

Portland State University

PDXScholar

---

Dissertations and Theses

Dissertations and Theses

---

1-1-2010

# Thermodynamic Effects of 5' and 3' Single Strand Dangling Ends on Short Duplex DNA

Rebekah Dickman

*Portland State University*

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/open\\_access\\_etds](https://pdxscholar.library.pdx.edu/open_access_etds)

**Let us know how access to this document benefits you.**

---

## Recommended Citation

Dickman, Rebekah, "Thermodynamic Effects of 5' and 3' Single Strand Dangling Ends on Short Duplex DNA" (2010). *Dissertations and Theses*. Paper 94.

<https://doi.org/10.15760/etd.94>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).

**Thermodynamic Effects of 5' and 3' Single Strand Dangling Ends on Short  
Duplex DNA**

**by  
Rebekah Dickman**

**A thesis submitted in partial fulfillment of the  
requirements for the degree of**

**Master of Science  
in  
Chemistry**

**Thesis Committee  
Albert S. Benight, Chair  
Dirk Iwata-Reuyl  
Andrew Rice**

**Portland State University  
2010**

## **Abstract**

Differential scanning calorimetry (DSC) melting analysis was performed on 27 short double stranded DNA duplexes containing 15 to 25 base pairs and short single stranded overhangs from one to 10 bases, on both ends. Molecules have two 5' dangling ends or one 5' and one 3' dangling end. For these molecules the duplex region was incrementally reduced from 25 to 15 base pairs with increased length of the dangling ends from one to 10 bases. A third set of molecules contained 21 base pair duplexes with a four base dangling end on either the 5' or 3' end. Blunt ended duplexes from 15 to 25 base pairs were also examined and served as control duplexes. DSC melting curves were measured in solution containing 85 mM, 300 mM or 1.0 M Na<sup>+</sup>. From these measurements, thermodynamic parameters for 5' and 3' dangling-ends as a function of end length were evaluated. Results showed the 5' ends were slightly stabilizing, and this stability was essentially constant with end length, while the 3' ends were generally destabilizing with increasing length of the end. This finding of lower stability for the 3' ends is consistent with results of published studies that have found 5' dangling ends to be more than or equally as stabilizing as 3' dangling ends. Our finding that 3' dangling ends are actually destabilizing for duplex DNA contrasts with published results. The 3' ends also display a stronger dependence on the [Na<sup>+</sup>]. In the lower Na<sup>+</sup> environment the 3' ends are more destabilizing than at the higher salt environments. Analysis of the thermodynamic parameters of the dangling ended duplexes as a function [Na<sup>+</sup>] indicated the 3' dangling end molecules behave differently compared to 5' dangling

ended and blunt ended duplexes. The net counterion release per phosphate upon melting the molecules having one 5' and one 3' end was approximately 15% smaller as a function of end length compared to the duplex having two 5' ends. Further analysis of the DSC evaluated thermodynamic transition parameter  $\Delta H_{cal}$ , and its relationship to the measured transition temperatures of the DNA molecules, provided an estimate on the excess heat capacity differences,  $\Delta C_p$ , between duplex and melted single strands for the dangling-ended molecules. The analysis revealed the molecules with one 5' and one 3' dangling end had very different  $\Delta C_p$  values compared to the blunt-ended molecule; while the molecules with two 5' ends have  $\Delta C_p$  that are essentially the same as the blunt-ended duplex. These observations are interpreted as differences in the interactions with  $\text{Na}^+$ , solvent and the terminal base pairs of the duplex for the 5' versus 3' dangling ends.

## Table of Contents

<b>Abstract .....</b>	<b>i</b>
<b>List of Tables:.....</b>	<b>v</b>
<b>List of Figures .....</b>	<b>vi</b>
<b>Chapter 1 - Introduction .....</b>	<b>1</b>
<b>1.1 Introduction .....</b>	<b>1</b>
<b>1.2 End Interactions.....</b>	<b>2</b>
Observed Stabilization.....	2
Direction of Attachment to DNA or RNA.....	2
Identity of The Terminal Stack.....	3
Spatial Geometries of Dangling Ends.....	5
<b>1.3 Hybridization Applications .....</b>	<b>5</b>
<b>Chapter 2: Materials and Methods .....</b>	<b>7</b>
<b>2.1 DNA.....</b>	<b>7</b>
Sequence Design .....	7
DNA Molecules .....	11
<b>2.2. Differential Scanning Calorimetry .....</b>	<b>12</b>
Introduction.....	12
DSC Sampling .....	12
DSC Data Analysis .....	13
<b>2.3 Solutions and Preparations .....</b>	<b>17</b>
Buffer Solutions .....	17
Quantification of Samples.....	17
Buffer Exchange .....	18
Gel Characterization.....	19
<b>Chapter 3: Results .....</b>	<b>21</b>
<b>3.1 Dangling End Thermodynamics .....</b>	<b>21</b>
Enthalpy-Entropy Compensation.....	23

Parsing Assumptions .....	25
Thermodynamic Formulation .....	26
Thermodynamic Contributions From 5' Dangling Ends .....	30
Thermodynamic Contributions from 3' Dangling Ends.....	31
<b>3.2 Validity Check For Parsing Assumptions.....</b>	<b>34</b>
Difference Equations For $n_L=4$ .....	34
Examination of Difference Equations.....	38
<b>3.3 Examination of Duplex Properties .....</b>	<b>41</b>
Counterion Binding Theory .....	41
Differential Counterion Release Upon Melting.....	43
Heat Capacities of the Melting Transitions .....	44
<b>Chapter 4 – Discussion and Interpretations .....</b>	<b>49</b>
<b>4.1 Literature Comparisons .....</b>	<b>49</b>
Comparisons for 5' Dangling Ends.....	49
Comparison of 3' Dangling Ends.....	51
<b>4.2 Factors Involved in Dangling End Behavior .....</b>	<b>53</b>
Stabilization Models.....	53
The Mystery of 3' Dangling Ends and the Set II Molecules .....	56
<b>4.3 Practical Applications for Probe Design.....</b>	<b>59</b>
Predictive Ability.....	59
<b>Chapter 5 – Summary and Conclusion.....</b>	<b>64</b>
<b>References .....</b>	<b>66</b>

## List of Tables:

Table 1: Sequences for Set I Molecules	9
Table 2: Sequences for Set II Molecules	10
Table 3: Sequences for Set III Molecules	10
Table 4: Sequences for Set IV Molecules	11
Table 5: Measured Thermodynamic Parameters for Sets I, II and IV in 85 mM Na <sup>+</sup>	21
Table 6: Measured Thermodynamic Parameters for Sets I, II and IV in 300 mM Na <sup>+</sup>	22
Table 7: Measured Thermodynamic Parameters for Sets I, II and IV in 1000 mM Na <sup>+</sup>	23
Table 8: Calculated Contributions From 3' and 5' Ends	33
Table 9: Set III Measured Thermodynamic Parameters and Calculated Values	36
Table 10: Example Calculation for Nearest Neighbor, n-n, Model	59
Table 11: Predicted Values for the Free Energy, $\Delta G$ , Using Discovered Parameters	62

## List of Figures

Figure 1: Schematic Representation of Molecule Sets and Design	8
Figure 2: Typical DSC Melting Curve Analysis	16
Figure 3: Enthalpy and Entropy Compensation $\Delta H$ vs. $T\Delta S$	24
Figure 4: Calculated Dangling End Parameters; $\delta X_{ss}^{5'}$ and $\delta X_{ss}^{3'}$ vs. $n_L$	29
Figure 5: Calculated $\Delta G_{ss}^{5'}$ and $\Delta G_{ss}^{3'}$ with Trends	33
Figure 6: Comparative Histogram of $\delta X_{ss}^{3'}$ and $\delta X_{ss}^{5'}$ for $n_L=4$	37
Figure 7: Schemes for Calculating $\delta X_{ss}^{3'}$ and $\delta X_{ss}^{5'}$ for $n_L=4$	40
Figure 8: Counterion Release; $\Delta\Psi$ vs. $n_L$	42
Figure 9: Heat Capacity; $\Delta H(T_m)$ vs. $T_m$	47
Figure 10: Difference in Heat Capacity; $\Delta\Delta H(T_m)$ vs. $T_m$	47
Figure 11: Stacking Interaction Differences for Dangling Ends in DNA and RNA	54



## **Chapter 1 - Introduction**

### **1.1 Introduction**

Discovery of the structure of the DNA double helix in 1953 changed the world of biological science in ways that few discoveries ever have<sup>2</sup>. Almost 60 years later our understanding of the interactions essential in maintaining the structure and stability of the fascinating and deceptively complex double helix of DNA is still far from complete. When two single strands of DNA anneal, a duplex can form depending on the complement of the two single strand sequences. Since the 60s, investigations of the thermodynamics involved in DNA strand annealing have been an active area of research<sup>3,4,5,6, 7</sup>. Considerable progress has been made in defining and determining sequence dependent parameters for characterizing the annealing or melting process and the overall stability of short duplex DNA.

Canonical binding is dictated by the specific arrangement of the four DNA bases, and follows the Watson-Crick (W-C) base pairing scheme where Adenine (A) binds Thymine (T) and Guanine (G) binds Cytosine (C). However, there are times when two sequences may not be perfectly complementary because of their individual specific sequences or different lengths. In such cases, perfect binding is not always possible or required, and the thermodynamics of duplex annealing may be altered. In this thesis work we set out to investigate how the occurrence of dangling ends, which are terminal unmatched single strand bases on the ends of the duplex, affects the thermodynamic stability of short DNA duplexes.

## **1.2 End Interactions**

### **Observed Stabilization**

Research on RNA dangling ends is more abundant than it is for DNA, due to the frequent appearance of RNA dangling ends in biological reactions such as the sequence dependent stabilization of the aminoacyl stem of tRNA by a single strand 3' sequence<sup>8</sup> and codon-anticodon dangling end interactions between tRNA and mRNA<sup>9</sup>. Observed increases in stability are postulated to be largely enthalpic in origin, a direct result from additional cross-strand stacking and favorable solvent interactions<sup>6,10,4,11</sup>. Specific interactions that explain the magnitude and degree of effect imparted by the addition of dangling ends to the duplex include, direction of attachment to DNA or RNA, terminal base-pair identity, specific dangling residue, and location of residue in relation to the terminal base pair.

### **Direction of Attachment to DNA or RNA**

Constraint of dangling bases to orientations that allow for better terminal base pair interactions, allow more stability to be granted to the duplex. Generally RNA dangling ends provide more stability than DNA dangling ends<sup>10,12</sup>. Correlation plots of hydrogen screening versus thermodynamic stabilization ( $\Delta G^\circ$ ) show that RNA ends are much more ordered than DNA ends<sup>13</sup>. This is thought to explain some of the observed stability differences for RNA versus DNA<sup>14</sup>.

Dangling residues have also been found to provide greater stability when attached to the 3' end of RNA than when they are attached to the 5' end<sup>10,4</sup>. Population dynamic studies of RNA molecules with dangling residues as seen via high-resolution NMR, have found that the addition of a dangling end on the 3' terminus results in more cooperative melting of the duplex<sup>15</sup>. Additionally femtosecond spectroscopic dynamic measurements for RNA with a 3' purine base dangling end, found that the majority population had significant overlap between the terminal base-pair and the dangling base<sup>16</sup>. In contrast the high resolution NMR showed that a 5' base addition provides more conformational freedom and the dynamic measurements showed the majority population to exist in largely unstacked states with regard to the terminal base pair. This suggests that for RNA additional residues act as an extension of the original duplex when attached to the 3' end, but not when attached to the 5' end<sup>15</sup>. Interestingly, observations of DNA dangling ends find the opposite to be true, bases added to the 5' end are more stabilizing than their 3' counterparts<sup>1</sup>.

### **Identity of The Terminal Stack**

In the nearest-neighbor model (n-n), hydrogen bonding between base pairs and stacking interactions between neighboring bases are considered to be the major sources of sequence dependent duplex stability<sup>5,17,18</sup>. The inclusion of an unpaired dangling end may provide the terminal base pair with interactions that mirror those of an additional base pair<sup>11</sup>. The dangling residue creates a shielded environment

resembling that of the interior of the duplex, in turn allowing increased stacking interactions which provide greater stabilization through screening from less favorable solvent interactions<sup>19</sup>. Depending on the particular identity of the dangling base, favorable interactions can augment duplex stability to an extent comparable to a normal W-C AT base pair<sup>4</sup>.

The fact that the terminal base pair can be stabilized through interactions with the dangling end immediately suggests the specific identity of the dangling end is of considerable importance in contributing to this stability. In general, the added stabilization afforded by a terminal dangling base follows the order (Purine)>T>C<sup>20,1,21</sup>. Apparently the larger bases (A or G) with a greater surface area, provide for better stacking interactions.

These stacking interactions between the bases that make up the terminal stack (the terminal base pair and dangling base) modulate the strength of the terminal interactions. For example, according to the n-n parameters that have been reported, the terminal stack <sup>5'</sup>AC/<sup>3'</sup>G, where the A is a 5' dangling end stacked on the C of a duplex terminating in a CG W-C base pair, does not confer the same stability as the inverted <sup>5'</sup>AG/<sup>3'</sup>C terminal stack, where the 5' dangling A stacks above the G on the terminal W-C GC base pair<sup>1</sup>. Differing specific interactions between the base of the dangling end and terminal base pair, as well as interactions with the hydration scheme of the terminal base pair, provide a plausible explanation for the observed sequence dependent stability.

## **Spatial Geometries of Dangling Ends**

The unpaired nature of the dangling end may allow for fewer geometric constraints and greater freedom in spatial orientation. While identities of the terminal base pair and dangling base provide plausible explanations for how stabilization could occur, spatial relationships provide evidential verification. Through X-ray and NMR studies of dangling ends adjoining the terminal base pair, no correlations were found between the sequence dependent gain in thermodynamic stability from a dangling end base and the intrastrand, interstrand, or total stacking areas, twist, slide, shift, xy, dx, rise, chi torsion or phase angle. Instead, the largest effect on duplex stability was found to be upon which end of the duplex (5' or 3') the dangling end resides<sup>22</sup>.

### **1.3 Hybridization Applications**

For applications in duplex hybridization it is essential to have a quantitative understanding of the effects of dangling ends on duplex stability. Hybridization reactions utilize a probe-target binding scheme. In these applications, the ability to predict how binding events will affect the thermodynamics of duplex formation is of extreme importance<sup>23</sup>. After selection of a target, alignment routines based on predictions of thermodynamic stability are employed to optimize probe design. An optimum probe-target binding scheme requires precise determination of the sequence specific stability of potential complexes that can form, including imperfect binding events<sup>24</sup>. To eliminate unwanted effects that can lead to erroneous and

ambiguous results, care is taken to carefully predict the thermodynamics of all desired (and undesired) pairing reactions.

Multiplex reactions use many probes designed to capture many targets, which may not be of the same length. Predictions for a system containing differing length probe and target molecules require the addition of various two-strand alignment schemes. Depending on the relative lengths and alignments, the resultant duplex can have two 5' dangling ends, two 3' dangling ends or one 5' and one 3' dangling end. While it is straightforward to calculate the thermodynamic stability of the duplex region given the sequence and n-n parameters, the model is only as good as the parameters it uses, thus it is essential to have a thorough understanding of the effects of dangling ends on duplex stability.<sup>25</sup>

Primary aims of this study were to obtain a better quantitative understanding of the contributions of single strand dangling-ends to the thermodynamic stability of short duplex DNA. To date, studies of DNA dangling ends that have been performed vary in the types of molecules that were examined. The influence of dangling ends on either end, especially as a function of length, is not well understood. To provide more precise predictive parameters, we systematically investigated the effects of length on duplex molecules, as a function of salt, having two 5' or one 5' and one 3' dangling ends.

## **Chapter 2: Materials and Methods**

### **2.1 DNA**

#### **Sequence Design**

DNA molecules that were designed and studied to evaluate the thermodynamic contributions of single strand dangling ends to short duplex DNA stability are grouped into four sets. Explicit sequences for the DNA molecules that were prepared and examined are shown in Tables 1-4. As depicted in Figure 1 duplexes with different types of single strand dangling ends were studied. The set I molecules shown in Table 1 are duplexes comprised of two annealed 25-base single strands. Sequence design was such that the two strands associate to form 10 different dangling ended duplexes having 24 to 10 base pairs with 1 to 10 bases as single strand 5' dangling ends. While strand length is constant, the duplex region decreases incrementally with a mirrored increased in length of the 5' single strand dangling ends. End sequence identities and the terminal stack (terminal base pair and first dangling base), remained the same for all molecules.

