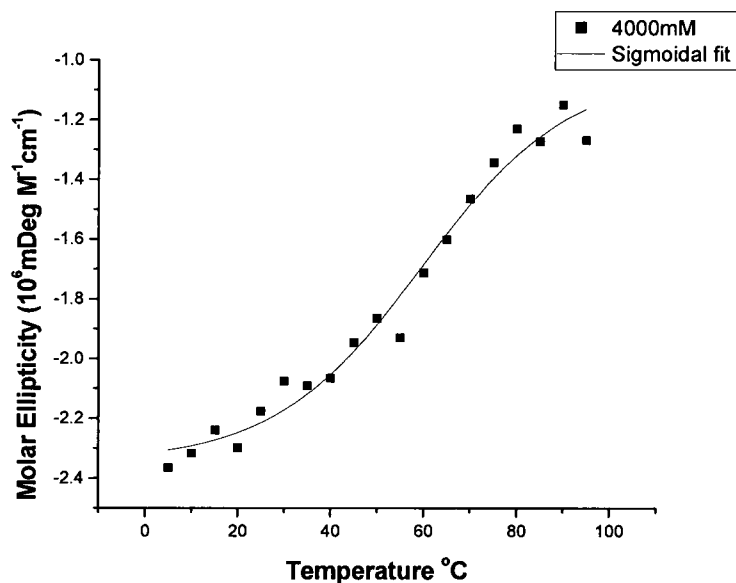


Figure 70: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 3000 mM K^+ , (B) van't Hoff plot

A.



B.

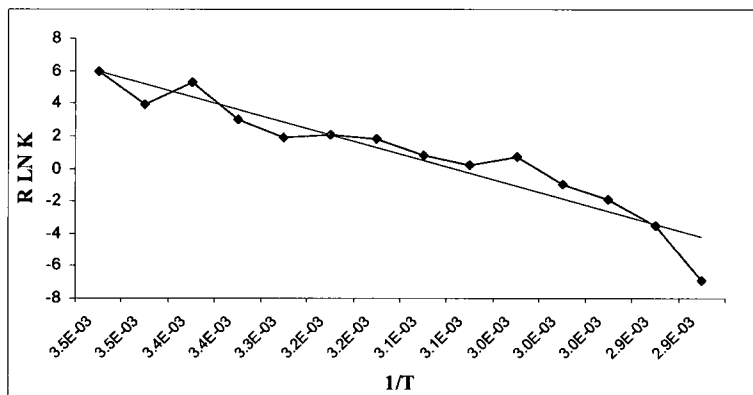


Figure 71: (A) Sigmoidal fit of the CD melting of the stem-loop DNA model in 4000 mM K^+ , (B) van't Hoff plot

Careful review of the aforementioned sigmoidal fit curves and van't Hoff plots (Figures 63 – 71) reveals several issues. The increased stability and the inability to define a sufficient upper baseline at the higher salt concentrations, makes it quite difficult to establish an apparent T_M values for these melts. The assignment of inappropriate T_M

values for the melts may contribute to errors in van't Hoff thermodynamic data. Systemic deviation from a simple sigmoidal shape curve was observed at most of the sigmoidal fit curves at all tested salt concentrations. Lastly, as van't Hoff analysis assumes two state transition, the results highlight clear deviation from that assumption. This is evident by the nonlinear van't Hoff plot obtained at all salt concentrations.

Table 4: van't Hoff thermodynamic parameters as a function of $[K^+]$

$[K^+]$	T_M (°C)	ΔH_{VH} (kcal mol ⁻¹)	ΔS_{VH} (kcal K ⁻¹ mol ⁻¹)	$\Delta G_{VH 42^\circ C}$ (kcal mol ⁻¹)
15 mM	41	-16.6	-52.0	-0.23
50 mM	49	-15.9	-49.2	-0.42
100 mM	56	-17.8	-55.6	-0.29
250 mM	60	-13.9	-43.7	-0.13
500 mM	64	-14.8	-46.4	-0.16
1000 mM	65	-13.3	-42.4	0.10
2000mM	63	-13.5	-42.1	-0.21
3000mM	64	-13.3	-41.4	-0.21
4000mM	64	-15.3	-47.7	-0.24

4.3.2.4 SVD

SVD is a tool for enumerating the number of species required to account for the spectral changes that accompany a transition. CD spectra were obtained at 95 different temperatures (5-99°C at 1°C increments) and collected in the wavelength range of 220-

320 nm, at 1 nm intervals. Therefore, the initial 101 x 95 data matrix A for SVD analysis was organized such that the rows consisted of data collected at 1 nm increments of wavelength (220-320 nm) and the columns at 1°C increments (5-95°C) for each set of data collected at a particular concentration of K^+ . The CD matrices were plotted in a 3D plot where X, Y and Z were ellipticity, wavelength and temperature respectively. (Figure 72)

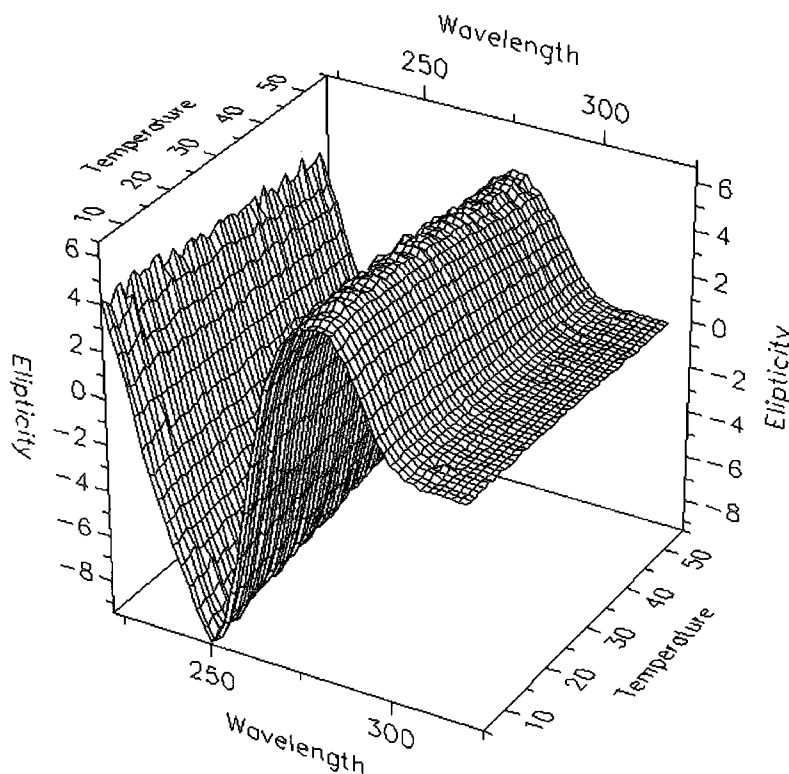


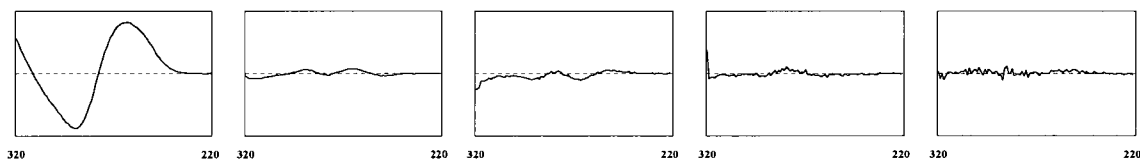
Figure 72: Representative 3D plot of the CD spectra

The constructed matrices were then subjected to singular-value decomposition (SVD) analysis to evaluate the number of significant spectral species during the transition. Two different software were used to perform the singular-value decomposition analysis, Olis global works and Matlab 6.5.0.1 (Mathworks).

4.3.2.4.1 SVD analysis using *Olis global works*

The magnitude of the singular values and the randomness of the corresponding spectral and kinetic eigenvectors were used to determine the number of possible spectral and kinetic eigenvectors were used to determine the number of possible significant species. The spectral and kinetic eigenvectors plots were assessed visually for randomness (Figures 73 - 81). From the aforementioned plots it is clear that three spectral species are required to describe the temperature-dependent spectral transitions. This is evident by the non-randomness of the spectral and kinetic eigenvectors plots for components one through three and the randomness of the spectral and kinetic eigenvectors plots for components four and five.

A. Spectral Eigenvector



B. Kinetic Eigenvector

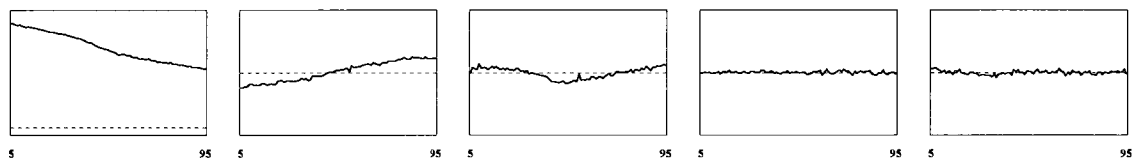
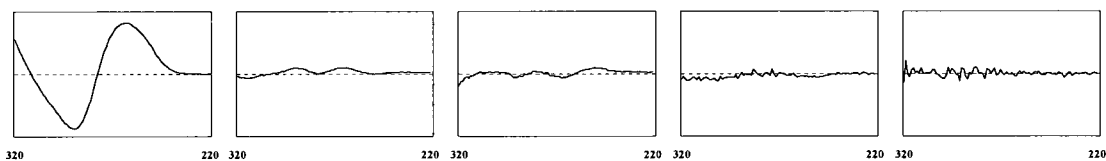


Figure 73: (A) Spectral Eigenvector of CD melting of DNA in 15mM K⁺, (B) Kinetic Eigenvector

A. Spectral Eigenvector



B. Kinetic Eigenvector

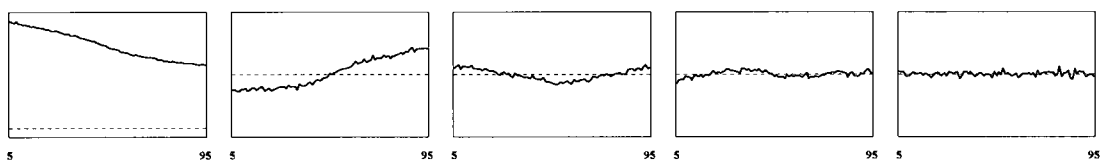
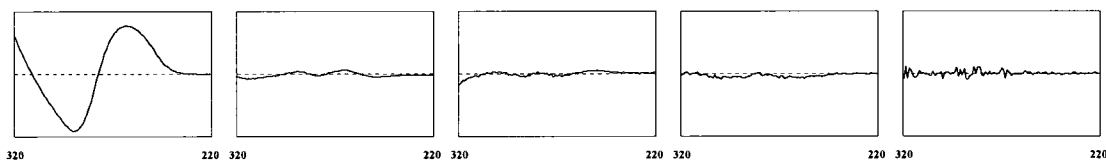


Figure 74: (A) Spectral Eigenvector of CD melting of DNA in 50mM K⁺, (B) Kinetic Eigenvector

A. Spectral Eigenvector



B. Kinetic Eigenvector

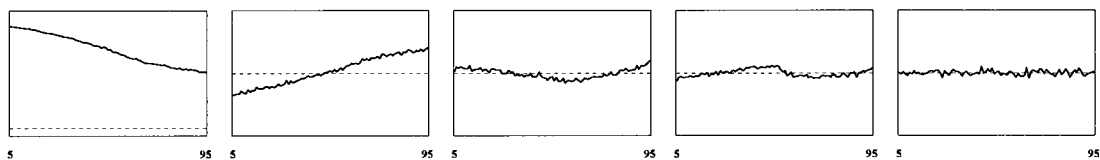
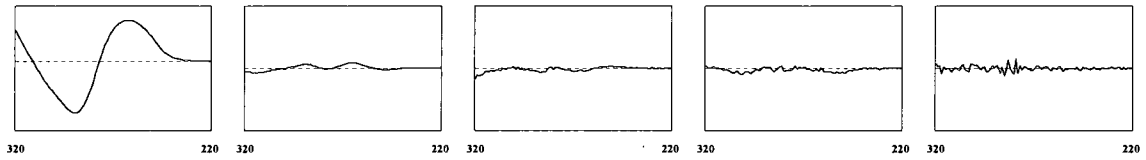


Figure 75: (A) Spectral Eigenvector of CD melting of DNA in 100mM K⁺, (B) Kinetic Eigenvector

A. Spectral Eigenvector



B. Kinetic Eigenvector

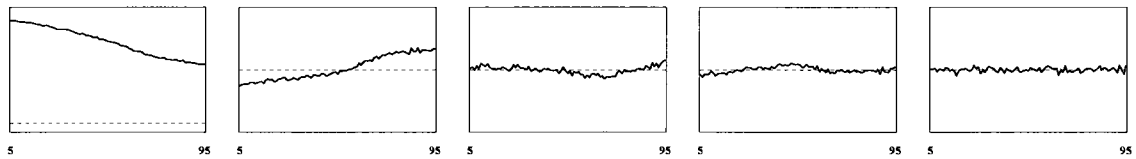
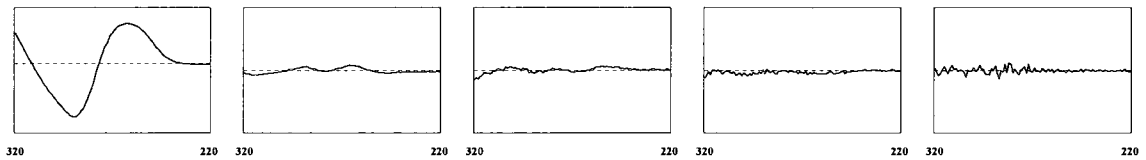


Figure 76: (A) Spectral Eigenvector of CD melting of DNA in 250mM K^+ , (B) Kinetic Eigenvector

A. Spectral Eigenvector



B. Kinetic Eigenvector

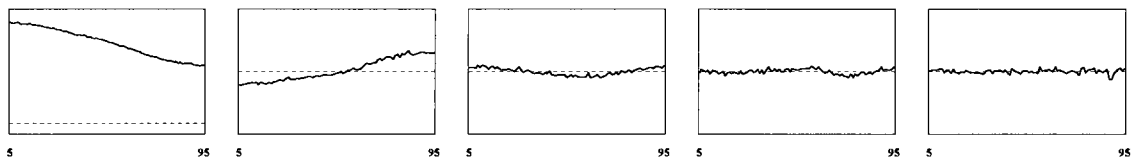
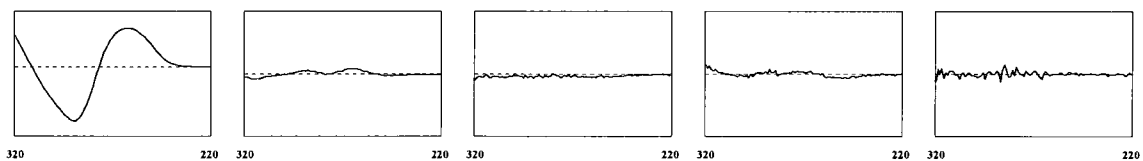


Figure 77: (A) Spectral Eigenvector of CD melting of DNA in 500mM K^+ , (B) Kinetic Eigenvector