The set II molecules, shown in Table 2, were formed by annealing two single strands of different lengths to form duplexes ranging in length from 24 to 17 base pairs, each with a 5' and 3' dangling end. For the set II molecules, as the duplex length decreases, the length of the 5' and 3' dangling ends increases incrementally from one to eight bases. Duplex regions for the set II molecules are exactly the same as the corresponding duplex regions in the set I molecules.

DNA molecules designated as set IV in Table 3 are the blunt-ended duplex controls, ranging in length from 25 to 15 base pairs with the same duplex sequences as the corresponding molecules of the same length from sets I and II. Set III contains the half molecules matching set I and set II where the duplex is 21 base pairs and the dangling ends have four bases. Each molecule in set III (Table 3) has one single strand dangling-end and the duplex sequence of the corresponding molecules in sets I and II.


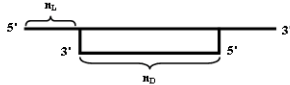
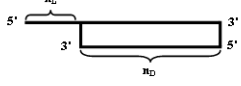
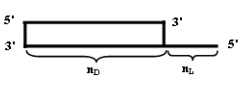
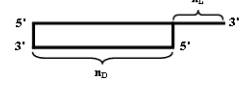
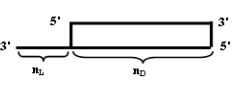

Set I		$\begin{array}{c cccccccccccccccc} n_D & 25 & 24 & 23 & 22 & 21 & 20 & 19 & 18 & 17 & 16 & 15 \\ \hline n_L & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{array}$
Set II		$\begin{array}{c ccccccc} n_D & 24 & 23 & 22 & 21 & 19 & 17 \\ \hline n_L & 1 & 2 & 3 & 4 & 6 & 8 \end{array}$
Set III	<div>(a) </div> <div>(b) </div> <div>(c) </div> <div>(d) </div>	$\begin{array}{c c} n_D & 21 \\ \hline n_L & 4 \end{array}$
Set IV		$\begin{array}{c cccccc} n_D & 24 & 23 & 22 & 21 & 17 & 16 \\ \hline n_L & 0 & 0 & 0 & 0 & 0 & 0 \end{array}$

Figure 1: Schematic representation of the various dangling ended DNAs that comprise the study.



Table 1: Duplexes with 5' dangling ends on both strands of the duplex. The terminal stack identity is retained for all molecules (5'AC/3'G).

**Table 1** Set I sequences, 5' overhangs

Sequences	Duplex Length ( $n_D$ )	Overhang Length ( $n_L$ )
5' -CATCATCGAACTCAGGTCTCACTTG GTAGTAGCTTGAGTCCAGAGTGAAC	25	0
5' - <b>AC</b> ATCATCGAACCAGGTCTCACTTG GTAGTAGCTTGGTCCAGAGTGAACA <b>A</b>	24	1
5' - <b>GAC</b> ATCATCGACCAGGTCTCACTTG GTAGTAGCTGGTCCGAAGTGAAC <b>AG</b>	23	2
5' - <b>AGACA</b> AGTGATCGCTGAGAGAGTTG GTTCACTAGCGACTCTCTCAAC <b>AGA</b>	22	3
5' - <b>TAGAC</b> ATGTGATCGCTAGTGAGATG GTACACTAGCGATCACTCTAC <b>AGAT</b>	21	4
5' - <b>CTAGAC</b> ATCTCACAGCGATCACTTG GTAGAGTGTGCTAGTGAAC <b>AGATC</b>	20	5
5' - <b>CCTAGAC</b> ATCTCACAGCGTAACTTG GTAGAGTGTGCGATTGAAC <b>AGATCC</b>	19	6
5' - <b>TCCTAGAC</b> ATCACAGGCGTTACATG GTAGTGTCCGCAATGTAC <b>AGATCCT</b>	18	7
5' - <b>ATCCTAGAC</b> ATCCGACTCTGCAATG GTAGGCTGAGACGTTAC <b>AGATCCTA</b>	17	8
5' - <b>GATCCTAGAC</b> ATTCTGAAGTCCAGTG GTAAGCTTCAGGTCAC <b>AGATCCTAG</b>	16	9
5' - <b>TGATCCTAGAC</b> ATAGCTGCACGTTG GTATCGACGTGCAAC <b>AGATCCTAGT</b>	15	10

Table 2: Duplexes with one 3' and one 5' dangling end on a single strand. Stack identity is conserved for 5' ends as (5'AC/3'G) and for 3' ends as (5'C/3'AG).

**Table 2** Set II sequences, 3'/5' overhangs

Sequences	Duplex Length ( $n_D$ )	Overhang Length ( $n_L$ )
5' - <b>ACATCATCGAACCAGGTCTCACTTGA</b> GTAGTAGCTTGGTCCAGAGTGAAC	24	1
5' - <b>GACATCATCGACCAGGCTTCACTTGAG</b> GTAGTAGCTGGTCCGAAGTGAAC	23	2
5' - <b>AGACAAGTGATCGCTGAGAGAGTTGAGA</b> GTTCACTAGCGACTCTCTCAAC	22	3
5' - <b>TAGACATGTGATCGCTAGTGAGATGAGAT</b> GTACACTAGCGATCACTCTAC	21	4
5' - <b>CCTAGACATCTCACAGCGTAACTTGAGATCC</b> GTAGAGTGTCGCATTGAAC	19	6
5' - <b>ATCCTAGACATCCGACTCTGCAATGAGATCCTA</b> GTAGGCTGAGACGTTAC	17	8

Table 3: Duplexes with one 3' or 5' dangling end and one blunt end on each molecule.

**Table 3** Set III sequences, single overhang

Sequences	Duplex Length ( $n_D$ )	Overhang Length ( $n_L$ )
5' - <b>TAGACATGTGATCGCTAGTGAGATG</b> GTACACTAGCGATCACTCTAC	21	4 (Set III a)
5' - CATGTGATCGCTAGTGAGATG <b>TAGA</b> GTACACTAGCGATCACTCTAC	21	4 (Set III c)
5' - CATGTGATCGCTAGTGAGATG GTACACTAGCGATCACTCTAC <b>AGAT</b>	21	4 (Set III b)
5' - CATGTGATCGCTAGTGAGATG <b>TAGA</b> GTACACTAGCGATCACTCTAC	21	4 (Set III d)
5' - CATGTGATCGCTAGTGAGATG GTACACTAGCGATCACTCATC	21	0 (Set IV)

Table 4: Blunt-ended control molecules. Duplex regions correspond to those of the set I and set II molecules.

<b>Table 4 Set IV sequences, controls</b>		
<b>Sequences</b>	<b>Duplex Length (<math>n_D</math>)</b>	<b>Overhang Length (<math>n_L</math>)</b>
5' -CATCATCGAACTCAGGTCTCACTTG GTAGTAGCTTGAGTCCAGAGTGAAC	25	0
5' -CATCATCGAACCAGGTCTCACTTG GTAGTAGCTTGGTCCAGAGTGAAC	24	0
5' -CATCATCGACCAGGCTTCACTTG GTAGTAGCTGGTCCGAAGTGAAC	23	0
5' -CAAGTGATCGCTGAGAGAGTTG GTTCACTAGCGACTCTCTCAAC	22	0
5' -CATGTGATCGCTAGTGAGATG GTACACTAGCGATCACTCATC	21	0
5' -CATCTCACAGCGATCACTTG GTAGAGTGTCGCTAGTGAAC	20	0
5' -CATCTCACAGCGTAACTTG GTAGAGTGTCGCATTGAAC	19	0
5' -CATCACAGGCGTTACATG GTAGTGTCGCAATGTAC	18	0
5' -CATCCGACTCTGCAATG GTAGGCTGAGACGTTAC	17	0
5' -CATTTCGAAGTCCAGTG GTAAGCTTCAGGTCAC	16	0
5' -CATAGCTGCACGTTG GTATCGACGTGCAAC	15	0

## DNA Molecules

DNA strands used to prepare all duplex molecules studied were purchased from Integrated DNA Technologies (IDT, Coralville, IA) and received following the standard desalting protocol performed by the supplier. As part of the design process all sequences were inspected for potential intramolecular hairpin formation using the IDT oligoanalyzer<sup>26</sup>. Acceptable requirements for designed sequences were that they supported only low stability intermolecular or intramolecular structures predicted to be unstable above 20°C. Sequence composition of the duplex region

was maintained around 50% GC base pairs and varied as little as possible between molecules to minimize for possible sequence dependent variations in thermodynamic stability.

## **2.2. Differential Scanning Calorimetry**

### **Introduction**

Thermodynamic parameters,  $\Delta H_{cal}$  and  $\Delta S_{cal}$ , of the heat induced melting transitions of duplex DNAs were evaluated from  $\Delta C_p$  versus temperature measurements, acquired as melting curves of the samples measured by differential scanning calorimetry (DSC). In a DSC melting curve experiment, the excess heat capacity,  $\Delta C_p$ , is continuously monitored over a selected temperature range. The DSC melting curve is a plot of measured  $\Delta C_p$  values versus temperature. Measurements of  $\Delta C_p$  as a function of temperature were made using a Nano-differential Scanning Calorimeter (Calorimetry Sciences Corporation, Provo, UT). The instrument utilizes a two-cell design and electronic comparison scheme to determine excess heat capacity,  $\Delta C_p^{ex}$ , of the sample as it is heated.

### **DSC Sampling**

Concentrations of DNA samples for DSC measurements were maintained

around 1.0 mg/ml (duplex concentration determined from using  $A_{260}$  reading). Prior to loading, DNA samples were degassed by bubbling with a fine stream of helium gas for 10 minutes. Measurements were generally collected over the temperature range from 10 to 120 °C. Experiments were conducted under positive pressure of three atmospheres with a heating rate of 2.0 °C/min. For all samples, multiple heating and cooling curves were collected, and the averaged data and associated errors were used for further analysis. Multiple but identical DSC instruments were used. To ensure machine reproducibility, sample melting curves were commonly collected from different instruments and results compared. All DNA melting curves, regardless of instrument or direction of scan, were highly reproducible.

Prior to melting a sample, the buffer-versus-buffer melting curve was measured over the same temperature range, using the same heating rate and  $[Na^+]$  buffer as the sample. This curve provided the standard baseline used in conjunction with the instrument software to standardize DSC melting curves for all DNA samples melted in that buffer. The baseline correction removes undesirable machine or other non-sample noise from the melting curve. A representative DSC experimental set and analysis is shown in Figure 2.

### **DSC Data Analysis**

Analysis of DSC melting data was performed using the CpCalc 2.1 software package supplied by the manufacturer for use with the DSC instrument (Applied Thermodynamics, Middlesex, NY). The buffer-versus-buffer baseline was used to

correct sample-versus-buffer scans by subtraction of the baseline from the average sample scan (Figure 2a). Resulting baseline-corrected,  $\Delta C_p^{ex}$  versus temperature curves were normalized for total DNA strand concentration, molecular mass, and cell volume. Values for DNA concentration (mg/mL), sample cell volume (0.3268 mL), and partial specific volume of the macromolecule (0.55 cm<sup>3</sup>/g) were input (Figure 2b). To calculate the thermodynamic parameters of the resultant standardized baseline-corrected curve, a progressive polynomial line was fit to connect linear regions in the lowest and highest temperature portions of the curve (Figure 2c). Integration of the area under the baseline-corrected curve provided a measurement of the calorimetric transition enthalpy,  $\Delta H_{cal}$ , given by

$$\Delta H_{cal} = \int_{T_1}^{T_2} \Delta C_p(T) dT \quad (1)$$

Where  $T_1$  and  $T_2$  are the beginning and ending temperatures of the DSC melting curve defining the temperature range examined. Individual temperature integration ranges varied and were chosen to provide a best fit to the linear regions around the melting curve as seen when using the integration algorithm (Figure 2d). The corresponding calorimetric entropy,  $\Delta S_{cal}$ , was determined by dividing  $\Delta C_p(T)$  by the temperature and integrating over the temperature range.

$$\Delta S_{cal} = \int_{T_1}^{T_2} \frac{\Delta C_p(T)}{T} dT \quad (2)$$

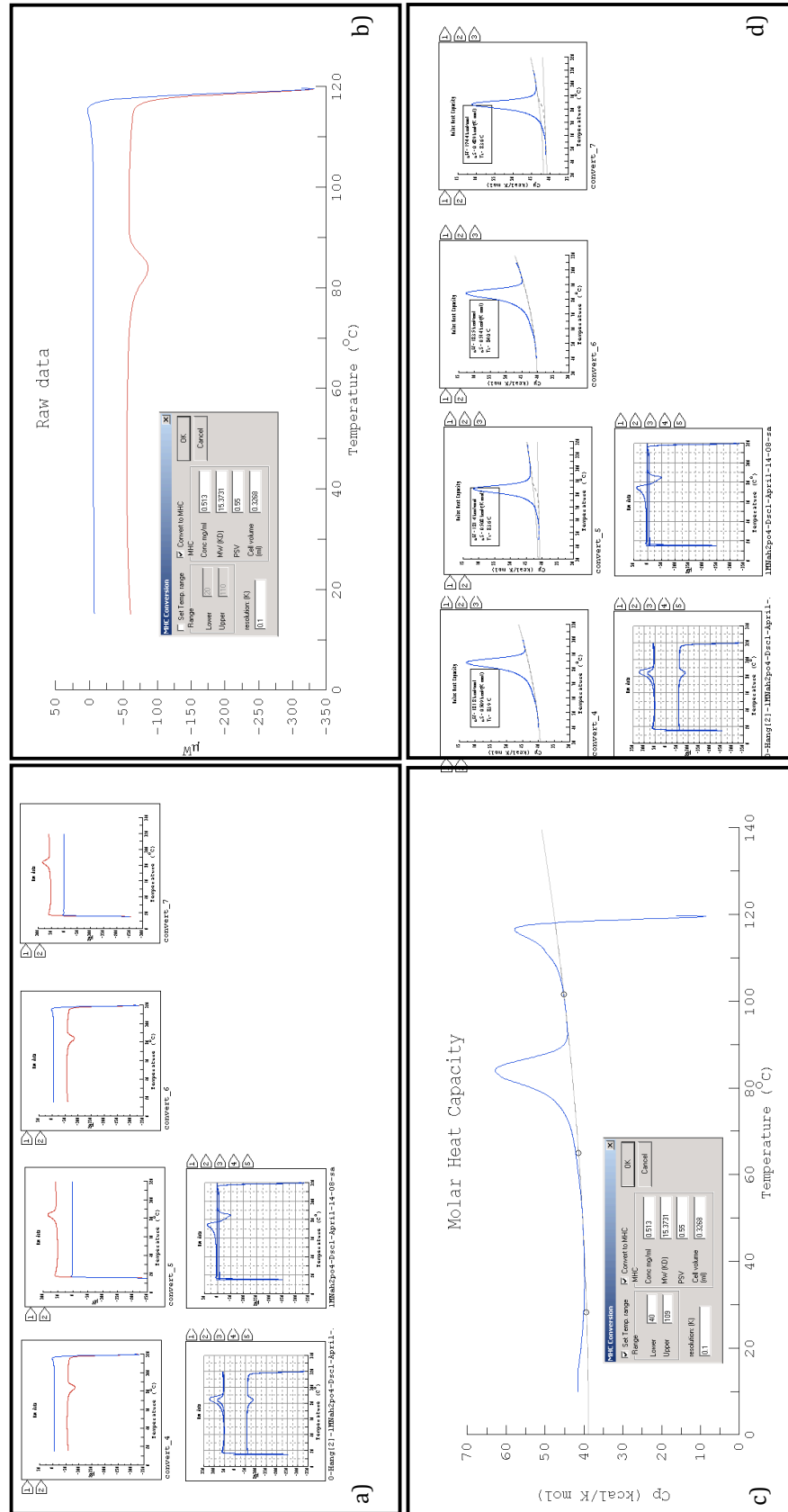
In the analyses that were performed, it was assumed in all cases the melting

transition occurs in an all-or-none, two-state manner. The transition temperature,  $T_m$ , is the temperature of the peak height maximum on the  $\Delta C_p$  versus temperature, DSC melting curve and represents the state in which half of the molecules are in their melted single strand form<sup>27</sup>. Reported values of the calorimetric free-energy,  $\Delta G_{cal}(T)$ , were determined at  $T = 298.15$  K by the Gibb's relation,

$$\Delta G_{25} = \Delta H_{cal} - T \Delta S_{cal} \quad (3)$$

The analysis follows the standard assumption that the overall difference in excess heat capacity from the beginning to the end of the melting transition is negligibly small, (i.e.  $\Delta C_p(T_2) - \Delta C_p(T_1) = 0$ )<sup>15</sup>. Estimates on the validity of this assumption can be made from analysis of variations of the evaluated enthalpy,  $\Delta H_{cal}$  and entropy  $\Delta S_{cal}$  with the measured transition temperature,  $T_m$ , for the molecules. The consequence of assuming  $\Delta C_p(T_2) - \Delta C_p(T_1) = 0$  is that evaluated thermodynamic parameters are most accurate in the transition region, and more importantly, that these parameters are assumed to be temperature independent.

Figure 2: Typical DSC analysis. (a) Multiple sample-buffer curves are collected and matched to a buffer-buffer curve. (b) Sample values of molecular weight, concentration, specific macromolecule volume and cell volume are input and the buffer curve is subtracted. (c) A baseline is fit to the upper and lower linear regions. (d) Thermodynamic parameters are collected and averaged.





## **2.3 Solutions and Preparations**

### **Buffer Solutions**

The buffer used in all melting experiments was sodium phosphate ( $\text{Na}_2\text{HPO}_4$ ) combined with varying amounts of sodium chloride,  $\text{NaCl}$ . Disodium ethylenediaminetetraacetic acid ( $\text{Na}_2\text{EDTA}$ ) was added to chelate trace metals and multivalent cations that could bind to the DNA and possibly influence the melting process. This buffer was chosen largely because during the course of a melting experiment, changes in  $\text{pK}_a$  with temperature are known to be negligibly small<sup>12,28</sup>.

Three different Sodium buffer solutions were used containing 85mM, 300mM and 1000 mM  $[\text{Na}^+]$  in 10 mM  $\text{Na}_2\text{PO}_4$ , 0.1mM EDTA. Solution pH was adjusted to between 7.2 and 7.4 using an Orion 4-Star pH/ Conductivity Meter (Thermo Electron Corporation, Beverly, MA).

### **Quantification of Samples**

Samples were ordered from IDT and received in single strand, dehydrated form. Prior to opening the sample tube it was spun in a bench top centrifuge at 4000 rpm for 2-5 minutes to ensure pellet presence and location. After centrifugation two mL of 85 mM  $\text{Na}^+$  buffer was added and the DNA sample was allowed to dissolve for at least one hour at room temperature. Concentrations of diluted single strand DNA solutions were determined from measured values of the absorbance at 260 nm,  $A_{260}$ , and appropriate values of the molar extinction coefficient provided by the supplier.

Absorbance measurements were made using a Hewlett-Packard 8452A Diode Array Spectrophotometer (Hewlett-Packard Corporation, Palo Alto, CA). Quartz cuvettes with path lengths of 1.0 cm were used. DNA samples were diluted 1:100 with buffer such that  $A_{260}$  values usually ranged from 0.2 – 0.9 OD. Single strands were mixed in a 1:1 ratio and the final duplex concentration was adjusted with buffer to be 1.0 mg/ml. All DNA samples were melted at the same duplex concentration

Duplexes were annealed at room temperature for at least one hour prior to characterization. After initial mixing, the absorbance at 260 nm was noted and compared to the absorbance at the high end of the temperature range after the melting transition. This value was then used to determine an accurate molar extinction coefficient for the duplex,  $\epsilon_{dup}$ . Subsequent concentrations were then found from the pre/post melting absorbances and  $\epsilon_{dup}$ .

To check for possible sample degradation, comparison of the pre and post-transition  $A_{260}$  values for the duplex were used. Sample quality was also assessed using vertical polyacrylamide gel electrophoresis (PAGE). Typical gels were 12% polyacrylamide.

### **Buffer Exchange**

Three buffer solutions of varying ionic strength were used. To change the solution buffer, samples were transferred to a DNA Centricon YM-3 or Amicon Ultra-4 centrifugal filter (Millipore, Bedford, MA) with a molecular weight cut off of 3,000

Daltons. Once transferred to the spin tube, samples were washed with 2 ml of nanopure water and then spun in an Eppendorf centrifuge for approximately 90 minutes at 4000-4500 rpm. After this time 1 ml additional nanopure water was added and the sample was spun to dryness.

Recovered samples were reconstituted in the desired buffer solution, and sample concentrations after buffer changes were determined by optical absorbance measurements at 260 nm. Typically, sample recovery after buffer exchange was greater than 95%.

### **Gel Characterization**

To ensure that melting did not degrade the sample, a characterizing gel was used. Samples were checked after initial melting and then select samples were tested at random throughout the study to determine if any decomposition had taken place. Electrophoresis was performed using a Hoefer MiniVE, vertical electrophoresis mini gel system (Hoefer, Inc, Holliston, MA). Constant voltage in the range of 100 - 150 V was supplied from a POWER-PAC300™ (BIO-RAD Company, Hercules, CA) power supply. Approximately 3µg of total DNA was suspended with 1.5µl TBE buffer and 1µl of indicator (by weight, 0.25% Bromophenol Blue and 0.25% Xylene Cyanol) and loaded onto a 12% polyacrylamide gel. Gels were subjected to electrophoresis until the Blue Juice™ indicator migrated about three quarters down the length of the gel. Gels were stained in Stains-All™ solution, for five to eight hours, followed by approximately one hour of destaining in nanopure

water. Stained gels were scanned using a HP Scanjet photo scanner.

## Chapter 3: Results

### 3.1 Dangling End Thermodynamics

Measured thermodynamic parameters for melting of the 27 duplexes studied are listed in Tables 5-7 according to their respective set designation and  $[\text{Na}^+]$  in which they were evaluated. In all salts, set IV values for  $n_L=5-7$  and 10 were extrapolated from a best-fit line of the measured points.

Table 5: Thermodynamic melting parameters for the set I, set II and set IV duplexes evaluated in 85 mM  $[\text{Na}^+]$ . Average errors:  $\Delta H$ : 3.9 kcal/mol,  $\Delta S$ : 11.8 e.u.,  $\Delta G$ : 0.8 kcal/mol.

	Overhang Length	Duplex Length	$\Delta H$ kcal/mol	$\Delta S$ e.u.	$\Delta G_{25}$ kcal/mol
Set I 5' overhangs	0	25	-186.9	-530.8	-28.6
	1	24	-185.0	-525.8	-28.2
	2	23	-179.3	-510.8	-27.0
	3	22	-171.9	-491.5	-25.4
	4	21	-161.5	-459.1	-24.6
	5	20	-161.5	-464.8	-22.9
	6	19	-151.3	-444.8	-18.7
	7	18	-141.1	-414.3	-17.6
	8	17	-135.2	-397.3	-16.7
	9	16	-120.3	-356.5	-14.0
Set II 3'/5' overhangs	10	15	-117.6	-347.3	-14.1
	0	25	-186.9	-530.8	-28.6
	1	24	-180.0	-520.0	-24.9
	2	23	-157.2	-457.0	-21.0
	3	22	-155.6	-454.8	-20.0
	4	21	-140.8	-416.7	-16.5
	6	19	-117.0	-345.3	-14.1
	8	17	-102.9	-304.0	-12.2
Set IV Controls	0	25	-186.9	-530.8	-28.6
	0	24	-171.3	-500.0	-22.2
	0	23	-159.2	-462.5	-21.3
	0	22	-151.9	-444.5	-19.4
	0	21	-142.3	-410.9	-19.8
	0	20	-138.7	-406.0	-17.7
	0	19	-130.2	-381.9	-16.4
	0	18	-121.8	-357.9	-15.1
	0	17	-109.7	-322.3	-13.6
	0	16	-103.3	-305.8	-12.1
	0	15	-96.3	-285.7	-11.1

Table 6: Thermodynamic parameters for the set I, set II and set IV duplexes evaluated in 300 mM [Na<sup>+</sup>]. Average errors:  $\Delta H$ : 4.5 kcal/mol,  $\Delta S$ : 12.9 e.u.,  $\Delta G$ : 0.7 kcal/mol.

	Overhang Length	Duplex Length	$\Delta H$ kcal/mol	$\Delta S$ e.u.	$\Delta G_{25}$ kcal/mol
Set I 5' overhangs	0	25	-192.3	-544.3	-30.0
	1	24	-191.1	-542.5	-29.3
	2	23	-183.0	-519.3	-28.2
	3	22	-180.5	-515.5	-26.8
	4	21	-163.2	-457.9	-26.7
	5	20	-148.7	-429.5	-20.7
	6	19	-141.0	-407.3	-19.5
	7	18	-137.4	-396.5	-19.2
	8	17	-135.8	-392.8	-18.7
	9	16	-113.5	-331.3	-14.8
	10	15	-113.3	-329.3	-15.2
Set II 3'/5' overhangs	0	25	-192.3	-544.3	-30.0
	1	24	-182.3	-516.7	-28.2
	2	23	-160.2	-454.3	-24.7
	3	22	-161.9	-460.0	-24.7
	4	21	-152.7	-434.1	-23.2
	6	19	-117.6	-339.5	-16.3
	8	17	-117.1	-340.0	-15.8
Set IV Controls	0	25	-192.3	-544.3	-30.0
	0	24	-169.6	-481.8	-25.9
	0	23	-146.5	-414.8	-22.8
	0	22	-149.4	-426.8	-22.1
	0	21	-152.5	-440.0	-21.3
	0	20	-137.3	-394.3	-19.8
	0	19	-129.2	-372.2	-18.2
	0	18	-121.1	-350.2	-16.7
	0	17	-122.2	-353.3	-16.8
	0	16	-91.9	-269.5	-11.6
	0	15	-96.8	-284.0	-12.2

Table 7: Thermodynamic parameters for the set I, set II and set IV duplexes evaluated in 1.0 M [Na<sup>+</sup>]. Average errors:  $\Delta H$ : 4.3 kcal/mol,  $\Delta S$ : 12.3 e.u.,  $\Delta G$ : 0.7 kcal/mol.

	Overhang Length	Duplex Length	$\Delta H$ kcal/mol	$\Delta S$ e.u.	$\Delta G_{25}$ kcal/mol
Set I 5' overhangs	0	25	-205.2	-578.9	-32.6
	1	24	-190.2	-535.0	-30.7
	2	23	-184.6	-517.0	-30.4
	3	22	-182.8	-513.8	-29.6
	4	21	-169.6	-469.8	-29.5
	5	20	-162.8	-460.0	-25.7
	6	19	-153.0	-440.0	-21.8
	7	18	-144.7	-412.5	-21.7
	8	17	-137.0	-390.0	-20.7
	9	16	-129.2	-371.3	-18.5
Set II 3'/5' overhangs	10	15	-121.5	-348.3	-17.7
	0	25	-205.2	-578.9	-32.6
	1	24	-181.8	-510.0	-29.7
	2	23	-175.8	-490.5	-29.6
	3	22	-174.8	-490.3	-28.6
	4	21	-156.8	-447.2	-23.5
	6	19	-131.5	-380.0	-18.2
	8	17	-125.2	-360.0	-17.8
Set IV Controls	0	25	-205.2	-578.9	-32.6
	0	24	-181.6	-508.3	-30.0
	0	23	-157.5	-440.3	-26.2
	0	22	-153.9	-433.3	-24.8
	0	21	-156.3	-444.0	-23.9
	0	20	-137.2	-389.3	-21.1
	0	19	-129.5	-368.9	-19.5
	0	18	-121.8	-348.4	-18.0
	0	17	-121.0	-346.8	-17.6
	0	16	-98.8	-286.8	-13.3
	0	15	-98.8	-287.0	-13.2

## Enthalpy-Entropy Compensation

Melting thermodynamic data evaluated for all duplex DNAs of the study are shown in Figure 3 where the transition enthalpy,  $\Delta H_{cal}$ , is plotted versus  $T\Delta S_{cal}$  for all [Na<sup>+</sup>]. As can be seen, these plots are well fit by a straight line with mean correlation coefficient,  $R^2 = 0.96$ . Such high linearity is typically observed for short duplex DNA melting and indicative of entropy-enthalpy compensation in the melting

of these duplexes. Although fundamentally not fully understood, such behavior is reportedly typical of the melting of short duplex DNA<sup>6,29,30</sup>. At this resolution of the data there is apparently only a slight dependence of the slopes and intercepts of these linear fits on the  $[\text{Na}^+]$ .

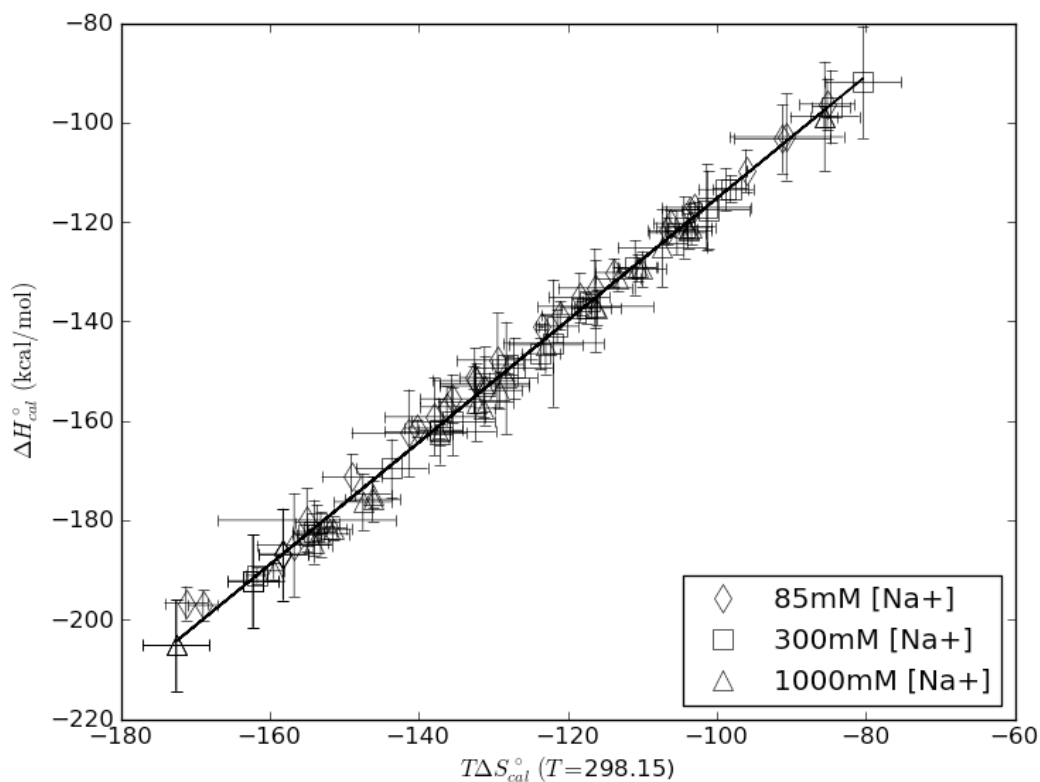


Figure 3: Plots of  $\Delta H_{cal}^{\circ}$  versus  $T\Delta S_{cal}^{\circ}$  ( $T = 298.15$  K) for the 27 duplex DNAs in 85 mM, 300 mM and 1.0 M  $[\text{Na}^+]$ .



## Parsing Assumptions

From the data displayed in Tables 5-7, quantitative estimates can be made on the thermodynamic contributions of 5' and 3' dangling ends to duplex stability. With the n-n parameters, the melting thermodynamics for short perfect match (blunt ended) duplexes can be accurately predicted<sup>4, 31</sup>. Inclusions of the single strand dangling end contributions are straightforward. The n-n model assumes the duplex region of the dangling ended molecule has the same calculated thermodynamic stability as the blunt ended duplex. Additional contributions from the terminal stack are then added to the blunt ended duplex calculation. Thus, it is assumed the thermodynamic melting transition parameters of short duplex DNAs can be reliably parsed into two distinct individual contributions, i.e. those from the duplex region, and those from any dangling end regions. By maintaining the duplex region and varying the ends, it should be possible to dissect out the end contributions. This is the basis of the analysis that was performed as described below. Note, reliability of the procedure assumes the interactions in the duplex and single strand regions are separable and parsed as described, and that the duplex region and its inherent structure and stability is unaffected (in a thermodynamic sense) by the dangling end or ends. Validity and implications of this assumption will be examined in the discussion presented in Chapter 4.