A. Spectral Eigenvector



B. Kinetic Eigenvector

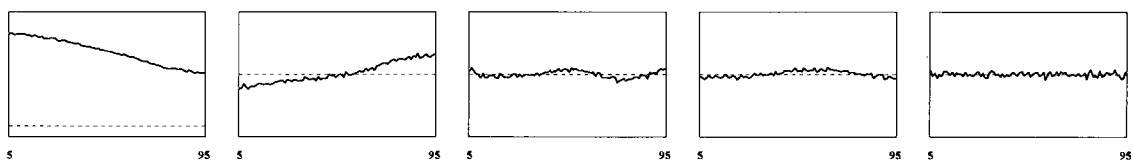
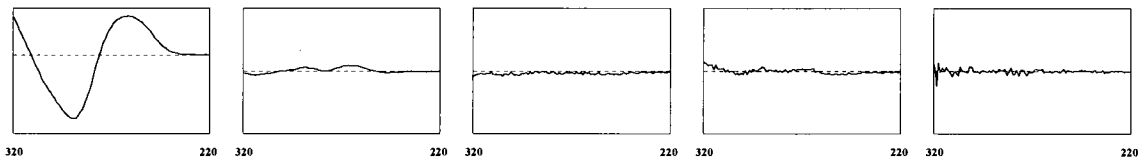


Figure 78: (A) Spectral Eigenvector of CD melting of DNA in 1000mM K⁺, (B) Kinetic Eigenvector

A. Spectral Eigenvector



B. Kinetic Eigenvector

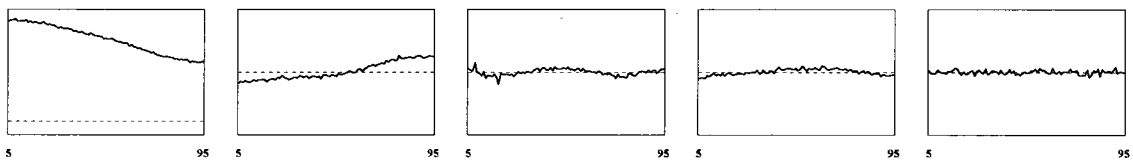
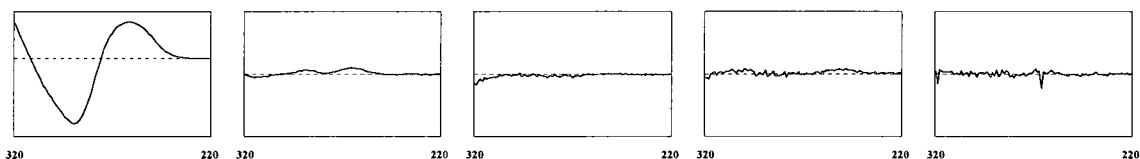


Figure 79: (A) Spectral Eigenvector of CD melting of DNA in 2000mM K⁺, (B) Kinetic Eigenvector

A. Spectral Eigenvector



B. Kinetic Eigenvector

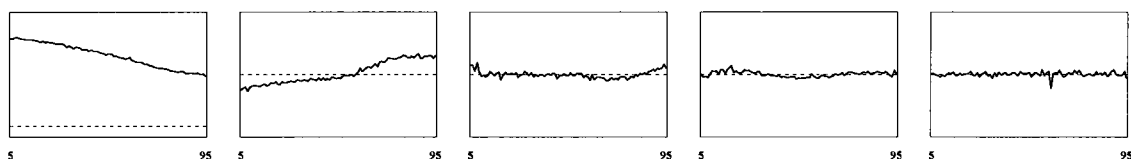
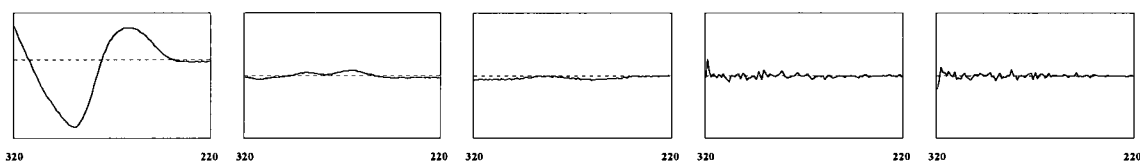


Figure 80: (A) Spectral Eigenvector of CD melting of DNA in 3000mM K⁺, (B) Kinetic Eigenvector

A. Spectral Eigenvector



B. Kinetic Eigenvector

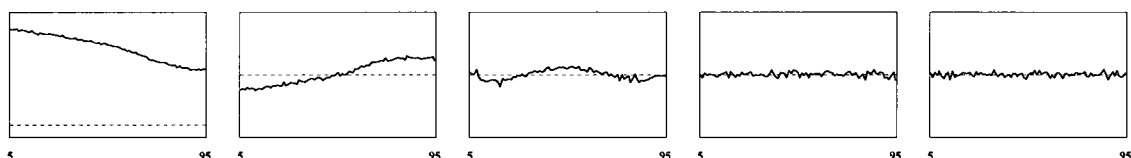


Figure 81: (A) Spectral Eigenvector of CD melting of DNA in 4000mM K⁺, (B) Kinetic Eigenvector

4.3.2.4.2 SVD analysis using Matlab

The S matrix was 95 x 95 and contained the singular values along its diagonal. The magnitudes of the singular values obtained from the S diagonal matrix were plotted as a function of the component number to obtain the residual plot (Figure 82). The

magnitudes of the singular values suggest that three major species are present during the melt of the stem-loop.

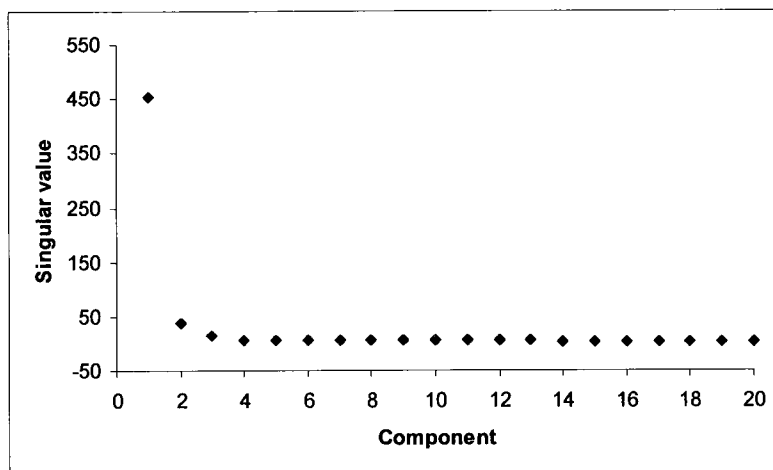


Figure 82: A representative residual plot (singular values magnitude as a function of component number) of CD melting of DNA in 50mM K⁺

The amplitude vectors (V coefficients) are directly proportional to the concentration of any given component. Figure 83 provide a representative plot of V coefficient for each significant component as a function of temperature for CD melting of DNA in 50mM K⁺.

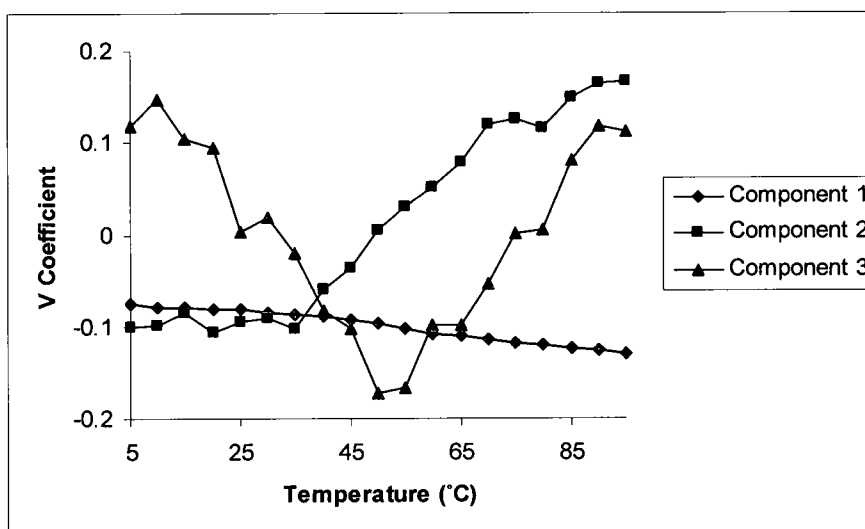


Figure 83: V coefficients for each significant component as a function of temperature of CD melting of DNA in 50mM K⁺

Residual contour plots for matrices with one through five singular values were assessed visually for randomness (Figures 84 - 88). From the aforementioned figures it is clear that residual contour plots for matrices generated with one, two and three singular values shows a degree of non-randomness. While, residual contour plots for matrices generated with four and five singular values are random. This could be interpreted as an indication of the possible presence of three major species during the melt of the stem-loop, namely components one, two and three.

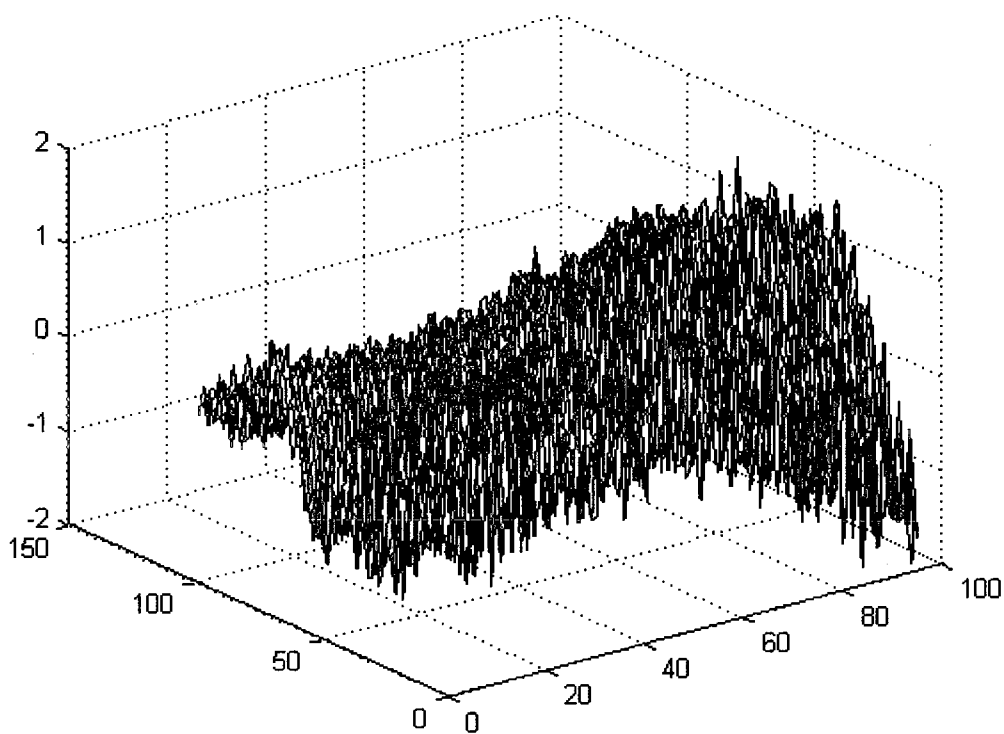


Figure 84: Residual contour plots for matrices generated with one singular value of CD melting of DNA in 50mM K⁺

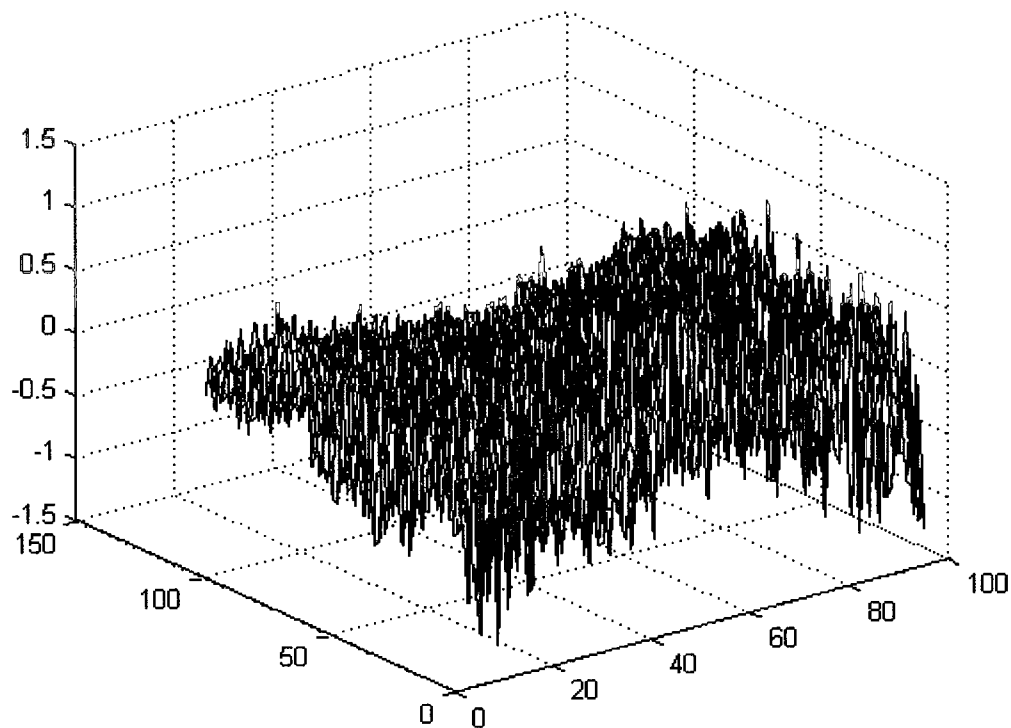


Figure 85: Residual contour plots for matrices generated with two singular value of CD melting of DNA in 50mM K⁺

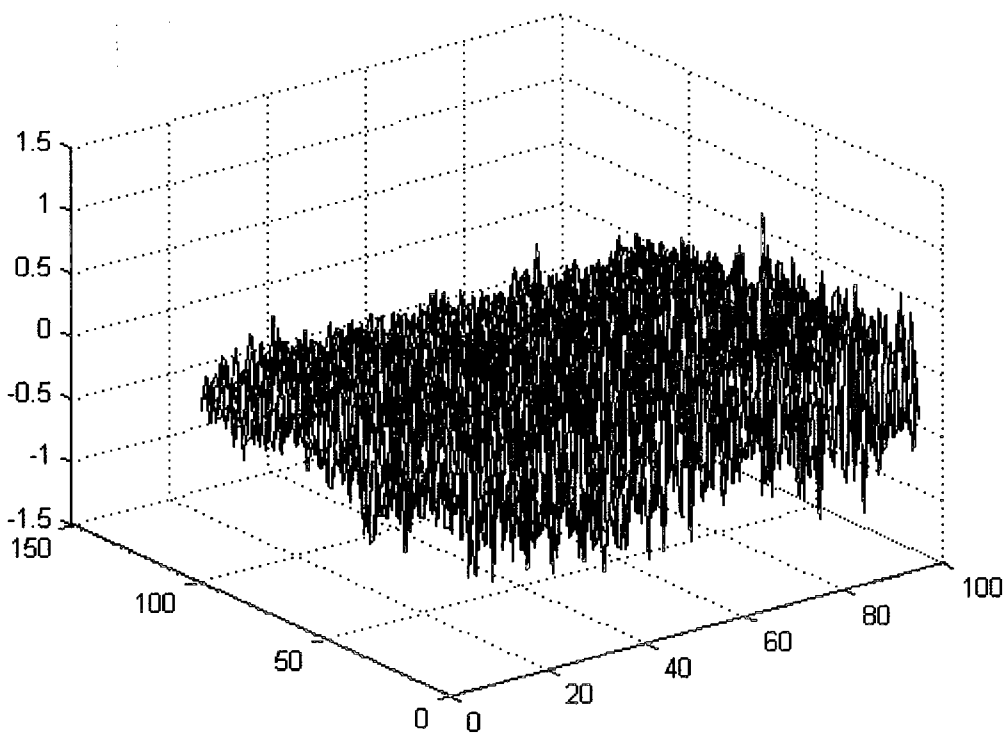


Figure 86: Residual contour plots for matrices generated with three singular values of CD melting of DNA in 50mM K⁺