## Thermodynamic Formulation

Given the set of DNA molecules and thermodynamic parameters evaluated, and under the umbrella of the aforementioned assumptions, the following descriptions can be written for the molecules of sets I-IV depicted in Tables 1-4 and Figure 1.

As shown in Figure 1, the set I molecules contain  $n_D$  base pairs and  $n_L$  base dangling-ends. Since the resulting duplexes were formed from the designed annealing of two 25 base single strands, the values of  $n_D$  and  $n_L$  are coupled, i.e. the smaller  $n_D$  the larger  $n_L$ . In general, the experimentally determined values of the melting transition thermodynamic parameters, the DSC measured enthalpy, entropy and free energy at 25°C, can be represented as  $\Delta X_{cal}^I \equiv \Delta H_{cal}^I, \Delta S_{cal}^I$ , or  $\Delta G_{25cal}^I$ , respectively. If the measured parameters,  $\Delta X_{cal}^I(n_D, n_L)$ , for the set I molecules arise from the individual energies of the duplex region and the two dangling end regions, the set I molecule with two identical 5'-dangling ends can be characterized as,

$$\Delta X_{cal}^I(n_D, n_L) = \Delta X_{cal}^{IV}(n_D) + 2\delta X_{ss}^{5'}(n_L) \quad (4)$$

Where  $\Delta X_{cal}^{IV}(n_D)$  is the measured thermodynamic parameter for the blunt-ended duplex molecule with the same  $n_D$  base pairs; and  $\delta X_{ss}^{5'}(n_L)$  is the thermodynamic contribution of a 5' single strand dangling end comprised of L bases. This term is taken twice because there are two 5' dangling ends in each set I molecule. In this development the explicit sequence dependence of the ends is not considered, but for

internal consistency the terminal stack identities are the same in all molecules. Furthermore, across all sets at each  $n_D$ , duplex sequences were the same.

The set II molecules differ from those in set I, in that they have only one 5' dangling end in addition to a 3' dangling end. Modifying equation (4) from above, the measured parameters,  $\Delta X_{cal}^{II}(n_D, n_L)$ , for the set II molecules with  $n_D$  duplex base pairs, and  $n_L$  single strand bases, can be expressed as follows,

$$\Delta X_{cal}^{II}(n_D, n_L) = \Delta X_{cal}^{IV}(n_D) + \delta X_{ss}^{5'}(n_L) + \delta X_{ss}^{3'}(n_L) \quad (5)$$

Where again,  $\Delta X_{cal}^{IV}(n_D)$  is the measured thermodynamic parameter for the blunt-ended duplex molecule with  $n_D$  base pairs;  $X_{ss}^{5'}(n_L)$  is the thermodynamic contribution of a 5' single strand dangling-end comprised of  $n_L$  bases and  $\delta X_{ss}^{3'}(n_L)$  is the thermodynamic contribution of a 3' single strand dangling-end comprised of  $n_L$  bases.

From the above expressions and the experimentally measured parameters in Tables 5-7, individual thermodynamic contributions of the ends  $\delta X_{ss}^{5'}(n_L)$  and  $\delta X_{ss}^{3'}(n_L)$ , can be determined from equations (4) and (5), viz.

From Eqn (4),

$$\delta X_{ss}^{5'}(n_L) = [\Delta X_{cal}^{II}(n_D, n_L) - \Delta X_{cal}^{IV}(n_D)] / 2 \quad (6)$$

With this and Eqn (5),

$$\delta X_{ss}^{3'}(n_L) = \Delta X_{cal}^{II}(n_D, n_L) - \Delta X_{cal}^{IV}(n_D) - \delta X_{ss}^{5'}(n_L) \quad (7)$$

The values of  $\delta X_{ss}^{5'}(n_L)$  and  $\delta X_{ss}^{3'}(n_L)$  evaluated in this manner, for X = H (a), S (b) and G<sub>25</sub> (c) are plotted versus  $n_L$  in Figure 4. These series of plots were generated from the thermodynamic data measured for the set I, set II and set IV molecules in 85 mM, 300 mM and 1.0 M Na<sup>+</sup> as summarized in Tables 5-7. Filled symbols in Figure 4 are the values of  $\delta X_{ss}^{5'}(n_L)$  and open symbols are the values of  $\delta X_{ss}^{3'}(n_L)$  at each [Na<sup>+</sup>].

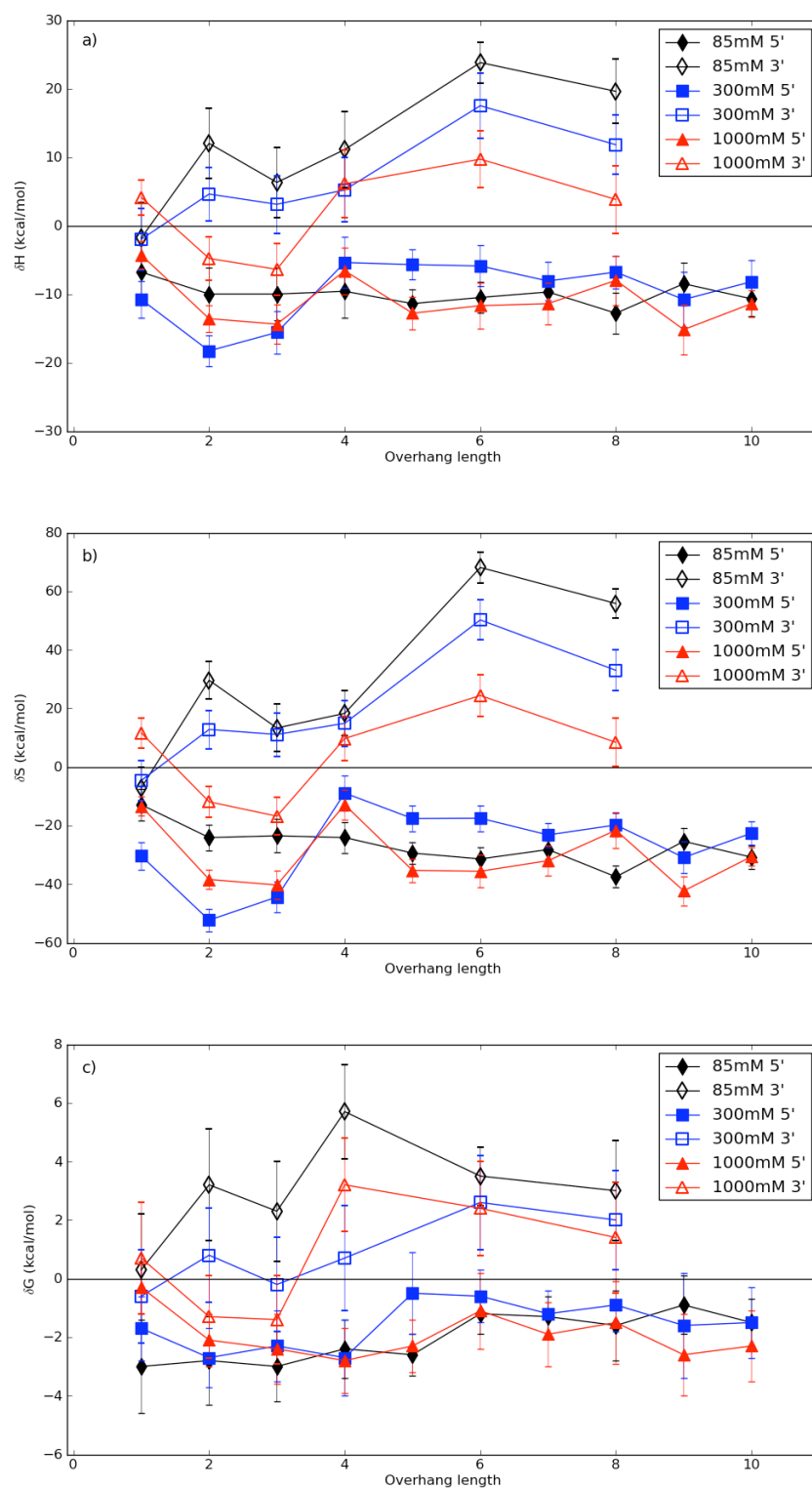


Figure 4: Comparison of the evaluated thermodynamic parameters  $\Delta H$  (a),  $\Delta S$  (b), and  $\Delta G$  (c) for the 5' and 3' dangling ends in sets I and II, in the three  $\text{Na}^+$  environments plotted versus overhang length,  $n_L$ .

## Thermodynamic Contributions From 5' Dangling Ends

In comparing the plots in Figure 4 several trends are clearly seen. The 5' ends ( $\delta X_{ss}^{5'}(n_L)$ ) are stabilizing and only weakly dependent on the  $[\text{Na}^+]$ . The values of  $\delta H_{ss}^{5'}(n_L)$  and  $\delta S_{ss}^{5'}(n_L)$  display similar trends, as a function of increasing end length,  $n_L$ , however the length dependence is not terribly strong. As  $n_L$  increases from one to three nucleotides, the values of  $\delta H_{ss}^{5'}(n_L)$  and  $\delta S_{ss}^{5'}(n_L)$  incrementally become more negative providing higher enthalpic and entropic stability to the duplex. For  $n_L \geq 4$ , these values level off and are essentially constant thereafter. The values in 300mM and 1.0 M  $\text{Na}^+$  are slightly higher than in 85 mM.

The behavior of evaluated free-energies,  $\delta G_{ss}^{5'}(n_L)$ , as a function  $n_L$  (Figure 4c) indicates, despite the varying  $\delta H_{ss}^{5'}(n_L)$  and  $\delta S_{ss}^{5'}(n_L)$  values for the different molecules, the parameters are apparently compensatory in such a way as to render  $\delta G_{ss}^{5'}(n_L)$  values essentially linear and slightly greater (negative) with increasing  $[\text{Na}^+]$ .

Thus to summarize, values of  $\delta G_{ss}^{5'}(n_L)$  as a function of  $n_L$  vary slightly with increasing  $n_L$ , but the dependence on  $[\text{Na}^+]$  is minimal and decreases with increasing  $[\text{Na}^+]$ . Taking the average of  $\delta G_{ss}^{5'}(n_L)$  over all values of  $n_L$ , in each  $[\text{Na}^+]$  environment, derived values for  $\delta G_{ss}^{5'}(n_L)$  were determined as summarized in Table 8. These can be used in the nearest-neighbor model to predict the thermodynamic

contributions of 5' dangling-ends to duplex stability.

### Thermodynamic Contributions from 3' Dangling Ends

In contrast, the 3' ends ( $\delta X_{ss}^{3'}(n_L)$ ) appear to be mainly destabilizing, and are more strongly effected by the  $[Na^+]$ , becoming less destabilizing with higher  $[Na^+]$ . Much like the 5' dangling end, the values of  $\delta H_{ss}^{3'}(n_L)$  and  $\delta S_{ss}^{3'}(n_L)$  display a mild trend as a function of increasing end length,  $n_L$ , only in the opposite direction (increasingly destabilizing). Here also, the length dependence of  $\delta X_{ss}^{3'}(n_L)$  is not terribly strong. Recall for the 5' end, when  $n_L$  increases from one to three nucleotides, the values of  $\delta H_{ss}^{5'}(n_L)$  and  $\delta S_{ss}^{5'}(n_L)$  follow a stabilizing trend, leveling off around  $n_L \geq 4$ . For the 3' ends in the 1 M  $Na^+$  conditions, this trend is conserved for  $n_L < 4$ . At  $n_L > 4$ , the values increase to become destabilizing and level off to a constant destabilization value at  $n_L \geq 5$ . In contrast, the values of  $\delta H_{ss}^{3'}(n_L)$  and  $\delta S_{ss}^{3'}(n_L)$  in 300 mM and 85 mM  $Na^+$  are initially near zero before they increase to a constant positive value for  $n_L = 2, 3$  and 4. After this they increase again and are essentially constant for  $n_L \geq 5$ .

The free energy,  $\delta G_{ss}^{3'}(n_L)$ , as a function of length, imparted by this type of end, increases to greater positive values with increasing end length up to  $n_L = 4$  and

is essentially constant thereafter. The Na<sup>+</sup> dependent stability is also much stronger and more pronounced than in their 5' counterparts. The lower Na<sup>+</sup> environment (85 mM) has an immediate destabilizing influence on duplex stability as reflected in the corresponding values of  $\delta G_{ss}^{3'}(n_L)$ . For 300 mM, no effect is found until  $n_L = 4$  at which point  $\delta G_{ss}^{3'}(n_L)$  becomes destabilizing. Conversely at 1.0 M Na<sup>+</sup>,  $\delta G_{ss}^{3'}(n_L)$  is initially stabilizing before reversing trends. Values for 300 mM and 1000 mM Na<sup>+</sup> are essentially the same after  $n_L = 4$ .

As described above the values of  $\delta G_{ss}^{3'}(n_L)$  display a more complicated dependence on  $n_L$  and [Na<sup>+</sup>] than their 5' analogues. Behavior of the data is reminiscent of a sigmoid growth function. Consequently for convenience in each Na<sup>+</sup> environment, we have parameterized a logistic equation to fit the data.

$$y = \left[ \frac{A}{1 + e^{-r(x-k)}} \right] + c \quad (8)$$

In this form, c shifts the curve vertically, and k horizontally while perpendicular rise is controlled by A and steepness by r. The curves were hand-fit to best describe the data. Values determined for the individual coefficients, c, k, r and A in each Na<sup>+</sup> environment are listed in the caption for Figure 5. These expressions can be used to predict the contributions of 3' dangling ends to the thermodynamic stability of duplex DNA. Values determined from these fits are summarized in Table 8.



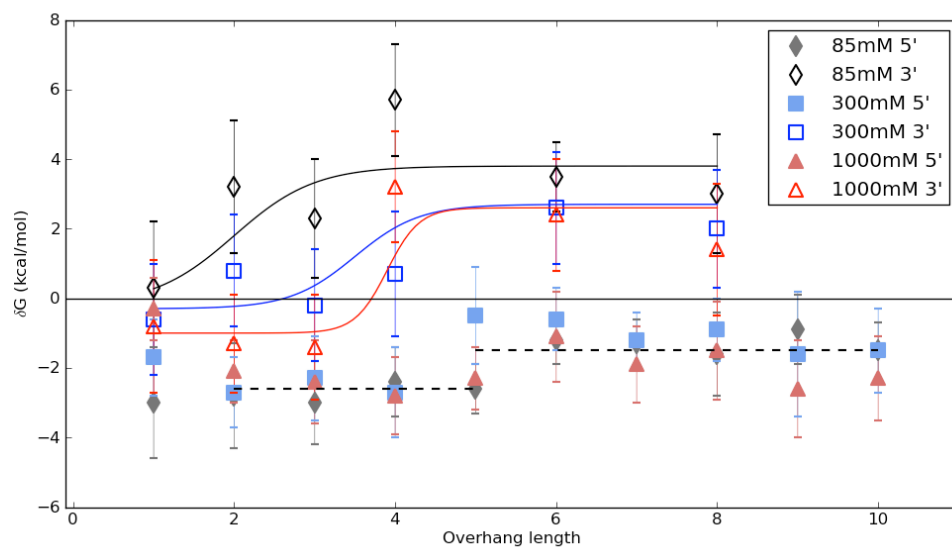


Figure 5: Plots of the relative free-energy contribution,  $\delta G$ , from the dangling ends versus overhang length ( $n_L$ ). Set I averages over all salts. Set II fits to logistical curve equation (8). Equation parameter values: (85 mM)  $A=4$ ,  $r=2$ ,  $c=-0.2$ ,  $k=1$ ; (300 mM)  $A=3$ ,  $r=2.5$ ,  $c=-0.3$ ,  $k=2$ ; (1 M)  $A=3.6$ ,  $r=5$ ,  $c=-1$ ,  $k=3.5$ . The value of  $\delta G_{ss}^{3'}(n_L = 1)$  for 100mM was adjusted within error parameters to improve fit.

Table 8: Calculated contributions for dangling ends determined from the analysis in Figure 5. Negative values are stabilizing and positive values are destabilizing for the duplex.

$\delta G_{25}(n_L)$ kcal/mol				
Contributions of 5' ends	$n_L$	85 mM	300 mM	1000 mM
	1	-3	-1.7	-0.3
	2-4	-2.6	-2.6	-2.6
	5	-2.6	-1.5	-1.5
	>5	-1.5	-1.5	-1.5
Contributions of 3' ends	1	1.8	-0.1	-1.0
	2	3.3	1.2	-1.0
	3	3.7	2.5	-0.7
	4	3.8	2.7	2.3
	6	3.8	2.7	2.6
	8	3.8	2.7	2.6

### 3.2 Validity Check For Parsing Assumptions

#### Difference Equations For $n_L=4$

The values given in Table 8 for  $\delta X_{ss}^{5'}(n_L)$  and  $\delta X_{ss}^{3'}(n_L)$  and their ultimate general utility in improving predictions of the thermodynamics for short duplex DNAs having single strand ends, relies on validity of the assumptions under which the parameters were evaluated. As stated previously evaluations of  $\delta X_{ss}^{5'}(n_L)$  and  $\delta X_{ss}^{3'}(n_L)$  are founded on the n-n model, in which the individual contributions of the ends can be parsed from those of the duplex region, Eqns (4) and (5). To test the validity of these underlying assumptions, the set III molecules were prepared and examined.

As seen by examination of Tables 1-3 and Figure 1, the molecules in set III are in a sense “half-molecules” of those in sets I and II with  $n_D=21$  and  $n_L=4$ . The set III molecules each contain one 5' or 3' dangling-end and one blunt end. The duplex and single dangling end sequences are the same as those in set I and set II. Through this design the set III molecules ( $n_D=21$ ) provide a secondary method of evaluating the parameters when  $n_L=4$ , i.e.  $\delta X_{ss}^{5'}(n_L=4)$  and  $\delta X_{ss}^{3'}(n_L=4)$ .

Utilizing equations (4) and (5), altered specifically for consideration of the set III molecules provides the following,

$$\Delta X_{cal}^{III}(n_D=21, n_L=4) = \Delta X_{cal}^{IV}(n_D=21) + \delta X_{ss}^{5'}(n_L=4) \quad (9a)$$

For the subset molecules a and b of set III having one 5' dangling-end and one blunt

end.

For the set c and d molecules of set III with one 3' dangling end and one blunt end,

$$\Delta X_{cal}^{III}(n_D = 21, n_L = 4) = \Delta X_{cal}^{IV}(n_D = 21) + \delta X_{ss}^{3'}(n_L = 4) \quad (9b)$$

Where  $\Delta X_{cal}^{IV}(n_D = 21)$  is the measured thermodynamic parameter for the blunt-ended duplex with 21 base pairs.

Using equations (9a) and (9b) and the pertinent results for sets I, II and IV (where  $n_D = 21, n_L = 4$ ) along with those from set III as summarized in Table 9, these molecules can then be used to obtain estimates on  $\delta X_{ss}^{5'}(n_L = 4)$  and  $\delta X_{ss}^{3'}(n_L = 4)$ .

Consider the following for  $\delta X_{ss}^{5'}(n_L = 4)$ ,

$$\delta X_{ss}^{5'}(n_L = 4) = \Delta X_{cal}^I(n_D = 21, n_L = 4) - \Delta X_{cal}^{IIa} \quad (10a)$$

or

$$\delta X_{ss}^{5'}(n_L = 4) = \Delta X_{cal}^I(n_D = 21, n_L = 4) - \Delta X_{cal}^{IIb} \quad (10b)$$

or

$$\delta X_{ss}^{5'}(n_L = 4) = \Delta X_{cal}^{II}(n_D = 21, n_L = 4) - \Delta X_{cal}^{IIIc} \quad (10c)$$

In a similar manner, consider the following for  $\delta X_{ss}^{3'}(n_L = 4)$ ,

$$\delta X_{ss}^{3'}(n_L = 4) = \Delta X_{cal}^{II}(n_D = 21, n_L = 4) - \Delta X_{cal}^{IIIa} \quad (11)$$

In total five equations to determine  $\delta X_{ss}^{5'}(n_L = 4)$  and three equations to determine  $\delta X_{ss}^{3'}(n_L = 4)$ , were generated. These values, calculated in each  $[Na^+]$ , are all displayed in Figure 6. The corresponding plots are given as  $[Na^+]$  increases from left

to right where  $\delta X = \delta H(a)$ ,  $\delta S(b)$ , and  $\delta G(c)$ . Stars on each Figure 6 plot denote values from Figure 4 at  $n_L = 4$ , on the right side for  $\delta X_{ss}^{3'}$  and the left side for  $\delta X_{ss}^{5'}$ .

Table 9: Thermodynamic data for Set III molecules.

[Na+]	Sequence	$\Delta H$	$\sigma$	$\Delta S$	$\sigma$	$\Delta G$	$\sigma$
<b>85mM</b>	Set I	-161.5	8.2	-459.1	9.3	-24.6	0.5
	Set II	-140.8	9.4	-416.7	15.7	-16.5	1.1
	Set III(a)	-153.4	11.9	-436.1	18.7	-23.4	0.5
	Set III(b)	-154.5	4.7	-440.2	19.4	-23.3	0.7
	Set III(c)	-136.3	3.2	-389.0	13.0	-20.3	0.9
	Set III(d)	-137.7	15.2	-392.1	13.6	-20.8	3.1
	Set IV	-142.3	6.8	-410.9	17.3	-19.8	0.5
<b>300mM</b>	Sequence	$\Delta H$	$\sigma$	$\Delta S$	$\sigma$	$\Delta G$	$\sigma$
	Set I	-163.2	8.4	-457.9	15.2	-26.7	1.2
	Set II	-152.7	2.8	-434.1	7.8	-23.2	1.4
	Set III(a)	-160.5	9.2	-456.3	6.0	-24.5	0.2
	Set III(b)	-161.5	5.3	-460.4	8.9	-24.2	0.2
	Set III(c)	-140.2	7.9	-396.4	10.0	-22.0	0.7
	Set III(d)	-143.1	13.0	-407.9	18.3	-21.5	1.2
	Set IV	-152.5	5.5	-440.0	20.1	-21.3	0.4
<b>1000mM</b>	Sequence	$\Delta H$	$\sigma$	$\Delta S$	$\sigma$	$\Delta G$	$\sigma$
	Set I	-169.6	3.9	-469.8	9.9	-29.5	0.4
	Set II	-156.8	4.4	-447.2	12.3	-23.5	0.4
	Set III(a)	-164.8	3.9	-468.5	10.1	-25.1	0.4
	Set III(b)	-165.4	1.2	-471.1	7.3	-24.9	1.2
	Set III(c)	-151.4	11.3	-430.2	13.9	-23.1	1.8
	Set III(d)	-154.5	7.2	-440.1	14.2	-23.3	2.9
	Set IV	-156.3	7.8	-444.0	17.0	-23.9	0.9

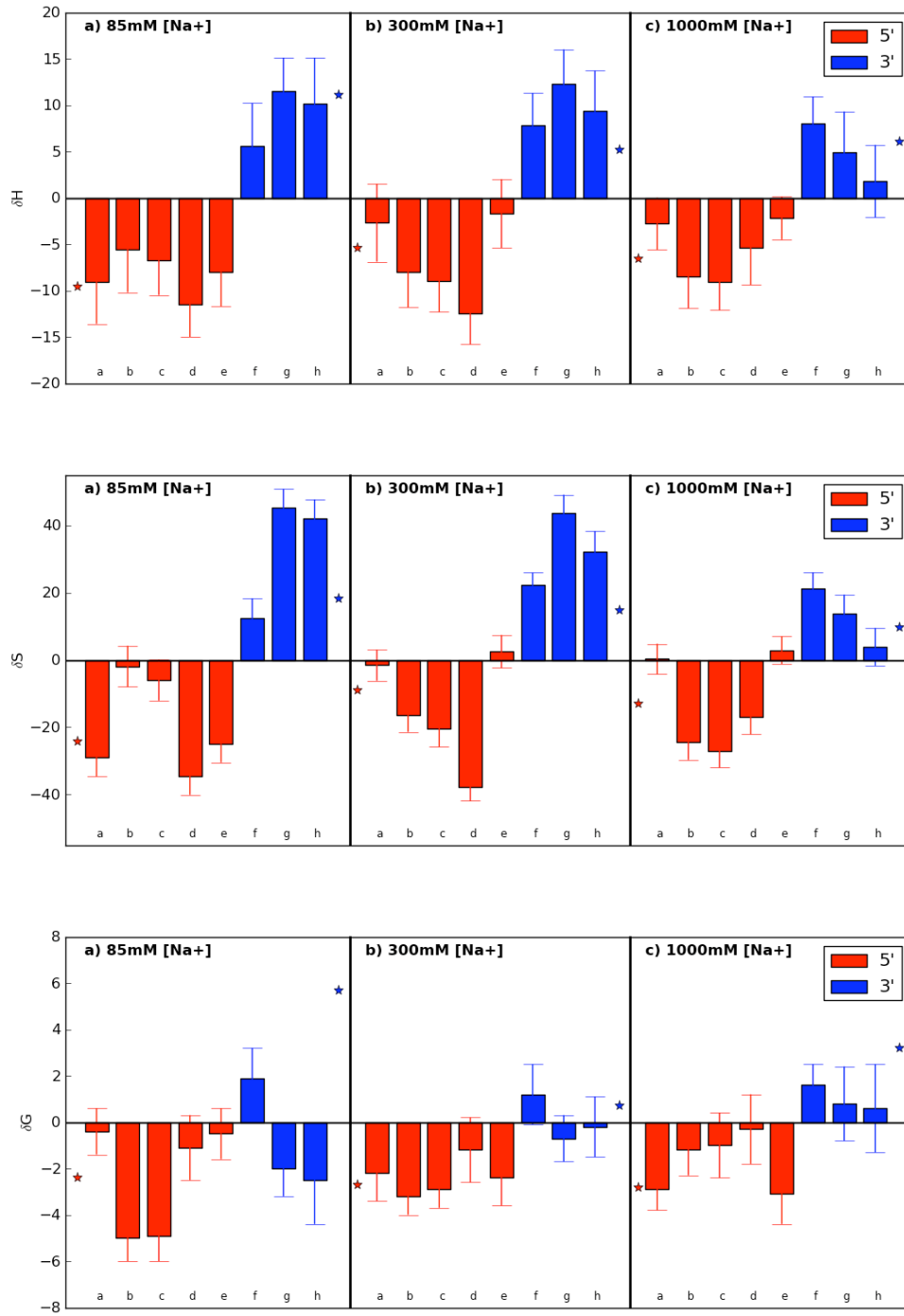


Figure 6: Histograms summarizing values of  $\delta X_{ss}^{5'}$  and  $\delta X_{ss}^{3'}$  determined from the thermodynamic data for the set III molecules and those relevant molecules from the sets I and II with  $n_D=21$  base pairs and  $n_L=4$ . Results from the different calculation schemas are shown. They are designated for  $\delta X_{ss}^{5'}$  (red) (a) I-IIIa, (b) IIIa-IV, (c) IIIb-IV, (d) II-IIIc, (e) I-IIIb, (\*) (I-IV)/2. For  $\delta X_{ss}^{3'}$  (blue) (f) II-IIIa, (g) IIIc-IV and (h) IIId-IV (\*). Stars show values determined from analysis of set I and II data, in Fig 5.

## Examination of Difference Equations

The thermodynamic parameters  $\delta X_{ss}^{3'}(n_L = 4)$  and  $\delta X_{ss}^{5'}(n_L = 4)$  depicted in Figure 6 were determined from consideration of different types of dangling ended duplex molecules. Due to the differing types of dangling ends for the molecules that were considered, i.e. two 5' dangling ends (set I), a 5' and 3' dangling end (set II), a single 5' or 3' dangling end (set III), or blunt ends (set VI), several methods were employed to calculate the values of  $\delta X_{ss}^{3'}(n_L = 4)$  and  $\delta X_{ss}^{5'}(n_L = 4)$ . Even so, and regardless of the particular calculation scheme, resultant values should be in reasonable agreement, provided assumptions invoked in their evaluation are valid. In all five basic schemes were employed as depicted in Figure 7.

Scheme 1 uses the difference between the set I molecules with two four base 5' dangling ends and the set III molecules (IIIa and IIIb), each having a single four base 5' dangling end, to produce  $\delta X_{ss}^{5'}(n_L = 4)$ . In scheme 2, values for  $\delta X_{ss}^{5'}(n_L = 4)$  and  $\delta X_{ss}^{3'}(n_L = 4)$  are derived from differences between set III molecules having a single 5' or 3' four base dangling end and the blunt-ended duplex from set IV. In scheme 3, the value for  $\delta X_{ss}^{3'}(n_L = 4)$  was determined from the difference between the set II molecule with a 5' and 3' dangling end and the set III molecule having one 5' dangling end (set III a). Likewise in scheme 3,  $\delta X_{ss}^{5'}(n_L = 4)$  was also evaluated from the difference between the set II molecule and the 3' dangling end molecule in set III (c). In scheme 4,  $\delta X_{ss}^{5'}(n_L)$  was determined from the difference between the

set I molecule (two 5' dangling ends) and the corresponding blunt-ended molecule from set IV, divided by two (as plotted in Figure 4). Finally, scheme 5 evaluated  $\delta X_{ss}^{3'}(n_L)$  from the difference between series II molecules (one 5' and one 3' dangling ends) and the blunt-ended molecule from set IV in addition to the value of  $\delta X_{ss}^{5'}(n_L)$  (as plotted in Figure 4). Schemes 4 and 5 were used for all calculations of  $\delta X_{ss}^{5'}(n_L)$  and  $\delta X_{ss}^{3'}(n_L)$  as a function of  $n_L$ , while schemes 1, 2 and 3 were used only for the series III molecules to validate the thermodynamic parameters derived from schemes 4 and 5 where  $n_L=4$ .

Examination of the difference plots in Figure 6 reveals trends reminiscent of those seen in Figure 4. To a first approximation, averages of the results from the different calculation schemes provide values in agreement for  $\delta X_{ss}^{3'}(n_L)$  and  $\delta X_{ss}^{5'}(n_L)$ . For all plots the 5' dangling ends appear to be stabilizing where the 3' dangling ends appear to be near zero or destabilizing. Entropy and enthalpy values display variation, but also good agreement within each calculation scheme, and compensate such that the free-energy values are in consensus. Exceptions are values for  $\Delta G_{ss}^{3'}$  at 85mM Na<sup>+</sup> and 300mM Na<sup>+</sup>, where the schemes produce seemingly contradictory results. Further, the plot for  $\Delta G_{ss}^{5'}$  in 85mM Na<sup>+</sup> also shows significant variability. Apparently in some cases the values obtained depend on the particular calculation scheme employed, suggesting there may be additional factors associated with dangling ends, particularly at the lowest [Na<sup>+</sup>], that are not adequately considered in

our evaluation procedure.



Scheme 1	$\text{Set I} = \frac{n_D}{n_L} \left  \begin{array}{c} 21 \\ 4 \end{array} \right $ $\delta X_{ss}^{5'} = \text{Set III a,b}$
Scheme 2	$\text{Set IV} = \frac{n_D}{n_L} \left  \begin{array}{c} 21 \\ 4 \end{array} \right $ $\delta X_{ss}^{5'} = \text{Set III a,b}$ $\delta X_{ss}^{3'} = \text{Set III c,d}$
Scheme 3	$\text{Set II} = \frac{n_D}{n_L} \left  \begin{array}{c} 21 \\ 4 \end{array} \right $ $\delta X_{ss}^{5'} = \text{Set III c}$ $\delta X_{ss}^{3'} = \text{Set III a}$
Scheme 4	$\delta X_{ss}^{5'} = \text{all Set I and IV}$
Scheme 5	$\text{Set VI} = \frac{n_D}{n_L} \left  \begin{array}{cccccc} 24 & 23 & 22 & 21 & 19 & 17 \\ 1 & 2 & 3 & 4 & 6 & 8 \end{array} \right $ $\delta X_{ss}^{3'} = \text{all Set II}$

Figure 7: Diagram of schemes used to calculate  $\delta X_{ss}^{5'}$  and  $\delta X_{ss}^{3'}$  values as seen in Figure 6 and Figure 4.