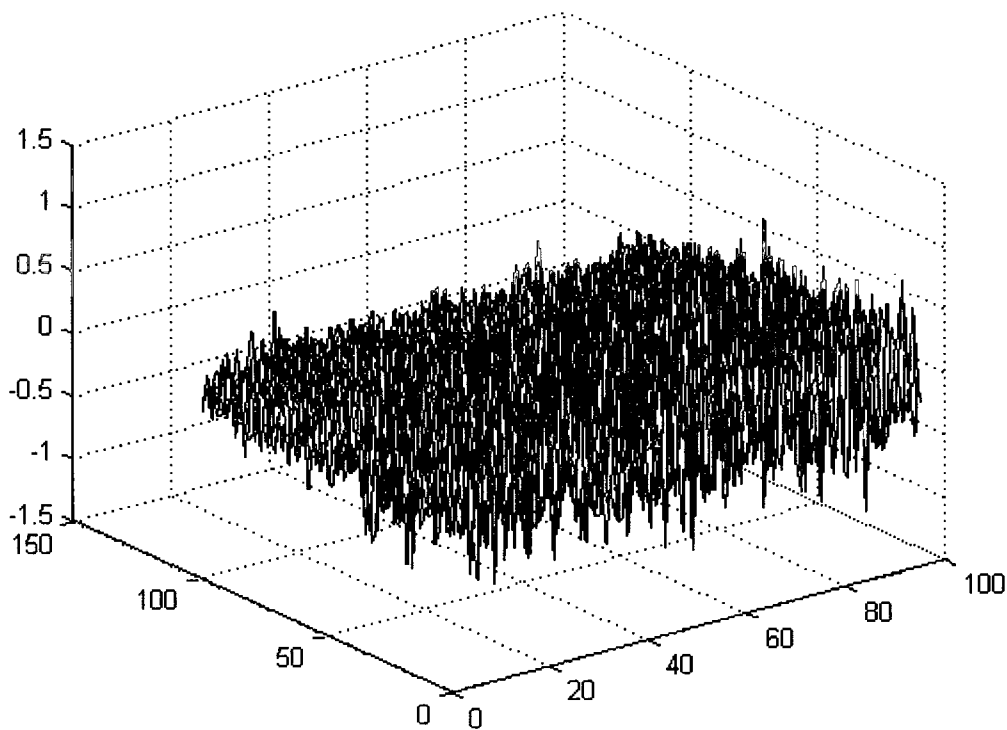


Figure 87: Residual contour plots for matrices generated with four singular values of CD melting of DNA in 50mM K⁺

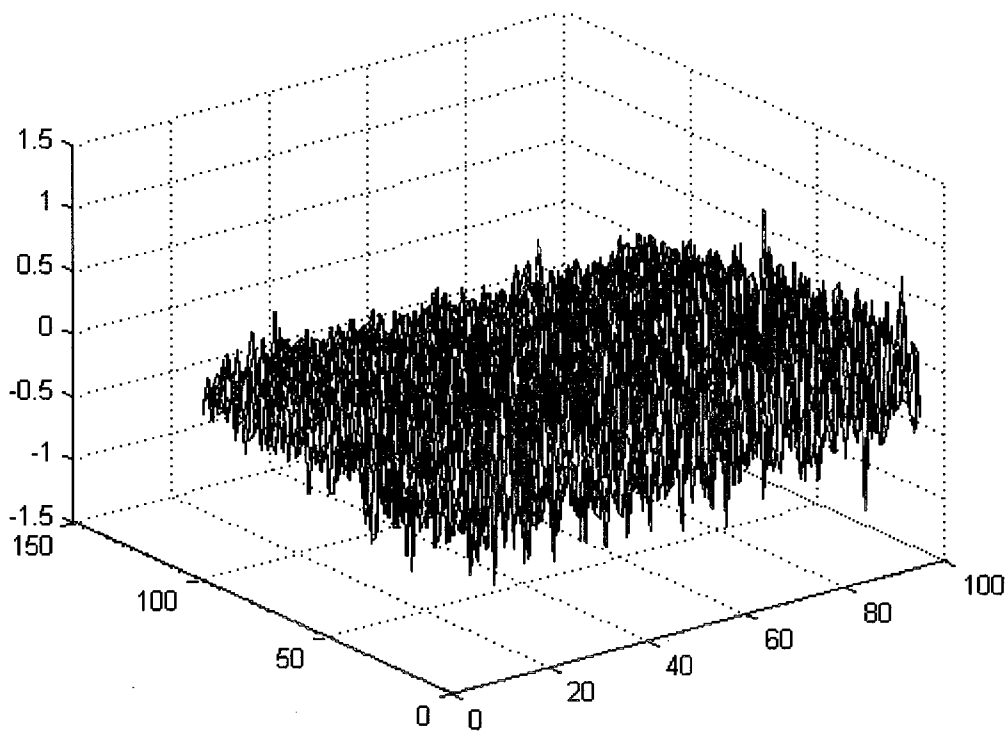


Figure 88: Residual contour plots for matrices generated with five singular values of CD melting of DNA in 50mM K⁺

Upon performing SVD analysis to all the temperature dependent CD scans, the obtained singular values and residual plots clearly points to three species at all salt concentrations.

4.3.3 Differential Scanning Calorimetry

DSC provides model independent approach to experimentally assess enthalpy as the heat capacity is directly measured as a function of temperature. Total of six DNA stem-loop melting DSC scans (sample vs reference, where the reference is the respective working KPBS buffer concentration) alternating between heating and cooling scans ramped between 5°C and 95°C were collected. The potassium ions concentrations utilized were 15, 50, 100, 250, 500, 1000, 2000, 3000 and 4000 mM K⁺. The data points were collected every 0.25°C with an associated temperature ramp rate of 0.25°C min⁻¹. Additionally, similar blank scans were collected (buffer vs reference, where the reference is the respective working KPBS buffer concentration). Figures 89 and 90 provide a representative repetitive DSC thermal melts for buffer-buffer and sample-buffer scans in 100 mM potassium ions respectively. The thermal melt process appears to be a reversible process at all salt concentration tested.

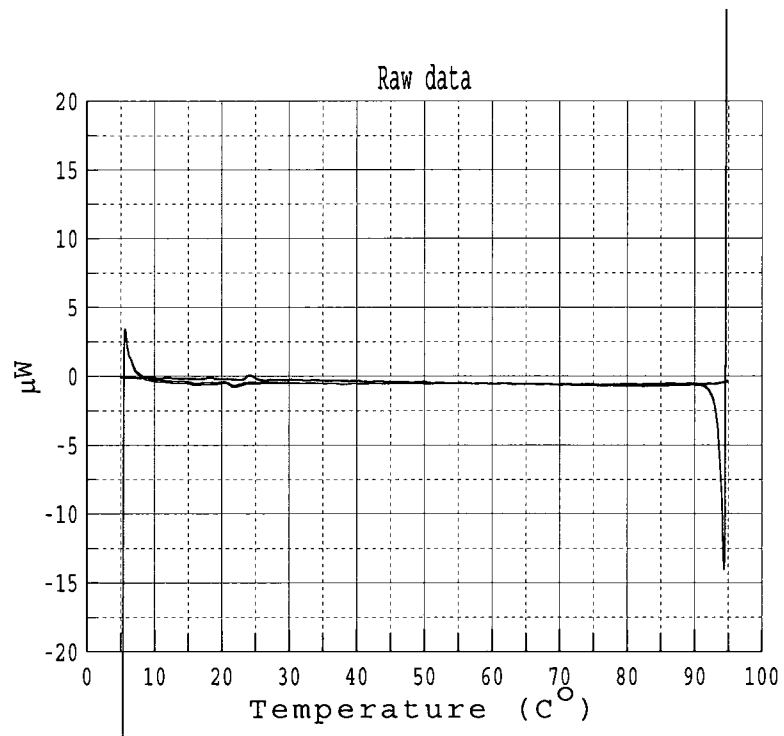


Figure 89: A representative repetitive DSC thermal melting scans (buffer-buffer) in 100 mM K⁺

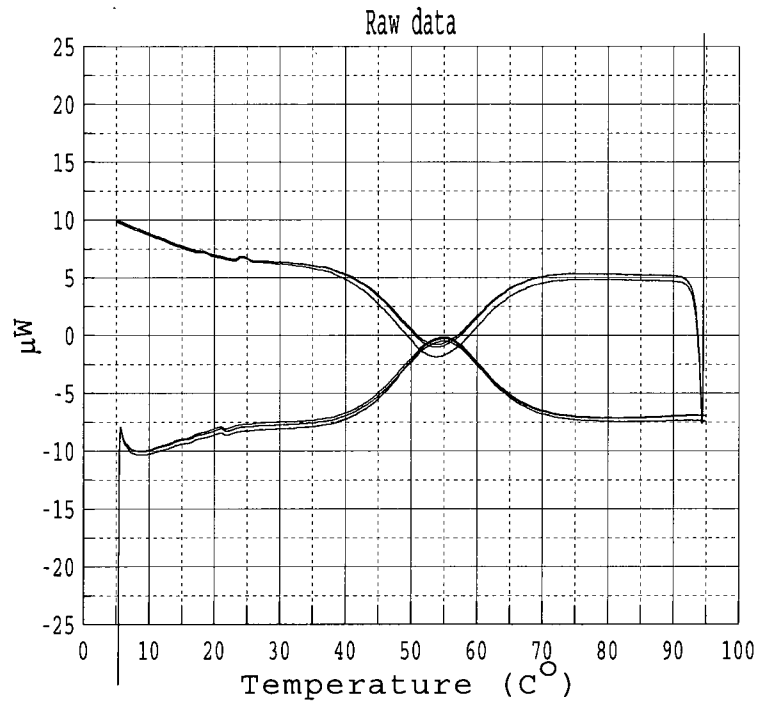


Figure 90: A representative repetitive DSC thermal melting scans (sample-buffer) in 100 mM K⁺

DSC thermal melts for the buffer-buffer were subtracted from the sample-buffer scans producing the molar heat capacity curve when plotted versus temperature (Figure 91).

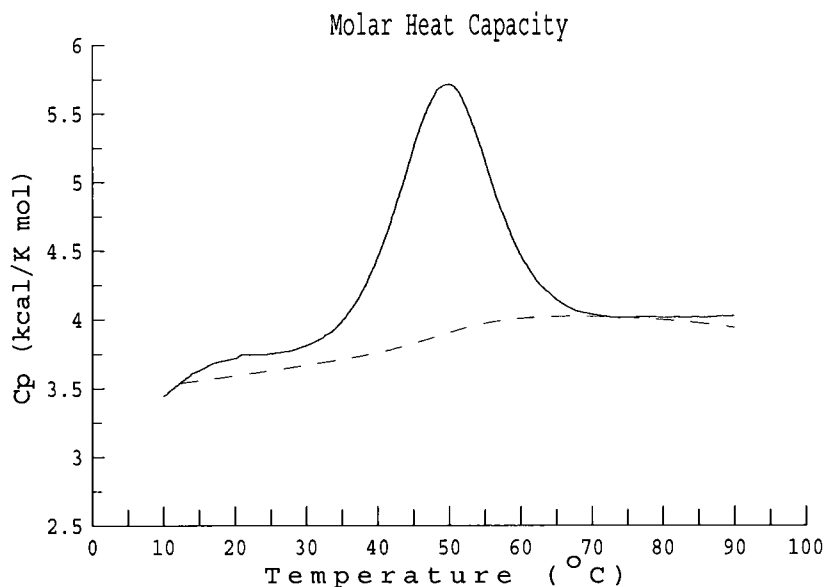


Figure 91: A representative molar heat capacity versus temperature plot in 100 mM K⁺ with a polynomial-linear baseline treatment

We can measure the latent heat of melting by measuring the area of the resulting peak.

We usually take the temperature at the apex of the peak as the melting temperature T_M where 50% of the DNA stem-loop is unfolded (Figure 92).

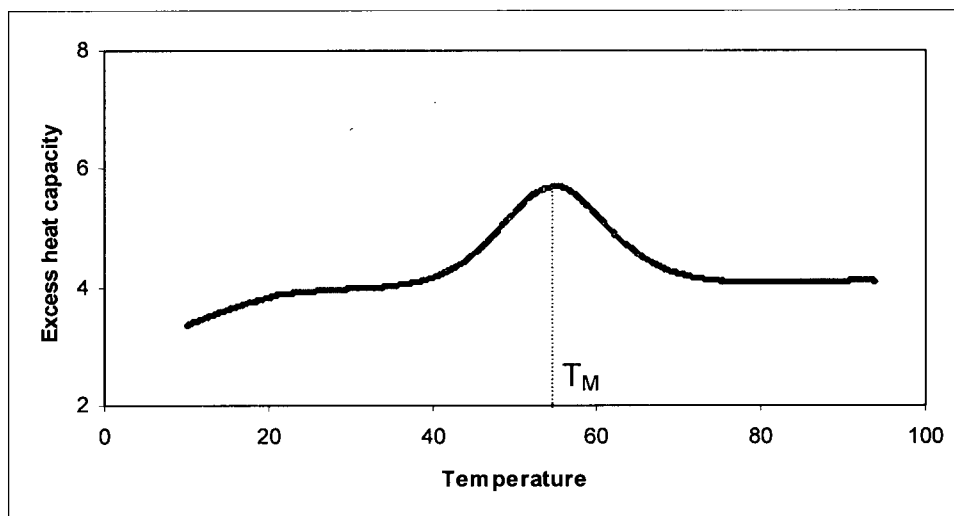


Figure 92: A representative DSC thermal plot in 100 mM K⁺

Figure 93 provides an overlay of the DSC thermoscans at all the various salt concentrations tested. The thermograms clearly indicate a salt dependence thermal stability that is evident by the increase in the thermal transition midpoint (T_M) as the salt concentration increase. Furthermore, all the collected DSC thermograms at all salt concentration clearly exhibit more than a single transition.