### 3.3 Examination of Duplex Properties

#### Counterion Binding Theory

In solution, binding of  $\text{Na}^+$  to both duplex and single strand DNA is driven primarily by the charge densities of the respective duplex and single strand species<sup>32,33,7</sup>. DNA duplexes have a relatively (significantly) higher charge density compared to single strands due to the higher negative charges of the double stranded phosphate backbone, mobile counterions, and water that surround the double helix<sup>34</sup>. Consequently, due to the reduced charge density of the melted single strands compared to the duplex, when a duplex denatures there is a net release of  $\text{Na}^+$ <sup>35,36</sup>.

The melting data and corresponding thermodynamic parameters that were

evaluated as a function of  $[Na^+]$  provide a means of quantitatively estimating the net  $Na^+$  release upon melting of short duplex DNAs as function of duplex and dangling end length. The release of  $Na^+$  upon melting,  $\Delta n$ , can be estimated assuming simple binding equilibrium and evaluated according the following<sup>34</sup>,

$$\alpha \Delta n = -\Delta H^\circ \left[ dT_m^{-1} / R \left( d \ln [Na^+] \right) \right] \quad (12)$$

Where  $\Delta H^\circ$  is the standard state enthalpy of dissociation of the duplex,  $R$  is the ideal gas constant and  $\alpha$  is a correction term for the sodium ion activity coefficient. A standard value of was assumed throughout. From slopes of plots of  $dT_m^{-1}(K)$  versus  $d \ln [Na^+]$  for the set I, set II and set IV duplexes estimates of  $\Delta n$  (the release of  $Na^+$ ) for each type of duplex were made as a function of duplex length. Generally, from linear fits of  $dT_m^{-1}(K)$  versus  $d \ln [Na^+]$  plots, the required slopes in Eqn (12),  $\left[ dT_m^{-1} / R \left( d \ln [Na^+] \right) \right]$  and  $\Delta n$  were evaluated. For comparison  $\Delta \Psi$ , the counterion release per phosphate,

$$\Delta \Psi = \Delta n / N \quad (13)$$

was plotted versus duplex length for the sets I, II and IV duplexes as shown in Figure 8. For these calculations,  $N$  is the number of phosphates in the duplex state including the single strand ends where present (i.e. in sets I and II).

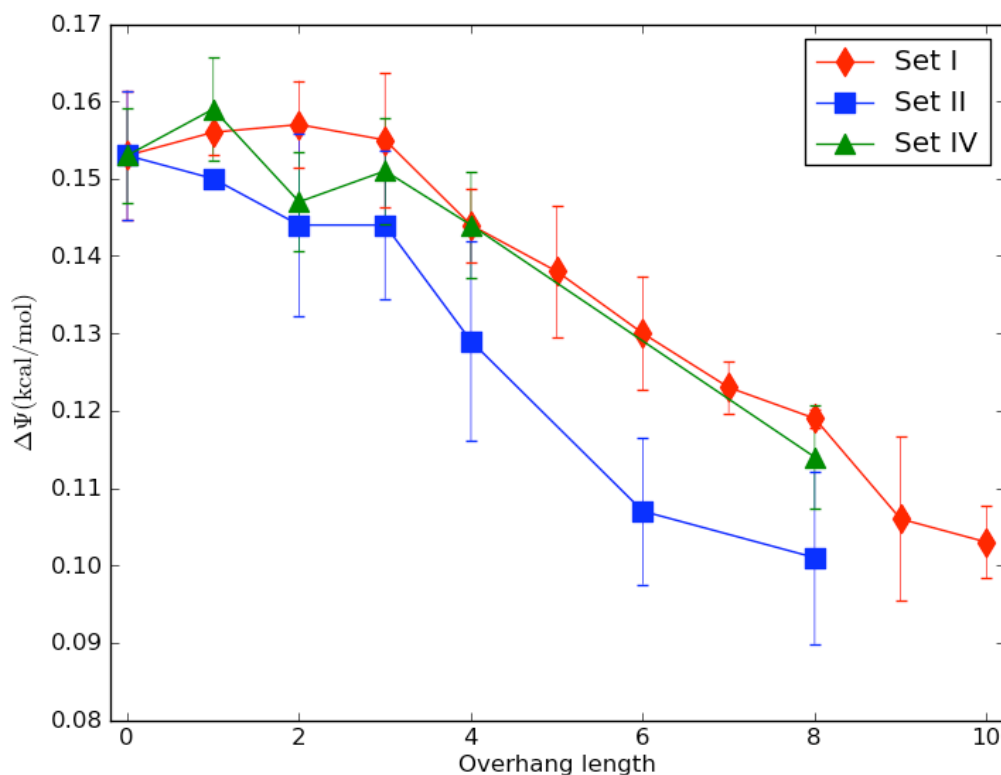


Figure 8:  $\Delta\Psi$  versus  $n_L$  for the three main molecule sets.

### Differential Counterion Release Upon Melting

Examination of the plots in Figure 8 reveals interesting behavior for  $\Delta\Psi$  as a function of decreasing duplex, and increasing dangling-end length, for the set I, set II and set IV duplexes. For the set I and set IV molecules,  $\Delta\Psi$  decreases with increasing duplex length and is essentially identical for both types of duplexes. For the set II molecules,  $\Delta\Psi$  decreases in the same trend, but is approximately 15% smaller at every point on the plot. These data indicate on average a net lower  $\text{Na}^+$  release per phosphate in melting for the set II molecules as compared to the set I

and set IV molecules, which are essentially indistinguishable on plots of  $\Delta\Psi$  vs.  $n_d$ .

Since relatively higher charge density of the duplex compared to the single strands is the underlying origin of counterion release upon duplex melting, the fact that the set I and set IV molecules display the same counterion release suggests the duplexes have the same charge density. More importantly, the ends must behave the same (in a counterion binding sense) and have the same charge density and counterion binding properties whether at the end of a duplex or in a melted single strand. In contrast, the observed behavior for the set II molecules suggests several possible scenarios. Either, the 3' dangling single strand ends bind less  $\text{Na}^+$  when on the end of a duplex compared to their single strand state, while the duplex binds counterions to the same extent; or the duplex adjoined by a 3' end is perturbed to an extent that decreases the local duplex charge density, which results in overall less counterion binding to the duplex state and consequently leads to a net lower counterion release upon melting. These scenarios will be considered again later.

### **Heat Capacities of the Melting Transitions**

In DSC experiments,  $\Delta H_{cal}$  is evaluated from the integrated area under the DSC melting curve,  $\Delta C_p(T)$  versus Temperature. As a result, the thermodynamic parameters of the melting transition,  $\Delta H_{cal}$  and  $\Delta S_{cal}$ , are only strictly accurate in the transition region. As a matter of practice, values of  $\Delta H_{cal}$  and  $\Delta S_{cal}$  are routinely used to predict thermodynamic stabilities of duplexes at temperatures (37°C) far

below the actual transition region (55-75°C). The accuracy of such predictions rely on the validity of the assumption that the evaluated parameters are temperature independent and the overall change in excess heat capacity,  $\Delta C_p$ , in the transition from duplex to single strands, is zero. If  $\Delta C_p \neq 0$ , then the thermodynamic parameters evaluated from analysis of the melting transition region may not be accurate for predictions at lower temperatures, where there is more biological relevance and where practical applications occur. The experimental transition enthalpy is given by,

$$\Delta H_{cal}(T) = \Delta H^\circ + \delta H + \Delta C_p (T - T^\circ) \quad (14)$$

where  $T^\circ$  is an arbitrary reference temperature and  $\delta H$  is a correction term for small variations in salt environments. Generally, for short duplex DNA, the two-state model is employed and it is assumed that the difference between  $T^\circ$  and  $T$  is small in the transition region where the thermodynamic parameters are evaluated, thus  $\Delta C_p = 0$ <sup>6</sup>. Although this assumption greatly simplifies parameter evaluations and calculations performed with them, over the past 15 years studies have reported the existence of a relative standard transition heat capacity for all duplexes DNA sequences. Estimates for the value are as high as 100 cal·deg<sup>-1</sup>·mol of base pair<sup>-1</sup> and vary slightly with sequence, salt<sup>30,19, 29</sup>. The average value of  $\Delta C_p = 64.6 \pm 21.4$  cal·deg<sup>-1</sup>·mol of base pair<sup>-1</sup> was reported to be a good approximation<sup>37</sup>.

Inclusion of the  $\Delta C_p$  parameter allows for more accurate predictions of the transition enthalpy and entropy at temperatures below the transition region<sup>19</sup>. For

dangling ended molecules, there is a question regarding their effect on, and contributions to  $\Delta C_p$ . To address this question, plots of the  $\Delta H_{cal}$  values versus  $T_m$  in all Na<sup>+</sup> environments for the molecules of sets I, II and IV were constructed as shown in Figure 9. Each set of molecules at all [Na<sup>+</sup>] display different slopes. Values of  $\Delta C_p$  determined from these slopes are 25.1 cal·deg<sup>-1</sup>·base pair<sup>-1</sup> for set I, 68.9 cal·deg<sup>-1</sup>·base pair<sup>-1</sup> for set II, and 31.2 cal·deg<sup>-1</sup>·base pair<sup>-1</sup> for set IV. These values differ somewhat from the average reported  $\Delta C_p$  value of  $64.6 \pm 21.4$  cal·deg<sup>-1</sup>·mol of base pair<sup>-1</sup>. However, if the values from the lowest [Na<sup>+</sup>] are not considered,  $\Delta C_p$  values of 54.6, 85.2 and 54.1 cal·deg<sup>-1</sup>·base pair<sup>-1</sup> are found for sets I, II and IV respectively, in reasonable agreement with the reported value.

To compare variations of the  $\Delta C_p$  values between the dangling ended molecules of sets I and II, and their blunt ended counterparts, set IV, the differences between the enthalpic parameters at the transition temperature of the dangling ended molecule,  $T = T_m^{dangling}$ , were found. For set I and set IV,

$$\Delta\Delta H(T = T_m^I) = \Delta\Delta H^\circ + \delta\delta H - \Delta C_p^I \times T_m^I + \Delta C_p^{IV} \times T_m^{IV} + \Delta\Delta C_p \times T_m^I \quad (15)$$

For the calculation, difference values are found from  $\Delta H^{IV}(n_d)$ , the enthalpy of the blunt molecules and  $\Delta H^I(n_d)$ , those of the dangling end molecules of the same duplex length.

In this expression there is a term independent of  $\Delta C_p$ , i.e.,

$$\Delta\Delta H^\circ + \delta\delta H - \Delta C_p^I \times T_m^I + \Delta C_p^{IV} \times T_m^{IV} \quad (16)$$

and a dependent term,

$$\Delta\Delta C_p \times T_m^I \quad (17)$$

This allows for evaluation of the differences in  $\Delta C_p$  for the dangling ended and blunt ended molecules.

$$\frac{d\Delta\Delta H(T = T_m^I)}{dT_m^I} = \Delta\Delta C_p \quad (18)$$

If there is no difference in the  $\Delta C_p$  between the sets, plots of  $\Delta\Delta C_p$  versus  $T_m^I$  would be expected to have zero slope, as observed in Figure 9 for the dangling ended molecules of set I. In contrast, if there is an observed slope of the plot of  $\Delta\Delta C_p$  versus  $T_m^I$  as observed in Figure 10 for the set II dangling ended molecules, there are additional factors from the dangling ends that contribute to  $\Delta C_p$  for these molecules. The slope of the plot in Figure 10 for the set II molecules provides an estimated value of  $\Delta\Delta C_p = 52.5$  cal·deg<sup>-1</sup>·base pair<sup>-1</sup>. This in excess to the  $\Delta C_p$  of the blunt ended molecules. If the average value of  $\Delta C_p$  is taken to be 64.6 cal·deg<sup>-1</sup>·base pair<sup>-1</sup> as reported, the total  $\Delta C_p \cong 117$  cal·deg<sup>-1</sup>·base pair<sup>-1</sup> for the set II molecules.

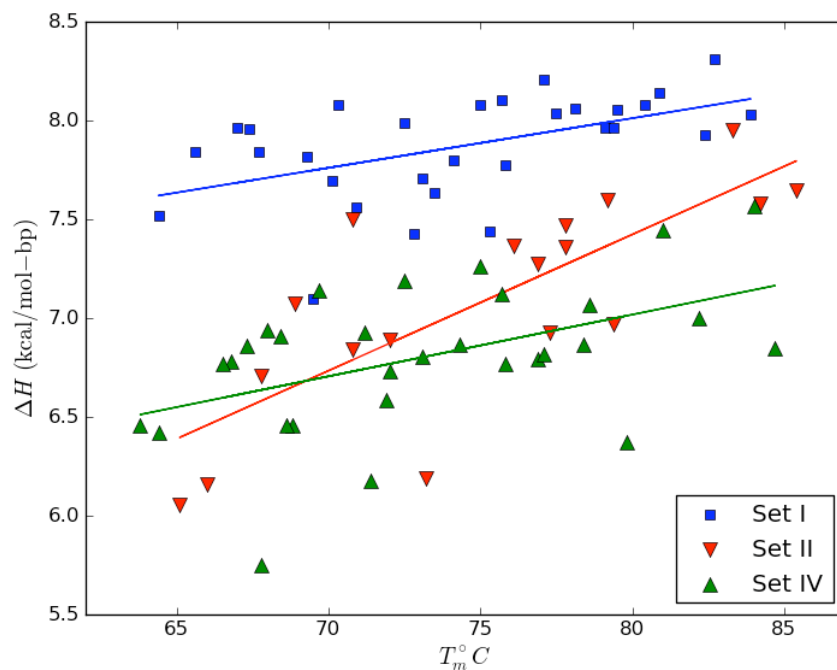


Figure 9: Plot of  $\Delta H(T_m)$  vs.  $T_m$  for each type of molecule in all salts. Set I molecules display a  $\Delta C_p$  value of  $25.1 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base pair}^{-1}$ , set II molecules  $68.9 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base pair}^{-1}$  and set IV molecules  $31.2 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base pair}^{-1}$

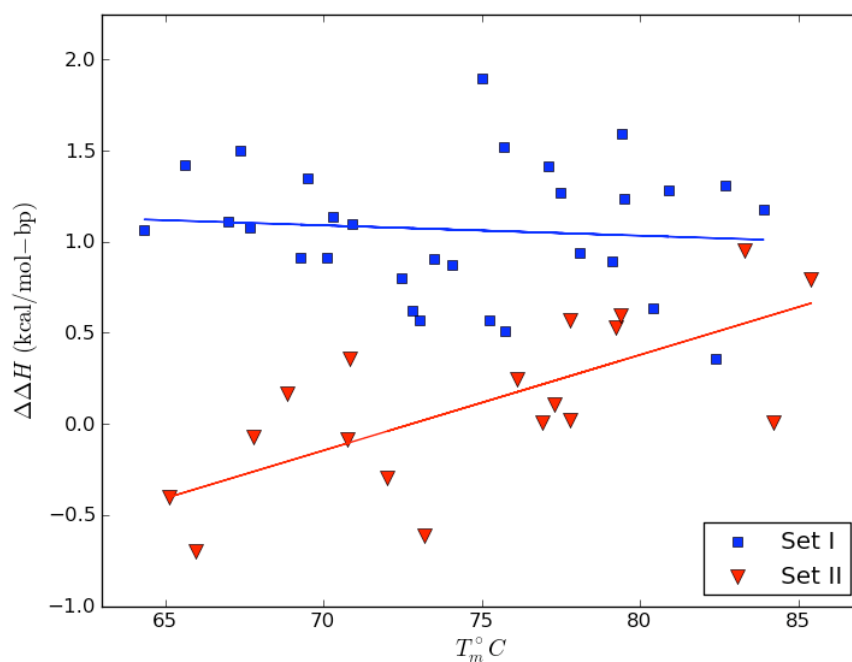


Figure 10: Plot of  $\Delta\Delta H(T_m)$  vs.  $T_m$  for each type of molecule in all salts. Set I molecules display a  $\Delta\Delta C_p = -5.8 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base pair}^{-1}$ , set II molecules  $52.5 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base pair}^{-1}$  as compared to the set IV molecules.