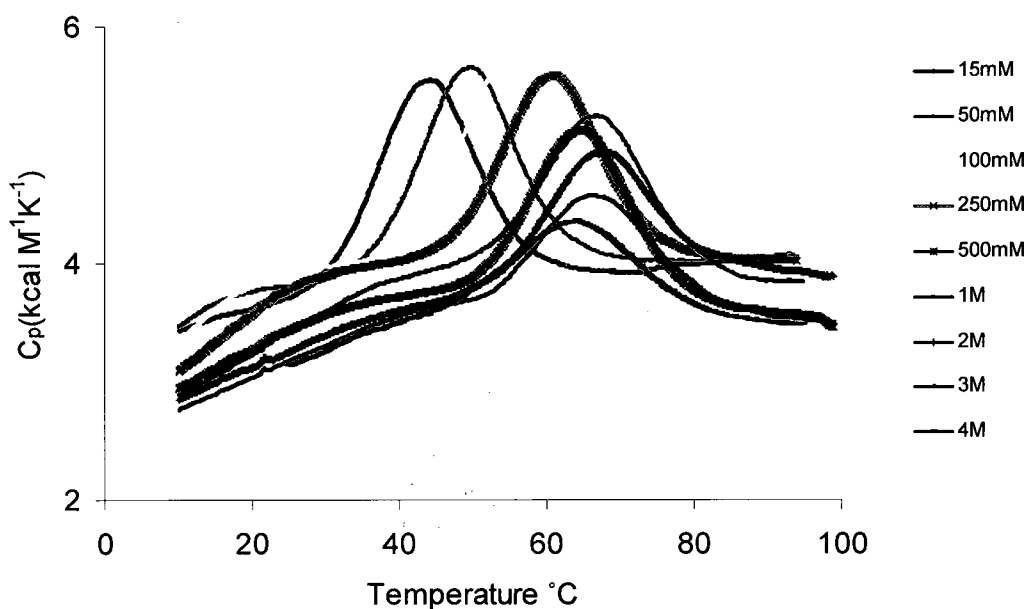


Figure 93: Baseline corrected DSC thermograms at various potassium ions concentration

Upon buffer subtraction and baseline correction, a multi Gaussian fit was applied to the DSC data using Origin 6.1. The data were best described into two distinctive Gaussian peaks indicating the presence of two distinct transitions T1 and T2 (Figure 94). The area under each Gaussian peak represents the enthalpy of the corresponding transition, while the apex of the peak represents the T_M for that transition. Similarly, the entropy of each transition could be calculated by plotting C_p/T versus T. Fit of the resultant plot with a multi Gaussian and determination of the area under each

Gaussian peak represents the entropy of the corresponding transition, while the apex of the peak represents the T_M for that transition (Figure 95).

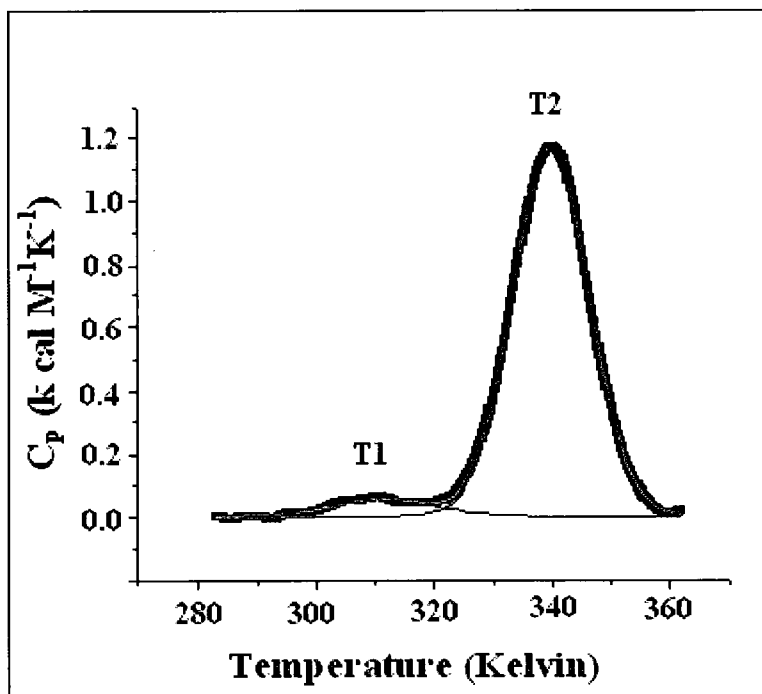


Figure 94: Deconvolution of a representative DSC thermogram obtained at 1M K^+ into two transitions to calculate calorimetric enthalpy for each transition

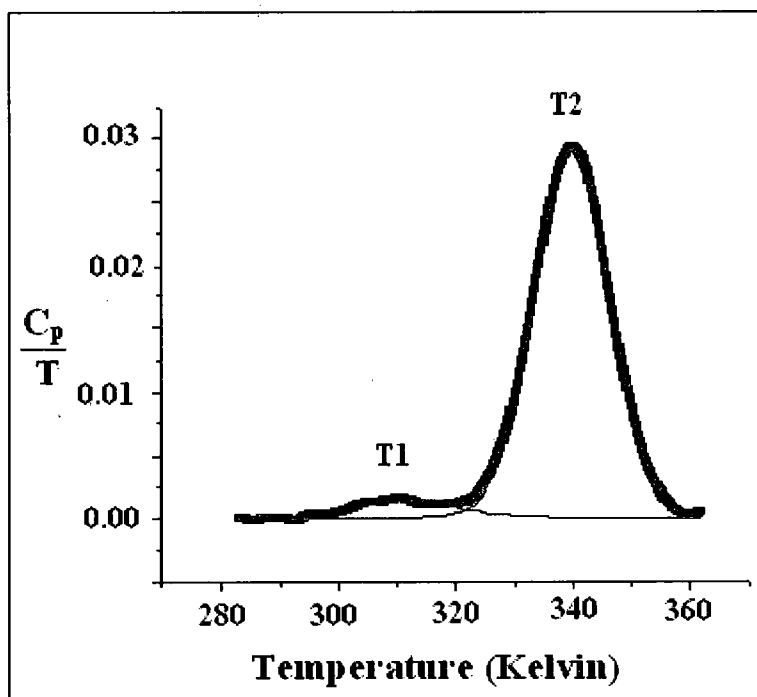


Figure 95: Deconvolution of a representative DSC thermogram obtained at 1M K^+ into two transitions to calculate calorimetric entropy for each transition

Table 5 presents the calorimetric thermodynamic data for transition I obtained from the deconvoluted DSC thermograms at various salt concentrations. ΔG was calculated at 42°C because this is the optimal growth temperature for *Halobacterium Salinarum*. All data is reported as the average from 6 scans. Table 6 and 7 presents the data for transition II and the overall thermodynamic parameters respectively.

Table 5: Transition I thermodynamic parameters

[K ⁺] (mM)	Transition I			
	T _M (°C)	ΔH_{Cal} (kcal/mol)	ΔS_{Cal} (cal/mol/K)	$\Delta G @ 42^\circ\text{C}$ (kcal/mol)
15	23.3 + 1.1	-0.6 + 0.1	-2.0 + 0.5	-0.06 + 0.02
50	16.7 + 0.4	-2.6 + 0.1	-8.9 + 0.3	-0.20 + 0.02
100	21.4 + 0.2	-2.8 + 0.3	-9.7 + 0.9	-0.28 + 0.04
250	27.0 + 0.2	-2.5 + 0.1	-8.6 + 0.5	-0.18 + 0.03
500	32.3 + 0.7	-1.6 + 0.2	-5.3 + 0.6	-0.04 + 0.01
1000	39.8 + 0.5	-2.1 + 0.2	-6.5 + 0.7	-0.00 + 0.01
2000	40.0 + 0.9	-1.3 + 0.3	-4.3 + 0.8	-0.01 + 0.01
3000	40.3 + 0.9	-1.7 + 0.1	-5.5 + 0.4	-0.02 + 0.02
4000	45.8 + 1.1	-1.1 + 0.3	-3.3 + 1.0	-0.03 + 0.02

Table 6: Transition II thermodynamic parameters

[K ⁺] (mM)	Transition II			
	T _M (°C)	ΔH _{Cal} (kcal/mol)	ΔS _{Cal} (cal/mol/K)	ΔG @ 42°C (kcal/mol)
15	43.4 + 0.2	-29.1 + 0.7	-91.8 + 2.2	-0.16 + 0.02
50	49.0 + 0.2	-24.5 + 0.3	-76.1 + 1.1	-0.52 + 0.02
100	54.2 + 0.2	-23.1 + 0.4	-70.9 + 1.3	-0.80 + 0.02
250	60.1 + 0.2	-21.2 + 0.3	-63.8 + 1.0	-1.09 + 0.02
500	64.0 + 0.2	-19.5 + 0.7	-57.8 + 2.2	-1.26 + 0.05
1000	66.2 + 0.2	-18.4 + 0.4	-54.3 + 1.3	-1.30 + 0.03
2000	66.7 + 0.3	-16.9 + 0.5	-49.9 + 1.5	-1.22 + 0.06
3000	66.0 + 0.2	-16.9 + 0.3	-50.0 + 0.9	-1.20 + 0.04
4000	64.3 + 0.3	-15.6 + 0.6	-46.3 + 1.8	-1.02 + 0.05

Table 7: DSC overall thermodynamic parameters

[K ⁺] (mM)	Overall			
	T _M (°C)	ΔH _{Cal} (kcal/mol)	ΔS _{cal} (cal/mol/K)	ΔG @ 42°C (kcal/mol)
15	43.2 + 0.4	-30.9 + 0.9	-97.8 + 3.1	-0.07 + 0.05
50	49.1 + 0.3	-29.1 + 0.3	-90.5 + 1.1	-0.61 + 0.05
100	54.2 + 0.2	-28.3 + 0.34	-86.3 + 1.0	-1.08 + 0.04
250	60.3 + 0.3	-24.7 + 0.4	-74.0 + 1.2	-1.36 + 0.03
500	64.0 + 0.3	-21.8 + 0.7	-64.8 + 1.9	-1.40 + 0.06
1000	66.4 + 0.3	-20.7 + 0.4	-61.2 + 1.3	-1.46 + 0.05
2000	66.5 + 0.4	-18.9 + 0.2	-55.3 + 0.4	-1.41 + 0.07
3000	66.0 + 0.3	-18.9 + 0.2	-55.5 + 0.7	-1.38 + 0.06
4000	64.3 + 0.5	-16.8 + 0.5	-49.7 + 1.4	-1.13 + 0.07

The T_M appears to be salt dependent for both transitions upto certain concentration as indicated by (Figure 96). However, ΔH and ΔS both appear to be salt independent for transition I but salt dependent for transition II (Figure 97 and 98).

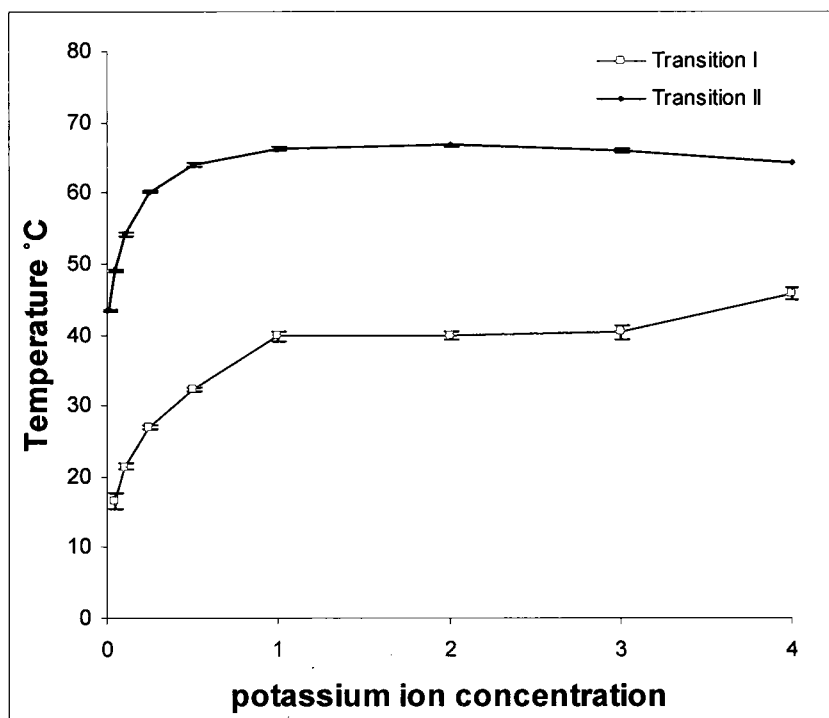


Figure 96: T_M dependency on salt concentration

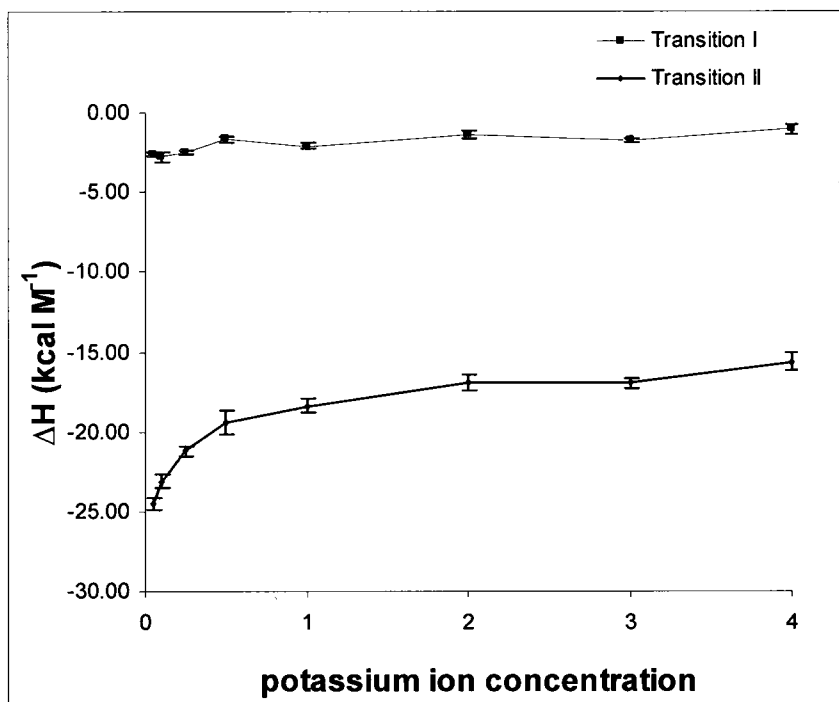


Figure 97: ΔH dependency on salt concentration

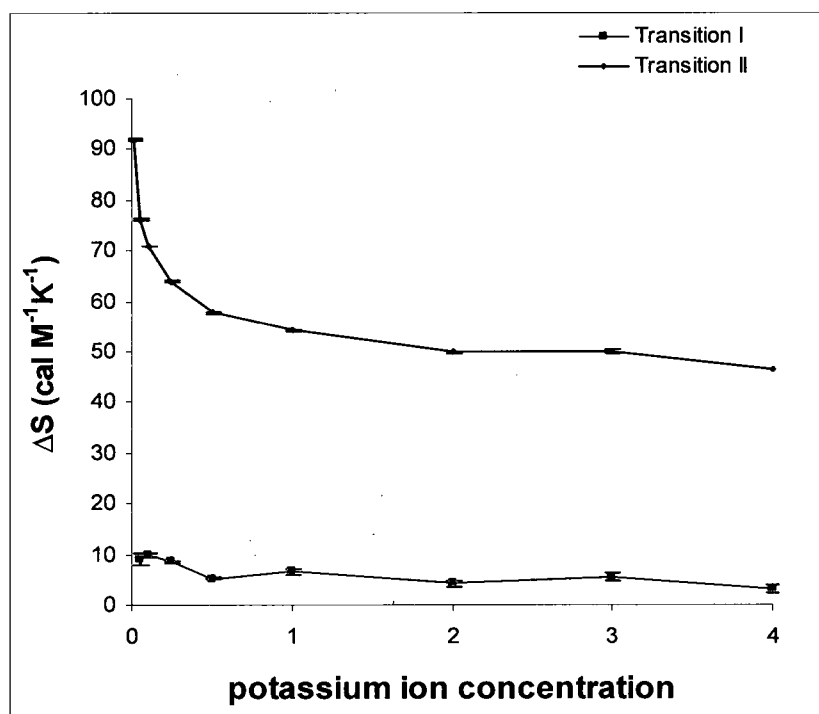


Figure 98: ΔS dependency on salt concentration

Comparison between the T_M values from the in silico method (Mfold), UV thermal melt (van't Hoff), CD thermal melt (van't Hoff) and the model independent DSC calorimetric values revealed that all the techniques provides relatively similar values within the margin of experimental error. This indicates that all the techniques are monitoring the same transitions (Figure 99).

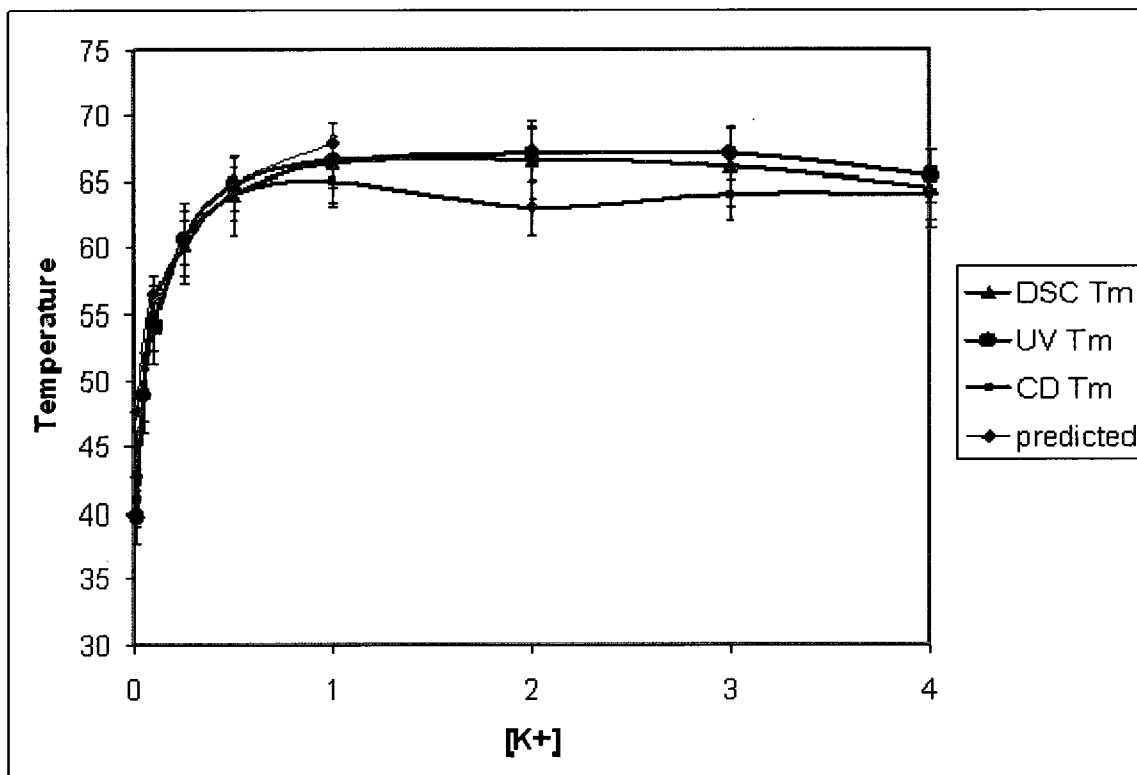


Figure 99: Comparison between the T_M values from different techniques

4.4 Biophysical analysis of the RNA model sequence

We clearly demonstrated that the DNA model can exist in conditions that mimic the *in vivo* conditions. Since the RNA model is transcribed as a single strand and can form all kinds of structures, we believe it is more relevant to perform the biophysical studies on the RNA model. Originally, we performed the studies on the DNA model as it is experimentally more robust.

4.4.1 UV Spectroscopy

4.4.1.1 Reversibility

Repetitive UV thermal melt at 260 nm were performed on the stem-loop RNA model as depicted in (Figure 100). Clearly, the repetitive scans illustrate that folding to unfolding transition is a reversible process that involves a hyperchromic shift. A total of four thermal scans alternating between heating and cooling scans ramped between 5°C and 95°C were collected. No hysteresis was observed between all the collected scans.

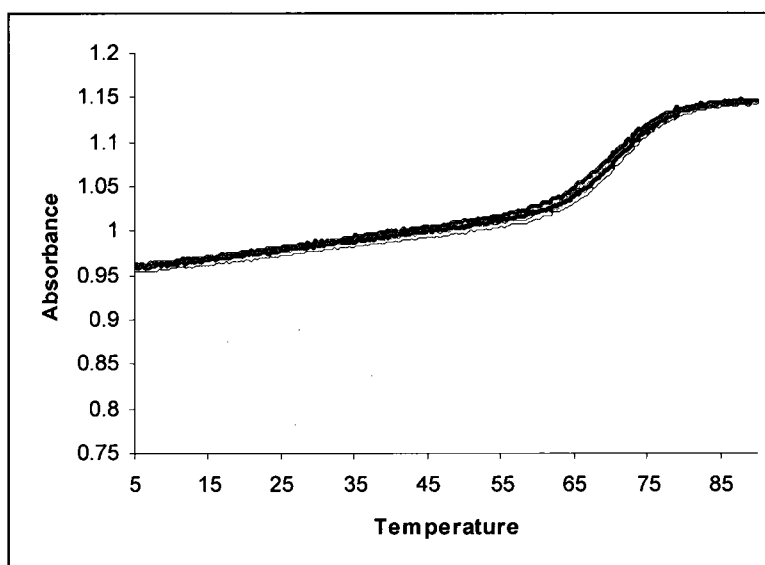


Figure 100: Repetitive RNA model heating and cooling scans at 260 nm in 1M [K⁺]

4.4.1.2 UV thermal melt of DNA stem-loop model

UV spectra of the DNA stem-loop sequences as a function of temperature collected at 260 nm are presented in (Figures 101). Total of four thermal scans alternating between heating and cooling scans ramped between 5°C and 95°C were collected at each potassium ions concentrations as follow 50, 250, 1000 and 2000 mM K⁺. The data points were collected every 0.25°C versus its respective working KPBS buffer concentration as a blank. Similar to the DNA model, the melting curves show an apparent transition

midpoint (T_M) that is directly proportional to the salt concentration up to certain salt concentrations.

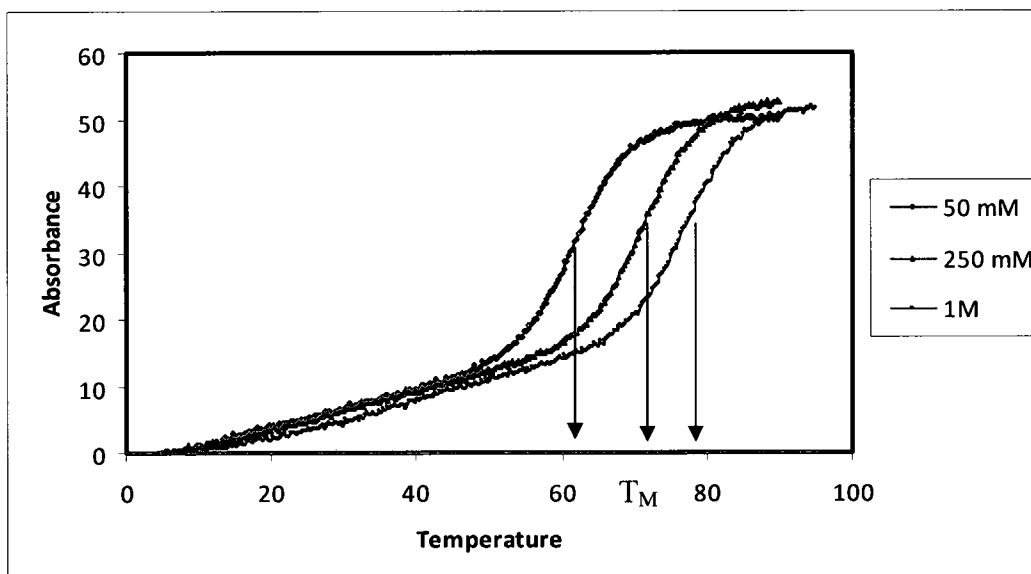


Figure 101: UV spectra at 260 nm of the RNA stem-loop model at different temperature and salt concentrations.

Close inspection of the melting curves reveal a hyperchromic transition indicating loss of base pairing.

4.4.1.3 van't Hoff Analysis

The collected data was subjected to van't Hoff analysis at various salt concentrations. The data was converted in terms of α (fraction folded RNA) and utilized to calculate K_F . van't Hoff plot ($R\ln K$ vs $1/T$) was constructed yielding a slope and an intercept essential for van't Hoff enthalpy and entropy determination respectively. Figure 102 provide a representative van't Hoff plot for the RNA model UV thermal melt. Table 8 presents the van't Hoff thermodynamic data obtained for the various salt

concentrations. T_M and ΔH_{VH} values were determined from UV absorbance data at 260 nm. All data is compiled from 4 scans.

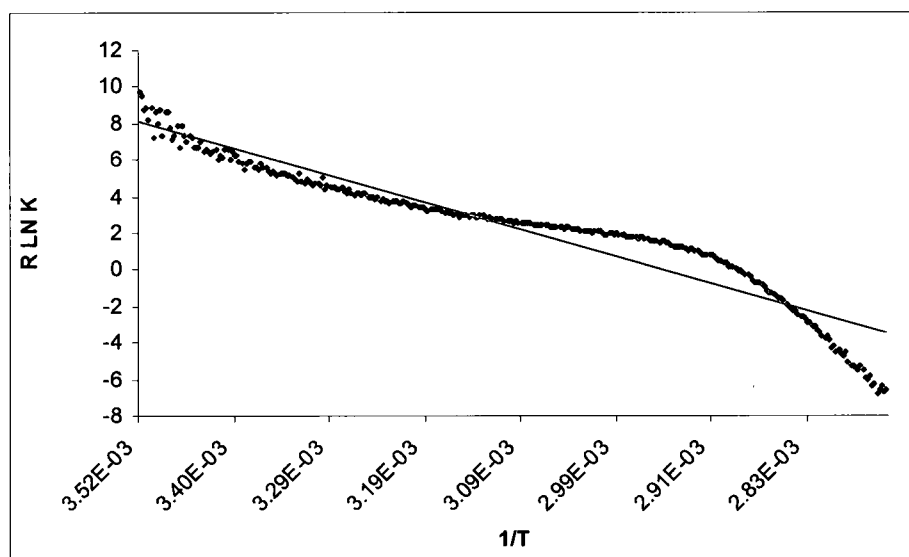


Figure 102: A representative van't Hoff plot of RNA model UV melting in 1M K^+

Table 8: van't Hoff thermodynamic parameters as a function of $[K^+]$

$[K^+]$	T_M ($^{\circ}C$)	ΔH_{VH} ($kcal\ mol^{-1}$)	ΔS_{VH} ($kcal\ K^{-1}\ mol^{-1}$)	$\Delta G_{VH}^{42^{\circ}C}$ ($kcal\ mol^{-1}$)
50 mM	61.4 ± 0.1	-20.3 ± 0.3	-63.0 ± 1.1	-0.41 ± 0.04
250 mM	70.7 ± 0.2	-13.4 ± 0.3	-40.1 ± 0.9	-0.77 ± 0.02
1000 mM	76.5 ± 0.2	-15.6 ± 1.6	-50.2 ± 3.2	0.27 ± 0.77
2000 mM	76.8 ± 0.5	-15.6 ± 0.5	-47.7 ± 1.9	-0.59 ± 0.17

Similar to the DNA model, the RNA model revealed the same issues discussed earlier and clearly demonstrated systemic deviation from a simple sigmoidal shape curve and a linear van't Hoff plot at all salt concentrations.

4.4.1.4 DNA Vs. RNA

Figure 103 provides graphic comparison between the T_M values obtained for the RNA and DNA model of the stem-loop. As differences in T_M is a direct link to differences in structure stability, it is clear that the RNA model exhibit more stable structure than the DNA model.

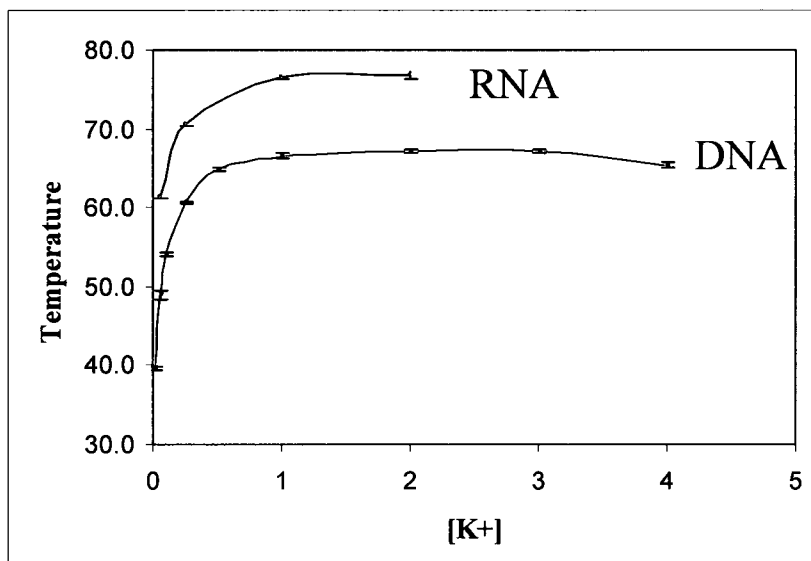


Figure 103: Comparison between the RNA and DNA T_M values

4.4.2 Circular Dichroism Spectroscopy

4.4.2.1 Temperature dependent CD Spectra of RNA stem-loop model

Figure 104 provides the CD spectrum of the RNA stem-loop at 5°C in 1M K^+ . A peak is observed at 266 nm and trough at 236 nm. 266 nm was selected to monitor and analyze the data as the more relevant wavelength to monitor the transition. The temperature dependent spectra of the RNA stem-loop model are provided in Figures 105 – 108 at different potassium ions concentrations as follow 50, 250, 1000 and 2000 mM K^+ , respectively.