## **Chapter 4 – Discussion and Interpretations**

### **4.1 Literature Comparisons**

#### **Comparisons for 5' Dangling Ends**

Factors involved in determining the relative influence a dangling end will have on a duplex include, identity of the molecule (DNA or RNA), terminal base pair identity, specific dangling residue, and location of residue in relation to the terminal base pair. In this study, molecule identity (DNA) and terminal stack identities were maintained to focus on effects related to dangling residue placement (5' or 3') as a function of length. Several previous studies on thermodynamic contributions of DNA dangling ends to duplex stability have been performed<sup>10,11,1,20,38,39</sup>. These examined different permutations of the combination of end length, end sequence and terminating duplex sequence. None have examined effects of end length stepwise out to a ten base dangling end ( $n_L = 10$ ), nor have they studied effects of length and type (5' or 3') in different salt environments. Although published studies of dangling ends have not addressed these specific factors, there are several places in which our results can be compared with those previously published.

The thermodynamic effects of all single base dangling ends were studied by Santa Lucia and coworkers through UV melting analysis<sup>1</sup>. In their systematic study, a duplex was fixed at eight base pairs and single dangling bases were attached to the 5' or 3' ends. All possible permutations of the terminating stack were examined and

sequence dependent thermodynamic parameters were evaluated. Relevant sequence dependent interactions determined in their study that can be compared with our results in a 1M Na<sup>+</sup> environment for the 5' ends, are those of the terminal stack 5'AC/3'G. Here a terminating CG base pair is adjoined on the 5' end by a single A base (Tables 1-3). After conversion to 25 °C, they report an observed stabilization of  $\Delta G_{25} = -1.21$  kcal/mol.

Our value of  $\delta G_{25}(n_L = 1)$  at 25°C for the 5'AC/3'G stack is only -0.3 kcal/mol, not nearly as stabilizing as reported and apparently an inconsistent result. However, for longer ends,  $n > 1$ , we found an average  $\delta G_{25}$  value of -2.6 kcal/mol, in better agreement with published results.

Nearly 20 years ago Doktycz et. al. published results of melting studies using DNA hairpin molecules<sup>20</sup>. The general sequence of, 5'(XY)<sub>2</sub>(GGATAC)<sub>2</sub>(T)<sub>4</sub>, naturally folds to form a duplex of six bases, with a T<sub>4</sub> single strand loop connecting one end of the duplex and a four base 5' dangling end on the other. Considering only the molecules with the same specific end sequences, i.e. those with a terminal stack of 5'AC/3'G, the reported  $\delta G_{25}$  gave an average value of -1.17 kcal/mol in 115 mM Na<sup>+</sup>. At  $n_L = 4$  in 100 mM Na<sup>+</sup> our data has coalesced to an average value of -1.5 kcal/mol which again is in reasonable agreement with the published result.

Ohmichi published results of melting studies of an eight base pair duplex molecule with 5' dangling ends varying from one to four bases<sup>10</sup>. The terminating sequence in their molecules was slightly different i.e. 5'AG/3'C instead of 5'AC/3'G,

disallowing an exact comparison. Nonetheless, the reported trends as a function of increasing length can be considered. For a single base 5' dangling end in 1.0 M Na<sup>+</sup>, a stabilization of  $\delta G_{25} = -0.3$  kcal/mol was reported. Addition of a second dangling base increased stabilization to -0.4 kcal/mol. This trend seemed to reach a consistent value around three bases with a value of -0.6 kcal/mol. This increase in stability coupled to the lengthening of the single strand dangling end is consistent with our observations.

For our data, in the high salt (1.0 M Na<sup>+</sup>), the first base was found to be stabilizing by only -0.3 kcal/mol while the addition of a second base contributes about -2 kcal/mol of stability. This effect is seen to a lesser extent in 300 mMol Na<sup>+</sup> where the first base is stabilized by -1.7 kcal/mol and the second about -2.6 kcal/mol. At 85mMol Na<sup>+</sup> the trend is reversed with the first base adding a stabilizing -3 kcal/mol and the second only -2.6 kcal/mol, although it should be noted in 85 mM Na<sup>+</sup>, the standard error is large enough that both values are essentially in agreement.

### **Comparison of 3' Dangling Ends**

Compared to 5' dangling ends, fewer studies of 3' dangling ends on the thermodynamic stability of duplex DNA have been performed. Overall, reports have found that 3' dangling ends make a favorable contribution to duplex stability, thus are stabilizing, but less so than their 5' counterparts. From the systematic study of

all dangling ends the temperature corrected value for the  $5'C/3'AG$  stack is  $\delta G_{25} = -1$  kcal/mol<sup>1</sup>. Comparing that to our adjusted result in 1.0 M Na<sup>+</sup> of  $\delta G_{25} = -0.8$  kcal/mol, in good agreement.

Ohmichi et. al. also evaluated effects of 3' dangling ends as a function of increasing length<sup>10</sup>. Again their system had a slightly different sequence  $5'G/3'CA$ , where the dangling A stacked on the C of the GC base pair instead of over the G, when the length increases from one to four bases in 1.0 M [Na<sup>+</sup>]. The reported  $\delta G_{25}$  values, are -0.5, -0.5, -0.8 to -0.8 kcal/mol, respectively. The values followed the same stabilizing trend as seen for increasing of length of the 5' dangling end. In 1.0 M [Na<sup>+</sup>], our observations follow the same trend as published findings. For  $n_L = 1, 2$  and 3 the ends are stabilizing with a slight trend toward increased stabilization. However, in our system at  $n_L = 4$  the trend reverses and becomes destabilizing. In particular, we observed that a four-base 3' dangling end is destabilizing by  $\delta G_{25} = +4.5$  kcal/mol in 1 M Na<sup>+</sup>. The lack of stabilization at this specific length may be due to the identity of the terminal stack that was used and might explain why Ohmichi et. al. did not observe such an event.

The dependence on [Na<sup>+</sup>] is also greater for the 3' versus the 5' end. At low salt concentrations the 3' dangling end was found to be extremely destabilizing. As the [Na<sup>+</sup>] increased, the effect was decreased especially for the longer ends. At  $n_L > 4$  there is little difference between the values in 300 mM and 1.0 mM [Na<sup>+</sup>]. Although unexpected, observations of a destabilizing dangling end are not

completely unpublished. The terminal stacks, 5'GT/3'A, 5'T/3'AC and 5'T/3'TA were found by SantaLucia to be destabilizing to a small degree, which suggests that our observed reverse in stability with longer dangling ends is not unfounded<sup>1</sup>.

## **4.2 Factors Involved in Dangling End Behavior**

### **Stabilization Models**

Our results indicate that 5' dangling ends are equally or more stabilizing than their 3' counterparts. This behavior has been previously documented and can possibly be explained by examining DNA single strand structure. Predictive algorithms treat single strand RNA and DNA as a random coil although it has been shown single strands possess some degree of order<sup>40,19</sup>. In its duplex state, RNA preferably adopts an A-form conformation while DNA preferably adopts the B-form. These orientations are conserved when in the single strand state. NMR studies of ssRNA and ssDNA hexamers with multiple adenine-adenine stacks, showed that in ssDNA, the imidazole stacks above the pyrimidine in the 5' to 3' direction, while in ssRNA the pyrimidine stacks above the imidazole in the 5' to 3' direction<sup>41</sup>. The 5' and 3' geometries are shown in Figure 11, for ssDNA (A) and ssRNA (B). When a non-paired base is added to the end of a helical duplex as a dangling end, these stacking preferences translate into bonding ability (C-F). They also provide a plausible explanation for differences in stability dependent on the location (3' or 5') of the dangling end.

A systematic review of crystal structures from the protein database (PDB) demonstrates for RNA, addition of a single strand base on the 3' end is positioned in such a way that it can freely interact with the hydrogen bonds of the terminal base pair. In contrast, a 5' base end is positioned away from the same hydrogen bonds, and therefore less likely to experience such stabilization. For DNA the situation is reversed. Adjoining the terminal base pair, a 5' end base interacts with the hydrogen bonds of the terminal base pair, but the 3' end base does not participate in such interactions with the 3' terminal base pair<sup>22</sup>. This is shown in Figure 11. Placement of the dangling base is optimal for terminal base pair interactions for DNA on the 5' end (C) and RNA on the 3' end (F). Conversely, when an additional base is added to the 3' end of DNA (E) or the 5' end of RNA (D) minimal overlap is seen, which must translate to fewer stabilizing interactions.

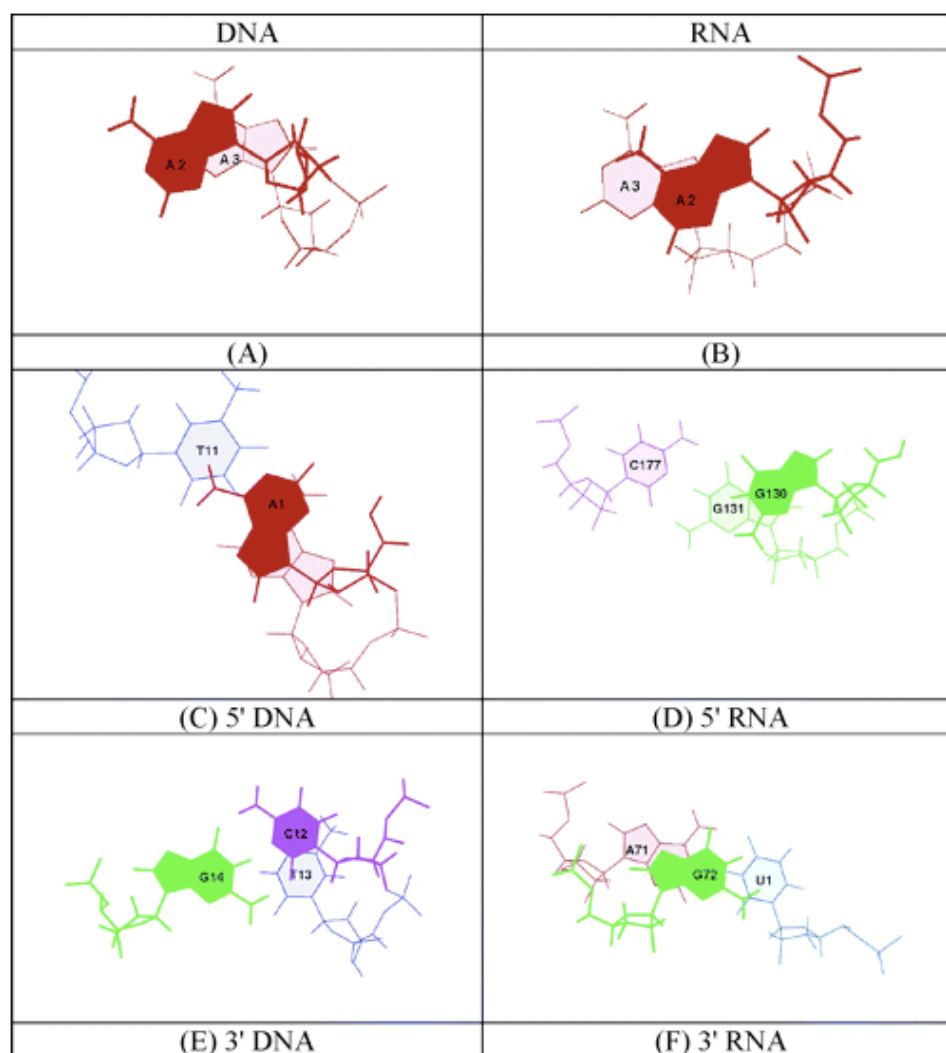


Figure 11: Directional stacking difference of ssDNA (A) and ssRNA (B) as viewed from the 5' end. Panels C (DNA) and D (RNA) show stacking geometries for a 5' dangling base as seen in the 5' direction. Panels E (DNA) and F (RNA) show stacking geometries for a 3' dangling base as seen from the 5' direction for DNA and 3' direction for RNA. (Isaksson, 2002. Used with permission)

## The Mystery of 3' Dangling Ends and the Set II Molecules

Our results indicate a 3' dangling end is generally destabilizing to a DNA duplex. Molecules used in this study were designed to mimic situations thought to occur commonly in multiplex hybridizations reactions on DNA microarrays, where it is not uncommon for a target and probe molecule to anneal so as to have additional flanking single strand regions. Such annealing can result in a duplex with both 5' and 3' dangling ends on the same strand. This is the design used in the set II molecules (Table 2). Subtle variations in the calculated thermodynamic results for the 3' dangling ends that appear to be method dependent (Figure 6), suggesting that additional factors may need to be considered in the set II molecules.

Values of the thermodynamic parameters  $\delta X_{ss}^{5'}(n_L)$  and  $\delta X_{ss}^{3'}(n_L)$  as a function of dangling end length,  $n_L$ , plotted in Figure 4, as compared with further analysis of  $\delta X_{ss}^{3'}(n_L = 4)$  in Figure 6, conclude that 3' dangling ends are more destabilizing in free energy at low  $[Na^+]$  when evaluated from results of the set IV and set II (double ended molecules) than they are from the set IV and set III (half molecules). The subtle and apparently anomalous difference in evaluated stability contributions of the 3' end obtained in the different molecular environments suggests our parsing assumption of the thermodynamic contributions of the dangling end and duplex regions may be compromised. If interactions between the dangling end and the adjoining duplex are not localized, as assumed in the n-n approach, and the entire duplex region is actually perturbed by the combination of dangling ends in the set II



molecules, that is, the 3' dangling end must affect stability of the entire duplex, then the assumed parsing of thermodynamic contributions between the duplex and dangling ends is no longer strictly valid. This notion of “global destabilization” is further supported by the plots of counterion release ( $\Delta\Psi$ ) as a function of dangling end length ( $n_L$ ) (Figure 8) and the heat capacity ( $\Delta C_p$ ) analysis (Figures 9 and 10) for the set II molecules.

Previous studies on counterion binding to duplex DNA suggest fewer  $\text{Na}^+$  ions bind near the ends of a molecule compared to in the middle<sup>34,42</sup>. Thus for the ends, differences in the dangling end versus blunt molecules should result in negligible changes in counterion binding, provided the duplex region is not additionally affected by the ends. This is supported by comparison of the blunt ended duplexes and set I molecules with 5' dangling ends. The net counterion release per phosphate is the same for the two sets (I and IV) of molecules and the plot of  $\Delta\Psi$  vs.  $n_L$  in Figure 8 are the same. In contrast for the set II molecules with one 3' and one 5' dangling end, the plot of  $\Delta\Psi$  vs.  $n_L$  is approximately 15% lower than for the set I and set IV molecules, i.e. there is a net lower  $\text{Na}^+$  release in melting the set II molecules compared to the set I and set IV duplexes at all values of  $n_L$ . This suggests that the duplex region for set II is altered in some way, creating differences in the associated counterion binding properties significant compared to the duplex regions of the set I and set IV molecules.

When DNA molecules anneal from their single state to their duplex state, there is a net change in solvent exposed surfaces. This change which is accompanied by the burying of hydrophobic residues, contributes to the  $\Delta C_p$ . Differences in the calculated  $\Delta\Delta C_p$  values for set I as compared to set II (Figure 10), further support the idea of subtle differences in the duplex regions of the molecules. When  $\Delta C_p$  values for the blunt molecules are compared to  $\Delta C_p$  values for the set I molecules, a constant difference of about 1 cal·deg<sup>-1</sup>·base pair<sup>-1</sup> is observed. There is also no marked length dependence, supporting the idea that most of the stability comes from interactions of the first base with the terminal base pair. In molecular models of dangling end-stacking geometries, molecules with the 5' dangling end have favorable stacking interactions, which translates to a slightly higher buried surface area.

Conversely when  $\Delta C_p$  values for the set II molecules are compared to the  $\Delta C_p$  of the blunt molecules, a non-constant difference is found. The slope of the difference line ( $\Delta\Delta C_p$ ) is 52.5 cal·deg<sup>-1</sup>·base pair<sup>-1</sup>, suggests as the length of the dangling ends is increased in the set II molecules, the buried surface area of the duplex region is decreased, making it less ordered. Together these mysterious results for the set II molecules suggest the duplex region is perturbed in this particular molecular environment.

### 4.3 Practical Applications for Probe Design

Multiplex hybridization reactions are designed to occur on microarrays in which single strands of DNA (probes) are affixed to a solid surface and then washed with sample DNA (targets). Probes are designed to specifically anneal with targets having complementary sequences. Binding can be detected using fluorescence or other signal generation systems. Microarray based assays can be used to probe the genome for specific sequences such as mutations that carry increased chances of developing a specific disease or disorder. The ability to design probes with extreme accuracy is imperative to successfully locate target sequences that can differ by as little as a single nucleotide. Thus, ascertaining the specific thermodynamics involved in probe/target alignment and being able to predict the energies of all possible alignments, such as small duplex binding regions flanked by non-bound (dangling ends) is key to optimal probe design. To wit, the more specific the predictive ability of thermodynamic binding properties, the more discrete and effective the probe design.