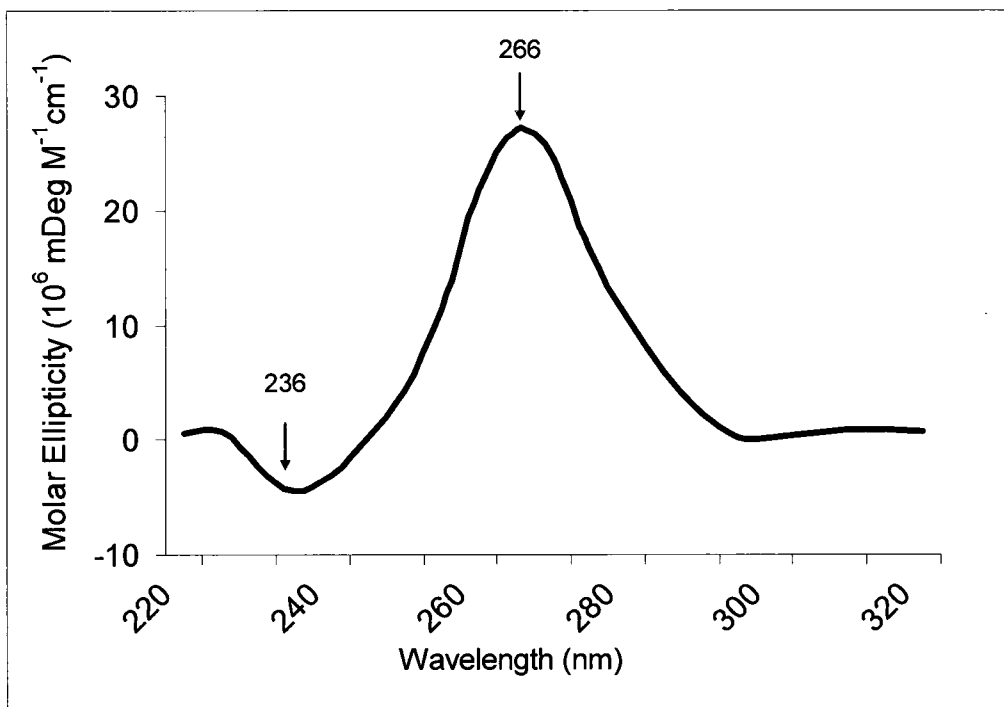


Figure 104: CD spectrum of RNA stem-loop in 1M K⁺ at 5°C.

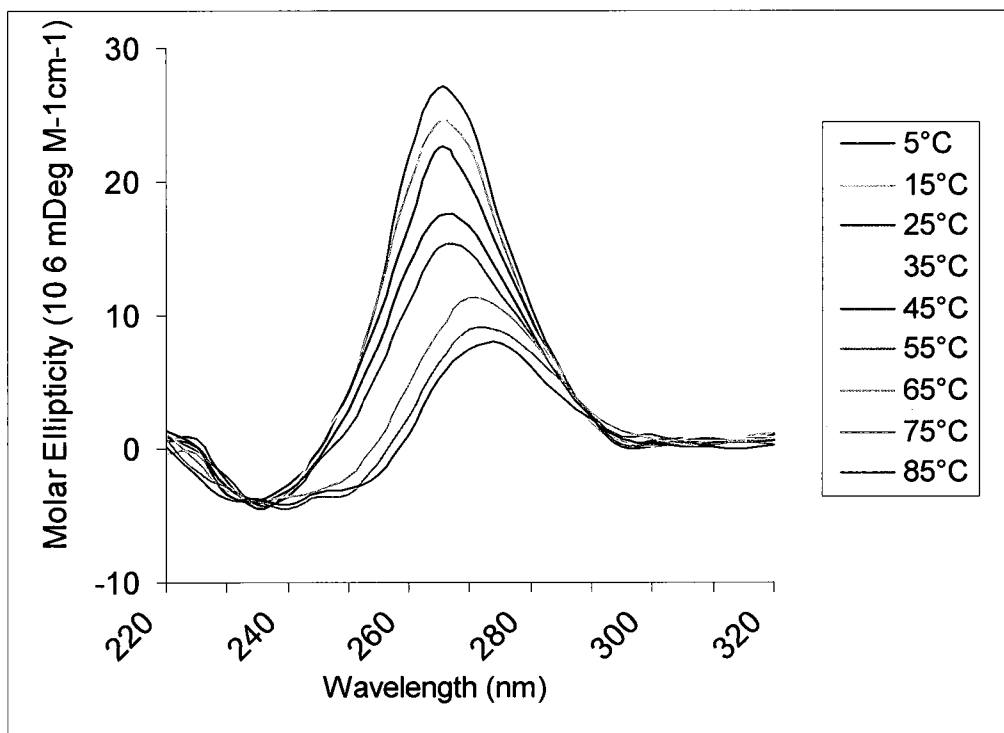


Figure 105: Temperature dependent CD spectra of RNA stem-loop in 50 mM K⁺

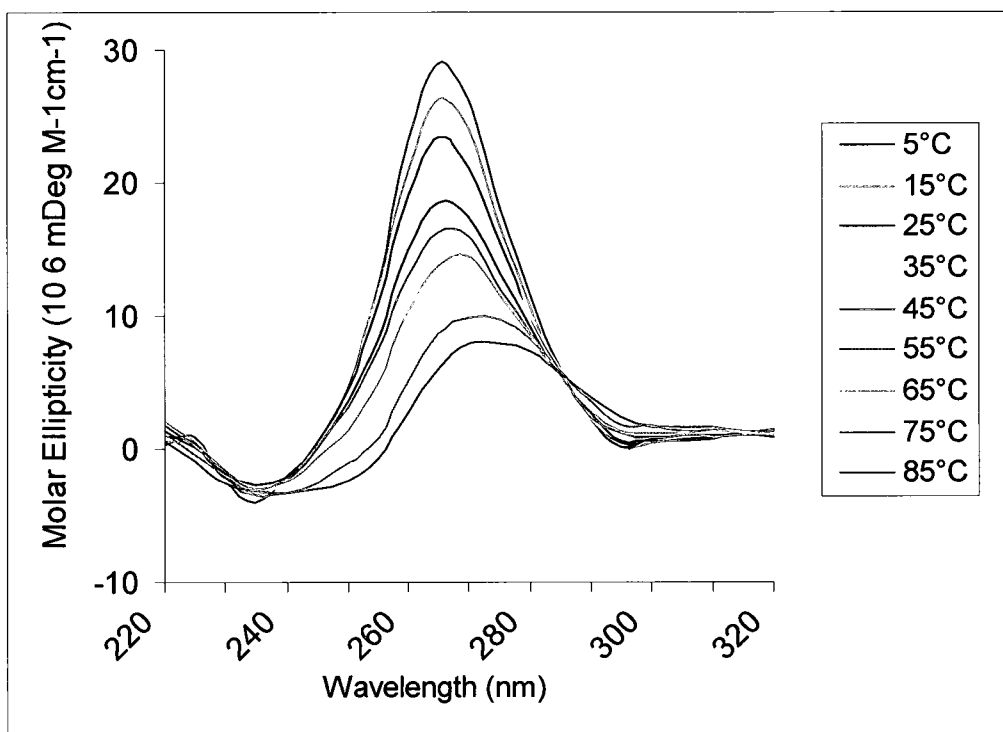


Figure 106: Temperature dependent CD spectra of RNA stem-loop in 250 mM K⁺

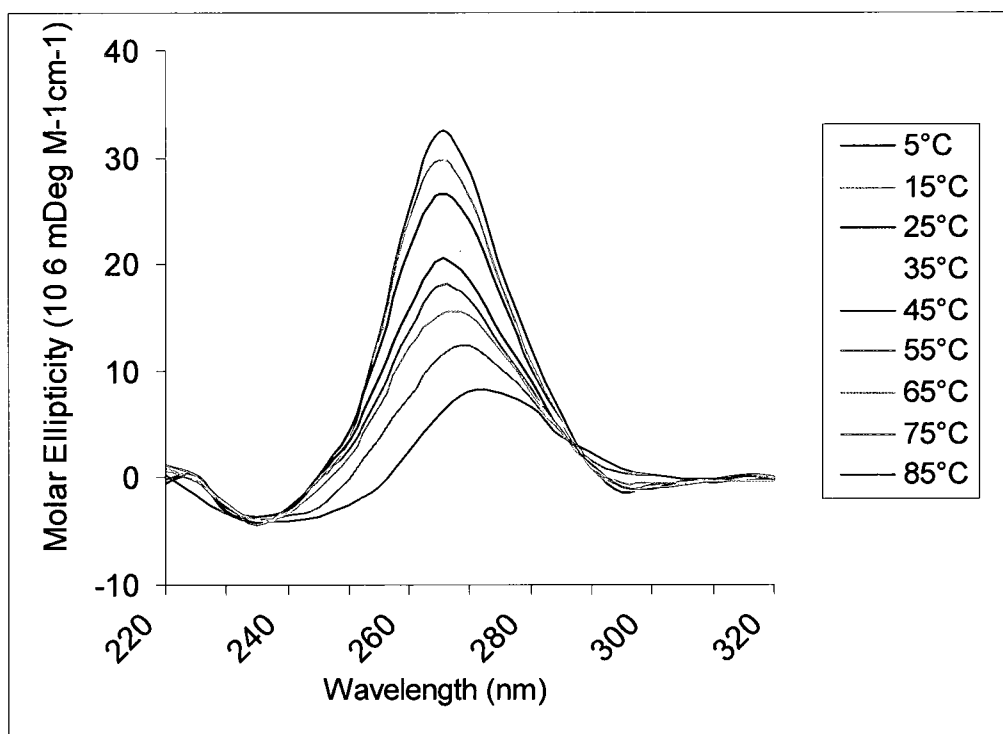


Figure 107: Temperature dependent CD spectra of RNA stem-loop in 1M K⁺

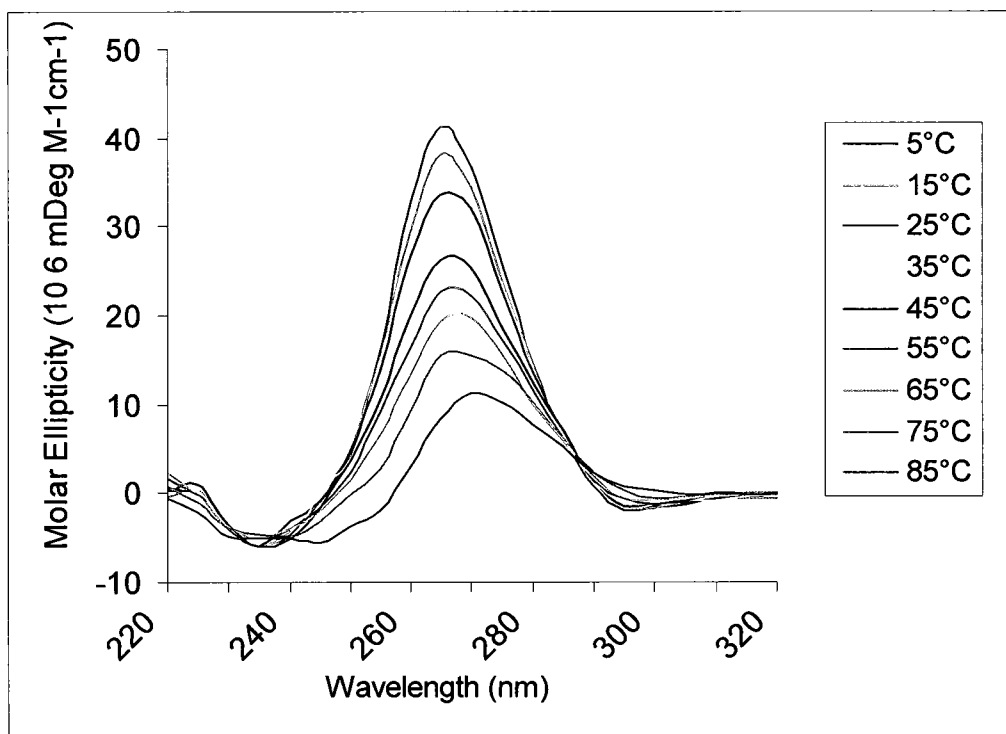


Figure 108: Temperature dependent CD spectra of RNA stem-loop in 2M K⁺

4.4.2.2 van't Hoff Analysis

The collected data was subjected to van't Hoff analysis at various salt concentrations. The data was converted in terms of α (fraction folded DNA) and utilized to calculate K_F . The van't Hoff plot ($\ln K$ Vs $1/T$) was constructed yielding a slope and an intercept essential for van't Hoff enthalpy and entropy determination respectively.

Figure 109 provide a representative van't Hoff plot for the RNA model CD thermal melt.

Table 9 presents the van't Hoff thermodynamic data obtained for the various salt concentrations.

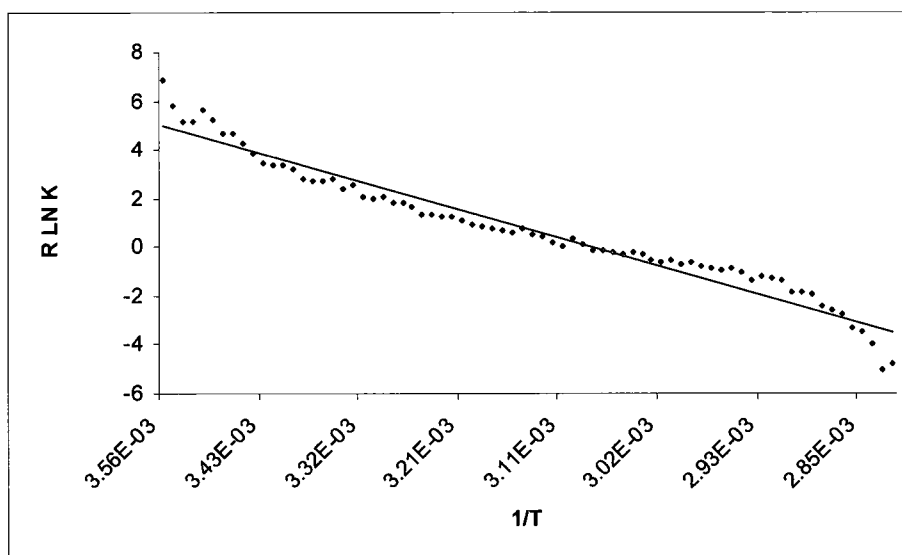


Figure 109: A representative van't Hoff plot of RNA model CD melting in 1M K⁺

Table 9: van't Hoff thermodynamic parameters as a function of [K⁺]

[K ⁺]	T _M (°C)	ΔH _{VH} (kcal mol ⁻¹)	ΔS _{VH} (kcal K ⁻¹ mol ⁻¹)	ΔG _{VH 42 °C} (kcal mol ⁻¹)
50 mM	61	-21.1	-66.1	-0.30
250 mM	70	-16.4	-50.9	-0.34
1000 mM	76	-13.7	-42.5	-0.32
2000 mM	75	-16.8	-51.6	-0.49

4.4.2.3 SVD

4.4.2.3.1 SVD analysis using *Olis global works*

The magnitude of the singular values and the randomness of the corresponding spectral and kinetic eigenvectors were used to determine the number of possible significant species. The spectral and kinetic eigenvectors plots were assessed visually for

randomness in the same manner as for the DNA model. It was determined that three spectral species are required to describe the temperature-dependent spectral transitions. This is evident by the non-randomness of the spectral and kinetic eigenvectors plots for components one through three and the randomness of the spectral and kinetic eigenvectors plots for components four and five.

4.4.2.3.2 SVD analysis using Matlab

The S matrix was 95 x 95 and contained the singular values along its diagonal. The magnitudes of the singular values obtained from the S diagonal matrix were plotted as a function of the component number to obtain the residual plot (Figure 110). The magnitudes of the singular values suggest that three major species are present during the melt of the stem-loop.

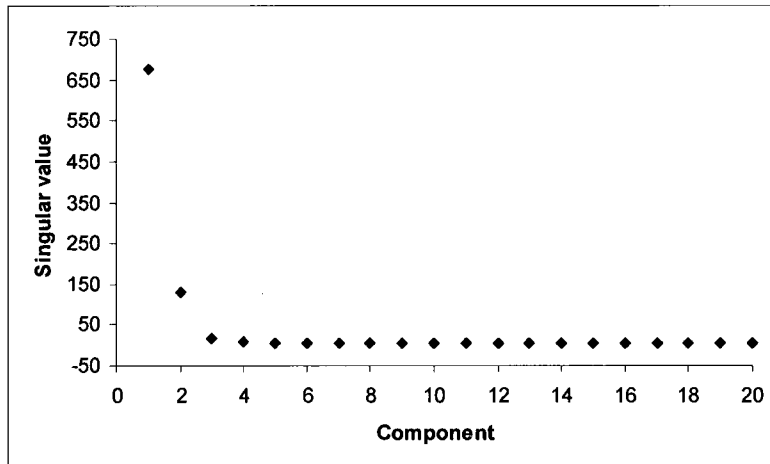


Figure 110: A representative residual plot (singular values magnitude as a function of component number) for CD melting of RNA in 1M K⁺

The amplitude vectors (V coefficients) are directly proportional to the concentration of any given component. Figure 111 provide a representative plot of V coefficient for each significant component as a function of temperature for CD melting of RNA in 1M K⁺.

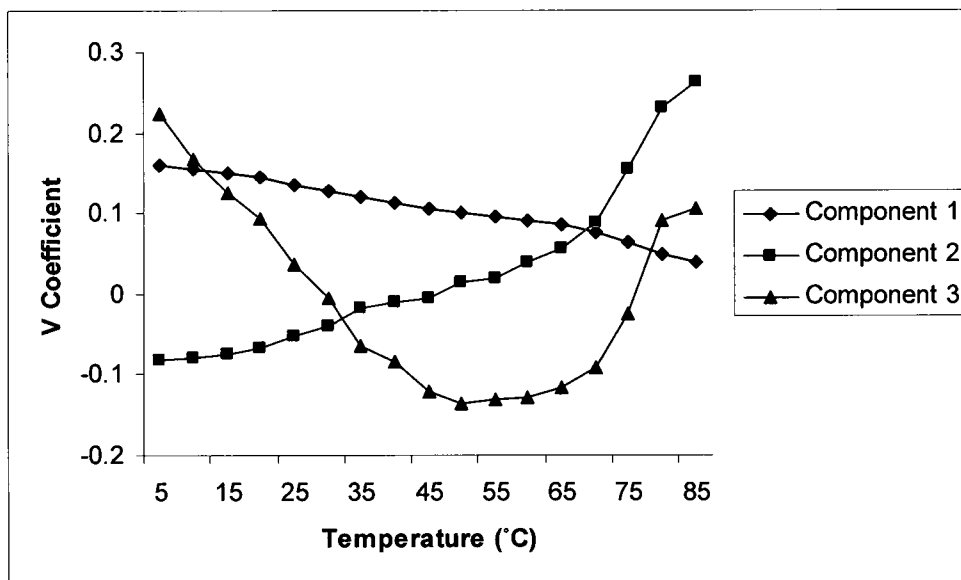


Figure 111: V coefficients for each significant component as a function of temperature for CD melting of RNA in 1M K⁺

Residual contour plots for matrices with one through five singular values were assessed visually for randomness (Figures 112 - 116). From the aforementioned figures it is clear that residual contour plots for matrices generated with one, two and three singular values shows a degree of non-randomness. While, residual contour plots for matrices generated with four and five singular values are random. This could be interpreted as an indication of the possible presence of three major species during the melt of the stem-loop, namely components one, two and three.

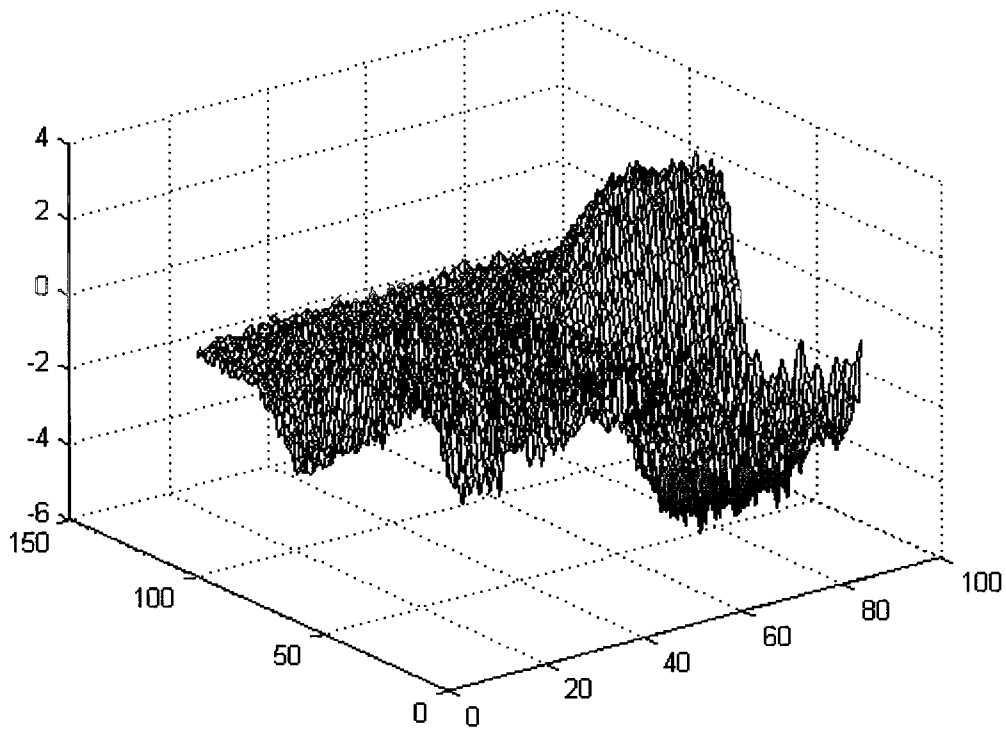


Figure 112: Residual contour plots for matrices generated with one singular value for CD melting of RNA in 50mM K⁺

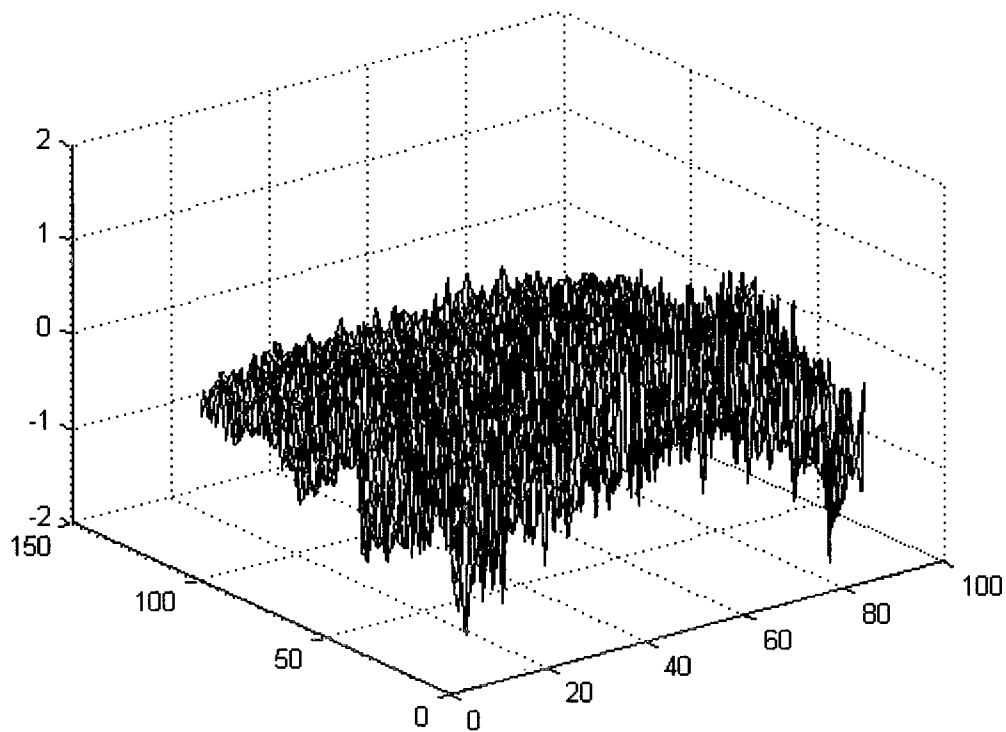


Figure 113: Residual contour plots for matrices generated with two singular values for CD melting of RNA in 50mM K⁺

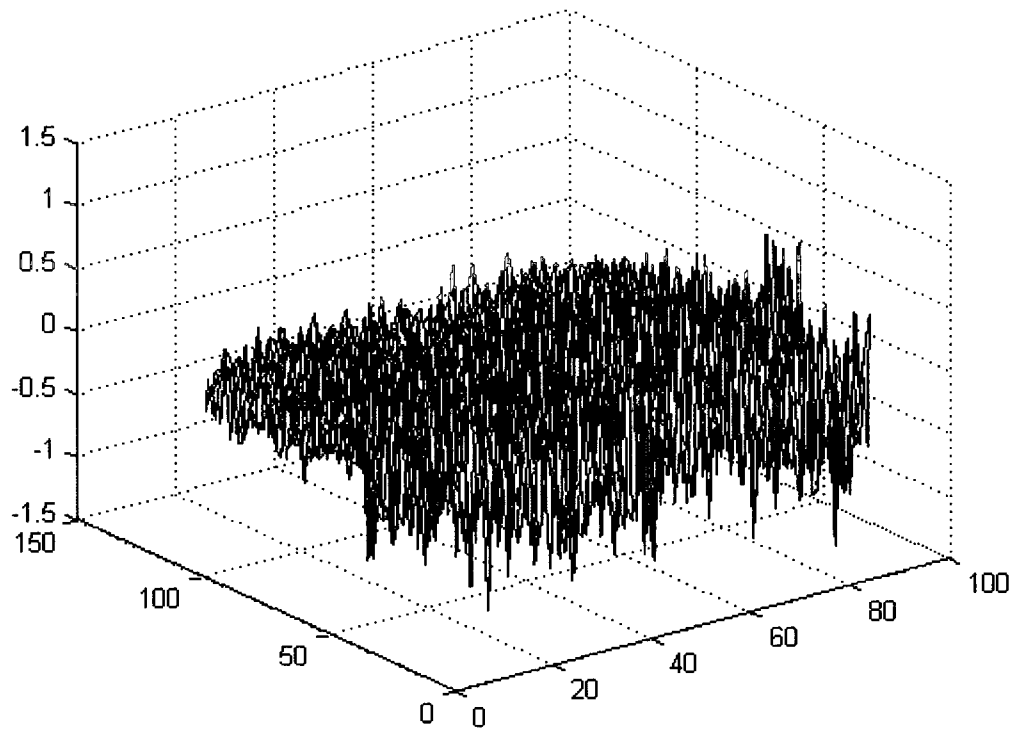


Figure 114: Residual contour plots for matrices generated with three singular values for CD melting of RNA in 50mM K⁺

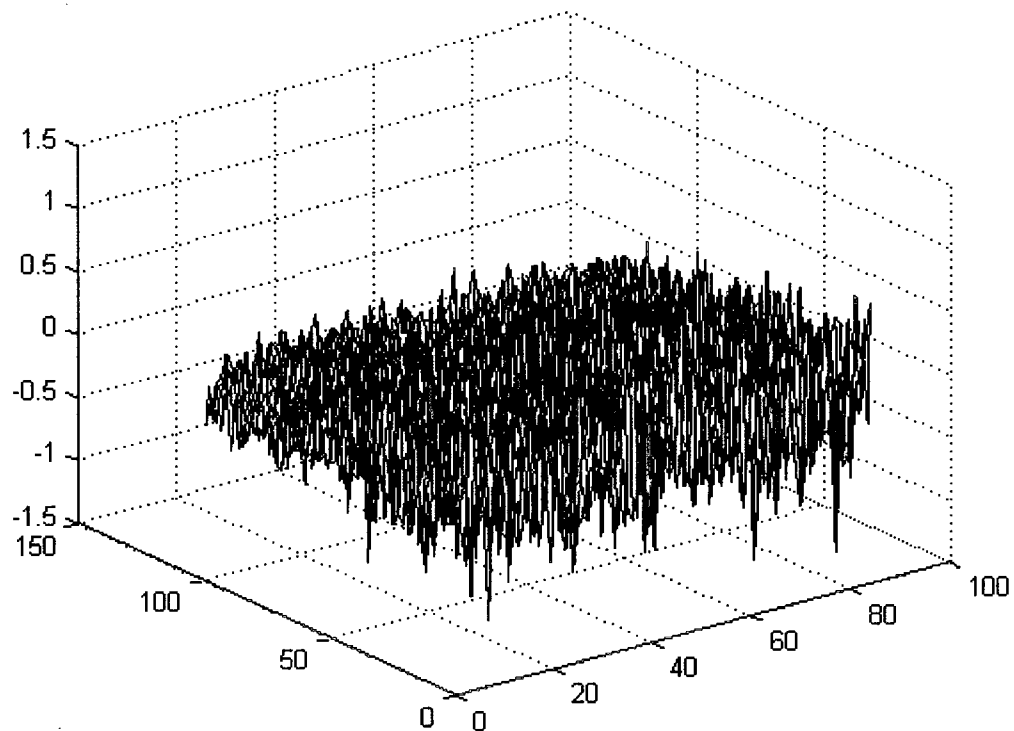


Figure 115: Residual contour plots for matrices generated with four singular values for CD melting of RNA in 50mM K⁺

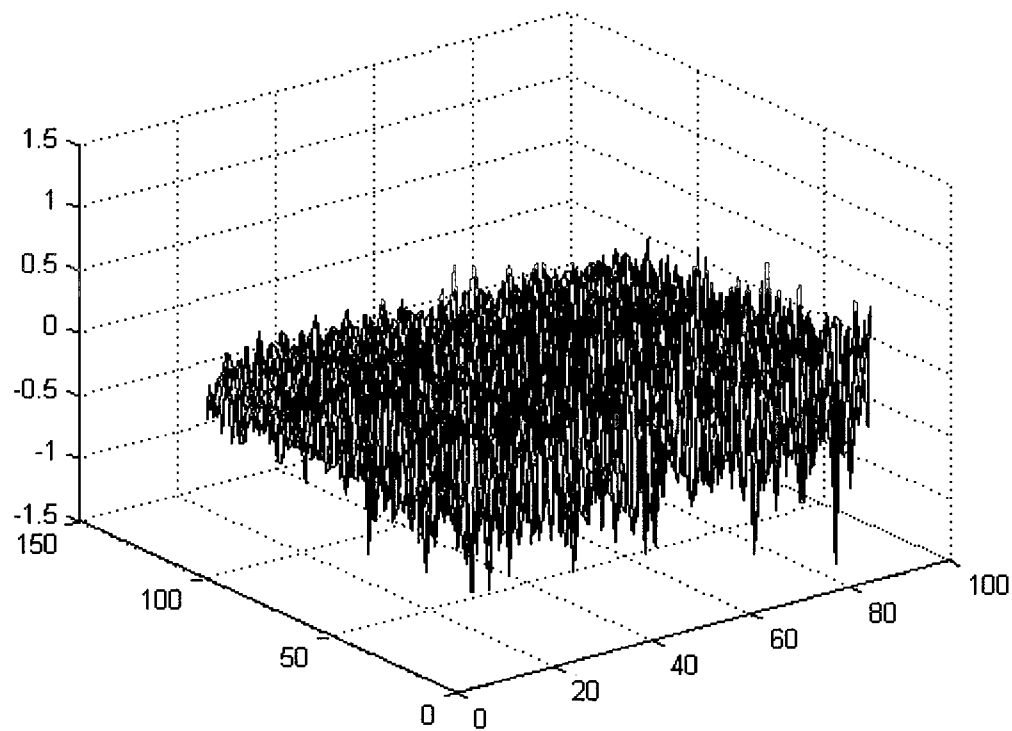


Figure 116: Residual contour plots for matrices generated with five singular values for CD melting of RNA in 50mM K⁺

5 Discussion

Understanding the mechanism of action of biological macromolecules requires the knowledge of their molecular structures and how they are coupled to the physical and chemical forces involved in biological milieu. We have applied the theories of thermodynamics to evaluate the structures and structural stability of nucleic acids. We have used the physical techniques of Ultraviolet-Visible spectroscopy, Circular Dichroism spectroscopy and Differential Scanning Calorimetry to evaluate biomolecular structure and stability of the nucleic acid secondary structural element, called a stem-loop. Singular Value Decomposition of the various collected spectroscopic data was used to monitor the equilibrium folding reactions of the nucleic acid structural element. The combination of these theories and techniques provided important secondary structural constraints from which detailed insights regarding macromolecular structure and stability were inferred.

In summary the consensus conclusion is that the energetically favorable DNA secondary stem-loop structure does exist under the biological conditions. The structure is stabilized at very high salt suggesting that it may play a role in *bop* gene expressivity. The T_M values determined by the CD and UV techniques were similar. However, the model independent enthalpy values obtained by DSC were different from the calculated van't Hoff enthalpy values determined from the model dependent spectroscopic analysis indicating the presence of more than two species that exist in equilibrium during the melting process. Consensus similar conclusion for the thermal stability of a RNA model of the stem-loop sequence was obtained. However, the RNA model of the stem-loop was significantly stable. We were also able to identify the presence and determine the number

of unfolding intermediates using van't Hoff and SVD analysis. When combined with the calorimetric data, we were able to propose an equilibrium model for the melting process.

5.1 Predictive Structural Analysis

Mfold algorithm predicts the possibility of two different secondary structures formation in the DNA model of the first twenty-five bases of the *bop* gene. One of the possible structures seems to be more energetically favorable than the other at least up to concentrations of 1.0 M NaCl (Figure 32). However, *Mfold* algorithm predicts the possibility of only one secondary structure formed in the RNA model of the first twenty-five bases of the *bop* gene (Figure 36). The predicted secondary structure in the RNA model is thermodynamically more stable than the DNA model.

In summary, the combinatorial analysis indicates that an energetically favorable stem-loop structure forms in all solution conditions examined. Comparisons of *in vitro* vs. *in silico* studies indicate a correspondence between theory and reality.

5.2 UV Spectroscopy

The simplest characterization of nucleic acids denaturing is via melting temperature, T_M . Repetitive UV thermal melts at 260nm illustrate the melting process is reversible process and involves a hyperchromic shift. The observed hyperchromic shift is a clear indication for the loss of base stacking interaction as the temperature increases. This provides strong evidence for the presence of some type of secondary structure. As the apparent T_M values dictate the stability of the DNA and RNA structures, it was determined that the stem-loop structure is stabilized at very high salt concentrations, a requirement for a role in *bop* gene expressivity. Also, when comparing different KCl

concentrations, studies have shown that KCl concentrations $>1\text{M}$ have no significant effect on the T_M values. This finding indicate that, at the optimum environmental conditions for the growth of *Halobacterium Salinarum* (42°C and 4M salt), the secondary structure is maximally stable.

Systemic deviation from a simple sigmoidal shape curve and calculated van't Hoff enthalpies indicate that the folding mechanism is not a two-state process.

5.3 Circular Dichroism Spectroscopy

The salt dependent studies clearly indicate that the transition is salt dependent. No isosbestic points were observed in the salt dependent scans. This indicates no transition between states or species as the salt concentration increases. The observed spectral changes are related to changes in the secondary structure geometry. There is salt contribution to a helical structure.

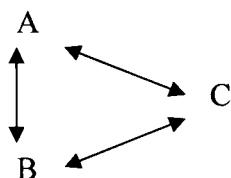
The thermal CD melts highlights the loss of helical structure as the temperature increases. The apparent T_M values calculated from the CD melts closely correlate to the apparent T_M values calculated from the UV melts. Both techniques are monitoring the same transition. The calculated thermodynamic data demonstrate at the optimum environmental conditions for the growth of *Halobacterium Salinarum* (42°C and 4M salt), the secondary structure is maximally stable. However, systemic deviation from a simple sigmoidal shape curve and calculated van't Hoff enthalpies indicate that the folding mechanism is a multiphasic process. SVD analysis of the CD spectral transitions confirm that the thermal melt of the stem-loop structure proceeds through a single significant folding intermediate at all salt concentration.

5.4 Differential Scanning Calorimetry

The model independent DSC thermograms clearly indicate at all salt concentrations the process exhibits multiple transitions. Upon deconvolution, the data was best described using two transitions. This confirms that the process proceeds through a single significant folding intermediate at all salt denoted as follow:



And / or



The T_M appears to be salt dependent for both transitions upto certain concentration as indicated in (Figure 96). However, ΔH and ΔS both appear to be salt independent for transition I but salt dependent for transition II. Comparison between the T_M values from the in silico method (*mfold*), UV thermal melts (van't Hoff), CD thermal melts (van't Hoff) and the model independent DSC calorimetric values revealed that all the techniques provides relatively similar values within the margins of experimental error. This indicates all the techniques are monitoring the same transitions. Analysis of the RNA sequence demonstrates that an RNA stem-loop structure can form and is significantly more stable than that the DNA structure at all salt concentrations evaluated.

6 References

1. Dunn, R.; McCoy, J.; Simsek, M.; Majumdar, A.; Chang, S.H.; RajBhandary, U.L.; Khorana, H.G. *Proc. Natl. Acad. Sci.* **1981**, 6744-6748.
2. Kolbe, M.; Besir, H.; Essen, L.O.; Oesterhelt, D. *Science* **2000**, 288 (5470): 1358-9
3. Oesterhelt, D. *Curr. Op. Struct. Biol.* **1998**, 8, 489-500.
4. Matheson, D.W.; Engel, W.K.; Derrer, E.C. *Neurology* **1976**, 26, 1182.
5. Garrett, R. H.; Grisham, C. M. *Biochemistry*, 2nd ed.; Brooks/Cole-Thomson Learning: California, 1999.
6. Mathews, C. K.; Van Holds, K. E.; Ahern, K. G. *Biochemistry*, 3rd ed.; Benjamin/Cummings, an imprint of Addison Wesley Longman: California, 2000.
7. Reusch, W. Natural Products *University of Massachusetts, Amherst* [Online] May 1, 2004, Nucleic Acids. <http://www.cem.msu.edu/~reusch/VirtualText/nucacids.htm> (accessed on Sep 23, 2009)
8. Antonacci, C., The biophysical analysis of the human telomeric repeat (TTGGG)₄ and characterization of porphyrin binding in potassium solutions. Ph.D. Dissertation, Seton Hall University, South Orange, NJ, 2007.
9. Woese, C. R.; Winker, S.; Gutell, R. R. *Proc. Natl. Acad. Sci.*, **1990**, 87, 8467-8471.
10. Tinoco Jr., I.; Bustamante, C. How RNA Folds. *J. Molecular Biology*, **1999**, 293, 271-81.
11. Magrum, Linda J.; Luchrsen, Kenneth R.; Woese, Carl R. *J. Mol Evol.*, **1978**, 11, 1-8.
12. Rengpipat, S.; Lowe, S. E.; and Zeikus, J. G. *Journal of Bacteriology*, **1988**, 3065-3071.
13. Stoeckenius, W.; Rowen, R. *Journal of Cell Biology*, **1967**, 34, 365-393.
14. Magyari, K.; Balint, Z.; Simon, V.; and Varo, G. The Photochemical Reaction Cycle of Retinal Reconstituted Bacteriorhodopsin. *J. Photochem. And Photobiol. B-Biology*, **2006**, 85, 140-144.
15. Turner, G. J. *Methods in Cell Biology*, **2008**, 84, 479-515.
16. Theoretical Biophysics Group. University of Illinois: Structure and Dynamics of the Purple Membrane. <http://www.ks.uiuc.edu/Research/newbr/> (accessed Apr 23, 2009).
17. Gropp, F. et al. *Mol. Microbiol.*, **1995**, 16, 357-364.
18. Baglia, N. S.; DasSarma, S. *Mol. Microbiol.*, **2000**, 36 (5): 1175-1183.
19. Winter-vann.; Marinez, L.; Parker, L.; Talbot, J.; and Turner, G. Halobacterium Salinarum, in Perspectives on solid state NMR in Biology; Kiihne, S.; and De Groot, H. J. M., Eds. 2001, pp 141-160
20. Kenkel, J. *Analytical Chemistry Refresher Manual*, first ed.; CRC Press: Florida, 1992.
21. Strong, F.C. *Anal. Chem.*, **1952**, 24, 338.
22. Cutnell, J.D.; Johnson, K.W. *Electromagnetic Waves, Physics*, 3rd ed.; John Wiley & Sons: New York, 1995; Chapter 24.

23. Tinoco, I.; Sauer, K.; Wang, J.C.; and Puglisi, J.D. *Physical Chemistry: Principles and Applications in the Biological Sciences*, 4th ed.; Prentice Hall: New Jersey, 2002; Chapter 10.
24. Freifelder, D. *Physical Biochemistry: Applications to Biochemistry and Molecular Biology*, 2nd ed.; W.H. Freeman and Company: New York, 1982; Chapter 16.
25. Spectroscopic Methods in Biochemistry: 4. CD Spectroscopy.
<http://www.uni-konstanz.de/FuF/Bio/folding/4-CD%20Spectroscopy%20r2.pdf>
(accessed on Aug 23, 2009)
26. Johnson, W.C. *CD of Nucleic Acids. In Circular Dichroism-Principle and Applications*; Nakanishi, K.; Berova, N.; Woody, R.W., Eds.; VCH Publishers: New York, 1994; pp 523-539.
27. Cantor, Charles R.; Warshaw, Myron M.; Shapiro, Herman. *Biopolymers*, **1970**, *9*, 1059-1077.
28. Markey, L.; Breslauer, K.J. *Biopolymers*, **1987**, 1601-1620.
29. SantaLucia Jr., J.; Hicks, D. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, 415-440.
30. *SVD and Signal Processing: Algorithms, Analysis and Applications*; Deprettere, F. Ed; Elsevier Science: North Holland, 1988.
31. Gene H. Golub, G. H.; Van Loan, C. F.; and Hopkins, J. *Matrix Computations*, first ed.; University Press: Baltimore, Maryland, 1983, pp 16-21, 29
32. Baldwin, J. M. *Embo J.*, **1993**, *12*, 1693 – 1703.
33. Baldwin, J. M. *J. Mol. Biol.*, **1997**, *272*, 144–164.
34. Kyte, J.; Doolittle, R. F. *J. Mol. Biol.*, **1982**, *157*, 105 – 132.
35. Zuker, M. *Nucleic Acids Res.*, **2003**, *13*, 3406-15.
36. Andronescu, M. Algorithms for predicting the secondary structure of pairs and combinatorial sets of nucleic acid strands. Master's Thesis, University of British Columbia, Canada, 2003.
37. Oesterhelt, D.; Stoeckenius, W. *Methods Enzymol.*, **1974**, *31*, 667-678.
38. Cutnell, J.D.; Johnson, K.W. *Electromagnetic Waves, Physics*, 3rd ed.; John Wiley & Sons: New York, 1995; Chapter 24.
39. Umland, J.B.; Bellama, J.M. *Atomic Structure, General Chemistry*, 2nd ed.; West Publishing Company: New York, 1996; Chapter 7.
40. Freifelder, D. *Physical Biochemistry: Applications to Biochemistry and Molecular Biology*, 2nd ed.; W.H. Freeman and Company: New York, 1982; Chapter 14.
41. Atkins, P.W. *Physical Chemistry*, 5th ed.; W.H. Freeman and Company: New York, 1994; Chapter 16.
42. Cantor, C.R.; Schimmel, P.R. *Biophysical Chemistry: The Behavior of Biological Molecules*, W.H. Freeman and Company: New York, 1980; Vol. 2; Chapter 8.
43. Freifelder, D. *Physical Biochemistry: Applications to Biochemistry and Molecular Biology*, 2nd ed.; W.H. Freeman and Company: New York, 1982; Chapter 16.
44. Baker, K. Singular Value Decomposition Tutorial. [Online] March 29, 2005, <http://www.fmrrib.ox.ac.uk/~tkincses/jc/SVD.pdf> (accessed Jan 15, 07).
45. Henry, E. R.; Hofrichter, J. *Methods in Enzymology*, **1992**, *210*, 129-192.
46. Johnson, W.C. *Methods in Enzymology*, **1992**, *210*, 426-447.
47. Sheardy, R.D.; Suh, D.; Kurzinsky, R.; Doktycz, M.J.; Benight, A.S.; Chaires, J.B. *J. Mol. Biol.* **1993**, *231*, 475-488.
48. Haq, I.; Chowdhry, B.Z.; Chaires, J.B. *Eur. Biophys.*, **1997**, *26*, 419-426.

49. Ott, J.B.; Boerio-Goates, J. *Chemical Thermodynamics: Principles and Applications Molecules*, Elsevier Academic Press: New York, 2000; Chapter 2.
50. CSC Model 6100 Nano II Differential Scanning Calorimeter User's Manual, Revision 1.1; Calorimetric Sciences Corporation, 1998.
51. Skoog, D.A.; Holle, F.J.; Nieman, T.A. *Principles of Instrumental Analysis*, 5th ed.; Harcourt Brace College Publishing: Philadelphia, 1998; Chapter 13.
52. Wang, G.; Vasquez, K.M. *Proc. Natl. Acad. Sci.*, **2004**, *37*, 13448-13453.
53. Watson, J. D.; Crick, F. H. *Nature*, **1974**, *248*, 765.
54. Owczarzy R., Vallone P.M., Gallo F.J., Paner T.M, Lane M.J., and Benight A.S. *Biopolymers*, **1997**, *44*, 217-239.
55. Rouzina, I., Bloomfield, V.A. Force-Induced Melting of the DNA Double Helix. *Biophysical Journal* **2001**, *Vol. 80*, 882–893.
56. Breslauer, K. J.; Franks, R.; Blockers, H.; Marky, L.A. *Proc. Natl. Acad. Sci.*, **1986**, *83*, 3746-3750.
57. Boysen, R. I.; Wang, Y.; Keah, H. H.; Hearn, M. T. W. *Biophysical Chemistry*, **1999**, 79-97.
58. DasSarma, S.; RajBhandary, U.L.; Khorana, H.G. *Proc. Natl. Acad. Sci.* **1984**, 125-129.
59. Leharne, S.; Chowdhry, B. Z. *Biocalorimetry: Applications of Calorimetry in the Biological Sciences*, John Wiley & Sons: Chichester, 1998; Chap 12.
60. Matheson, A. T.; Sprott, G. D.; McDonald, I. J.; Tessier, H.; *Microbiol.* **1976**, *6*, 780-786.
61. John, D.M.; Weeks, K.M. *Protein Science* **2000**, *9*, 1416-1419.