#### Predictive Ability

In his unified theory of DNA, Santa Lucia outlines the basic method for calculating the free energy of melting a duplex molecule using the n-n model as,

$$\Delta G^\circ = \Delta G_{initiation}^\circ + \sum \Delta G_{stack}^\circ + \Delta G_{additional}^\circ \quad (19)$$

For the duplex region this model uses combinations of the ten possible n-n values (

$\Delta G_{stack}^\circ$ ) that have been experimentally determined by a number of independent labs and are generally in good agreement with each other<sup>12</sup>.  $\Delta G_{initiation}^\circ$  is the cost required to begin the annealing process and has a positive free energy contribution. This value has recently been determined by our group, as well as other investigators<sup>43,44</sup>. The last term  $\Delta G_{additional}^\circ$ , encompasses any extra terms such as those that arise from symmetry considerations, a terminating AT base pair or single strand dangling ends. Once collected, all terms are summed to estimate the total calculated free energy,  $\Delta G^\circ$ .

An example calculation using the n-n model is shown for the DNA sequence 5'CATGC/3'GTACG. When broken down into n-n doublets, four are present; CA/GT, AT/AT, TC/AC and GC/CG. Summation of the n-n doublet energies along with the initiation free energy, results in the n-n calculated stability parameters summarized in Table 10.

Table 10: Contributions and summation of the n-n sequence dependent parameters for the sequence 5'CATGC/3'GTACG.

Stack	$\Delta H$	$\Delta S$	$\Delta G_{25}$
CA/GT	-8.4	-22.4	-1.7
AT/TA	-7.2	-20.4	-1.1
TG/AC	-8.5	-22.7	-1.7
GC/CG	-9.8	-24.4	-2.5
$\Delta G_{initiation}$	0.2	-5.7	1.9
Sum	-33.7	-95.6	-5.2

Thus, this sequence has a predicted  $\Delta G_{25} = -5.2$  kcal/mol

Current prediction programs using the n-n model for calculating thermodynamic properties are limited by the quality of the parameters they are based on. Mfold is one such program readily accessible via the Internet<sup>45</sup>. Their two-state model allows the sequence of two strands of DNA to be input, then the computation algorithm tests for the most stable structure formed and calculates the thermodynamic stabilities using tabulated n-n parameter values. To include effects of dangling ends, the specific value of -1.18kcal/mol for the 5' dangling end (5'AC/3'G) and -1.05kcal/mol for the 3' dangling end (5'C/3'AG) are added to the n-n calculations<sup>1</sup>. Potential differences in the thermodynamic effects of dangling ends longer than one base are not explicitly considered.

To test the applicability and utility of our evaluated dangling end parameters, five sequences with a duplex region of 21 base pairs were designed. Calculated thermodynamic transition parameters were generated using two methods, i.e. Mfold and the n-n model as discussed above (Table 10). Sequences differed in their design by placement and length of the dangling end. The sequences were: (1) the 21 base blunt ended control, (2) a 21 base duplex with two four base 5' dangling ends, (3) a 21 base duplex with two 5' dangling ends with five and eight bases respectively, (4) a 21 base duplex with one four base 3' and one four base 5' dangling end, and (5) a 21 base duplex with one five base 3' dangling end and one seven base 5' dangling end. Results shown in Table 11.

Examination of the results in Table 11, for 5' ends the n-n model predictions

and Mfold predictions are quite comparable. In contrast, for the 3' ended molecules the Mfold predictions are greater (more stabilizing) than predicted using our 3' end parameters. Since there have been no reported measurement of these specific 3' ends destabilizing the duplex, it is not surprising the standard program (Mfold) overestimates the stability for those molecules.

Sequence Identity	85 mMol [Na+]			300 mMol [Na+]			1000 mMol [Na+]		
	Exp	Mfold	NN	Exp	Mfold	NN	Exp	Mfold	NN
5'- C(X) <sub>19</sub> G G(X) <sub>19</sub> C	-19.8	-18.6	-19.2	-21.3	-20.9	-21.5	-23.9	-22.4	-23
5'- TAGAC(X) <sub>19</sub> G G(X) <sub>19</sub> C AGAT	-24.6	-20.7	-24.4	-26.7	-23.1	-26.7	-29.5	-25.8	-28.2
5'- TAGGAC(X) <sub>19</sub> G G(X) <sub>19</sub> CAAGTCGAT	*	-20.7	-23.5	*	-23.1	-24.9	*	-25.8	-26.4
5'- TAGAC(X) <sub>19</sub> GAGAT G(X) <sub>19</sub> C	-16.5	-20.5	-18.1	-23.2	-22.9	-21.4	-23.5	-24.6	-23.3
5'- ATCGAC(X) <sub>19</sub> GAGAATC G(X) <sub>19</sub> C	*	-20.5	-18.0	*	-22.9	-20.5	*	-24.6	-22.1

Table 11: Sequences used to predict  $\Delta G_{25}$  kcal/mol. In all sequences the duplex region consists of <sup>5</sup>CATGTGATCGCTAGTGAGATG<sup>3</sup> and its perfect-match complement. SantaLucia n-n values used<sup>1</sup>. All predicted values were corrected with an averaged difference for set IV, ( $\Delta G_{exp} - \Delta G_{mfold}$ ), (85mM=+6.2), 300mM=+6.7), and(1000mM=+7.8)

## Chapter 5 – Summary and Conclusion

Using highly sensitive Differential Scanning Calorimetry, we have explored the thermodynamic properties of dangling ends for small DNA duplexes. Application of the dangling end parameters to the nearest-neighbor model can help in determining more accurate and reliable thermodynamic parameters for systems in which DNA of non-uniform lengths anneal to each other. Increasing the precision of anticipated reactions allows for better resolution in probe and target design. Ultimately these results can be used to create higher resolution predictions for multiplex DNA hybridization reactions.

Duplexes were designed to evaluate the subtle consequences of a sliding probe/target alignment scheme where different sequence lengths can cause 5'/5' or 5'/3' unpaired bases to appear on the end of the DNA duplex. It was demonstrated that 5' dangling ends stabilize the DNA duplex whereas 3' dangling ends were destabilizing for low salts and end lengths greater than four bases. A slight increase in 5' stability was noted for dangling ends greater than one base, after which it was constant. In 1 M [Na<sup>+</sup>], increasing the length of the 3' dangling end had a relatively constant stabilizing effect, but became destabilizing around four bases. The lower Na<sup>+</sup> conditions, 85 mM and 300 mM, showed a general destabilization effect, which also leveled out after four bases. The discovery of destabilizing 3' ends are contrary to previously published values and are suggested to occur through perturbations of the duplex region in molecules with a 5' and 3' end on the same strand.



Apparently thermodynamic stability of dangling ends depends on the interplay between 5' and 3' placement versus molecular environment. Although understanding of the interactions is improving, unraveling this interplay will require additional studies of dangling ends with a focus on different salt conditions.

## References

1. Bommarito, S.; Peyret, N.; SantaLucia, J., Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Research* **2000**, 28 (9), 1929-1934.
2. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953**, 171 (4356), 737-8.
3. Devoe, H.; Tinoco, I., Jr., The stability of helical polynucleotides: base contributions. *J Mol Biol* **1962**, 4, 500-17.
4. Freier, S. M.; Sugimoto, N.; Sinclair, A.; Alkema, D.; Neilson, T.; Kierzek, R.; Caruthers, M. H.; Turner, D. H., STABILITY OF XGCGCP, GCGCYP, AND XGCGCYP HELIXES - AN EMPIRICAL ESTIMATE OF THE ENERGETICS OF HYDROGEN-BONDS IN NUCLEIC-ACIDS. *Biochemistry* **1986**, 25 (11), 3214-3219.
5. Breslauer, K. J.; Frank, R.; Blocker, H.; Marky, L. A., PREDICTING DNA DUPLEX STABILITY FROM THE BASE SEQUENCE. *Proceedings of the National Academy of Sciences of the United States of America* **1986**, 83 (11), 3746-3750.
6. Petersheim, M.; Turner, D. H., BASE-STACKING AND BASE-PAIRING CONTRIBUTIONS TO HELIX STABILITY - THERMODYNAMICS OF DOUBLE-HELIX FORMATION WITH CCGG, CCGGP, CCGGAP, ACCGGP, CCGGUP, AND ACCGGUP. *Biochemistry* **1983**, 22 (2), 256-263.
7. Owczarzy, R.; Vallone, P. M.; Gallo, F. J.; Paner, T. M.; Lane, M. J.; Benight, A. S., Predicting sequence-dependent melting stability of short duplex DNA oligomers. *Biopolymers* **1997**, 44 (3), 217-239.
8. Limmer, S.; Hofmann, H. P.; Ott, G.; Sprinzl, M., THE 3'-TERMINAL END (NCCA) OF TRANSFER-RNA DETERMINES THE STRUCTURE AND STABILITY OF THE AMINOACYL ACCEPTOR STEM. *Proceedings of the National Academy of Sciences of the United States of America* **1993**, 90 (13), 6199-6202.
9. Yoon, K.; Turner, D. H.; Tinoco, I.; Haar, F. V. D.; Cramer, F., KINETICS OF BINDING OF U-U-C-A TO A DODECANUCLEOTIDE ANTICODON FRAGMENT FROM YEAST TRNAPHE. *Nucleic Acids Research* **1976**, 3 (9), 2233-2241; Ayer, D.; Yarus, M., THE CONTEXT EFFECT DOES NOT REQUIRE A 4TH BASE PAIR. *Science* **1986**, 231 (4736), 393-395.
10. Ohmichi, T.; Nakano, S.; Miyoshi, D.; Sugimoto, N., Long RNA dangling end has large energetic contribution to duplex stability. *Journal of the American Chemical*

Society **2002**, 124 (35), 10367-10372.

11. Senior, M.; Jones, R. A.; Breslauer, K. J., INFLUENCE OF DANGLING THYMIDINE RESIDUES ON THE STABILITY AND STRUCTURE OF 2 DNA DUPLEXES. *Biochemistry* **1988**, 27 (10), 3879-3885.
12. SantaLucia, J.; Hicks, D., The thermodynamics of DNA structural motifs. *Annual Review of Biophysics and Biomolecular Structure* **2004**, 33, 415-440.
13. Snoussi, K.; Leroy, J. L., Imino proton exchange and base-pair kinetics in RNA duplexes. *Biochemistry* **2001**, 40 (30), 8898-8904.
14. Perez, A.; Noy, A.; Lankas, F.; Luque, F. J.; Orozco, M., The relative flexibility of B-DNA and A-RNA duplexes: database analysis. *Nucleic Acids Research* **2004**, 32 (20), 6144-6151.
15. Petersheim, M.; Turner, D. H., PROTON MAGNETIC-RESONANCE MELTING STUDIES OF CCGGP, CCGGAP, ACCGGP, CCGGUP, AND ACCGGUP. *Biochemistry* **1983**, 22 (2), 269-277.
16. Liu, J. D.; Zhao, L.; Xia, T. B., The dynamic structural basis of differential enhancement of conformational stability by 5' - and 3' -dangling ends in RNA. *Biochemistry* **2008**, 47 (22), 5962-5975.
17. Marky, L. A.; Breslauer, K. J., CALORIMETRIC DETERMINATION OF BASE-STACKING ENTHALPIES IN DOUBLE-HELICAL DNA-MOLECULES. *Biopolymers* **1982**, 21 (11), 2185-2194.
18. Aida, M., AN ABINITIO MOLECULAR-ORBITAL STUDY ON THE SEQUENCE-DEPENDENCY OF DNA CONFORMATION - AN EVALUATION OF INTRA-STRAND AND INTER-STRAND STACKING INTERACTION ENERGY. *Journal of Theoretical Biology* **1988**, 130 (3), 327-335.
19. Holbrook, J. A.; Capp, M. W.; Saecker, R. M.; Record, M. T., Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: Interpretation in terms of coupled processes of formation and association of single-stranded helices. *Biochemistry* **1999**, 38 (26), 8409-8422.
20. Doktycz, M. J.; Paner, T. M.; Amaratunga, M.; Benight, A. S., THERMODYNAMIC STABILITY OF THE 5' DANGLING-ENDED DNA HAIRPINS FORMED FROM SEQUENCES 5'-(XY)<sub>2</sub>GGATAC(T)<sub>4</sub>GTATCC-3', WHERE X, Y = A, T, G, C. *Biopolymers* **1990**, 30 (7-8), 829-845.
21. Guckian, K. M.; Schweitzer, B. A.; Ren, R. X. F.; Sheils, C. J.; Tahmassebi, D. C.; Kool, E. T., Factors contributing to aromatic stacking in water: Evaluation in the

context of DNA. *Journal of the American Chemical Society* **2000**, 122 (10), 2213-2222.

22. Isaksson, J.; Chattopadhyaya, J., A uniform mechanism correlating dangling-end stabilization and stacking geometry. *Biochemistry* **2005**, 44 (14), 5390-5401.

23. Fish, D. J.; Horne, M. T.; Brewood, G. P.; Goodarzi, J. P.; Alemayehu, S.; Bhandiwad, A.; Searles, R. P.; Benight, A. S., DNA multiplex hybridization on microarrays and thermodynamic stability in solution: a direct comparison. *Nucleic Acids Research* **2007**, 35 (21), 7197-7208.

24. Horne, M. T.; Fish, D. J.; Benight, A. S., Statistical thermodynamics and kinetics of DNA multiplex hybridization reactions. *Biophysical Journal* **2006**, 91 (11), 4133-4153.

25. Riccelli, P. V.; Hall, T. S.; Pancoska, P.; Mandell, K. E.; Benight, A. S., DNA sequence, context and multiplex hybridization reactions: Melting studies of heteromorphous duplex DNA complexes. *Journal of the American Chemical Society* **2003**, 125 (1), 141-150.

26. Oligoanalyzer, I., <http://www.idtdna.com/scitools/scitools.aspx>.

27. Wetmur, J. G., DNA PROBES - APPLICATIONS OF THE PRINCIPLES OF NUCLEIC-ACID HYBRIDIZATION. *Critical Reviews in Biochemistry and Molecular Biology* **1991**, 26 (3-4), 227-259.

28. Good, N. E.; Winget, G. D.; Winter, W.; Connolly, T. N.; Izawa, S.; Singh, R. M., Hydrogen ion buffers for biological research. *Biochemistry* **1966**, 5 (2), 467-77.

29. Rouzina, I.; Bloomfield, V. A., Heat capacity effects on the melting of DNA. 1. General aspects. *Biophysical Journal* **1999**, 77 (6), 3242-3251.

30. Wu, P.; Nakano, S.; Sugimoto, N., Temperature dependence of thermodynamic properties for DNA/DNA and RNA/DNA duplex formation. *European Journal of Biochemistry* **2002**, 269 (12), 2821-2830.

31. SantaLucia, J., A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, 95 (4), 1460-1465.

32. Manning, G. S., Counterion condensation theory constructed from different models. *Physica A* **1996**, 231 (1-3), 236-253.

33. Tan, Z. J.; Chen, S. J., Nucleic acid helix stability: Effects of salt concentration, cation valence and size, and chain length. *Biophysical Journal* **2006**, 90 (4), 1175-1190.

34. Owczarzy, R.; You, Y.; Moreira, B. G.; Manthey, J. A.; Huang, L. Y.; Behlke, M.

A.; Walder, J. A., Effects of sodium ions on DNA duplex oligomers: Improved predictions of melting temperatures. *Biochemistry* **2004**, 43 (12), 3537-3554.

35. Nakano, S.; Fujimoto, M.; Hara, H.; Sugimoto, N., Nucleic acid duplex stability: influence of base composition on cation effects. *Nucleic Acids Research* **1999**, 27 (14), 2957-2965.

36. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Graslund, A., Ionic effects on the stability and conformation of peptide nucleic acid complexes. *Journal of the American Chemical Society* **1996**, 118 (24), 5544-5552; Owczarzy, R.; Moreira, B. G.; You, Y.; Behlke, M. A.; Walder, J. A., Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry* **2008**, 47 (19), 5336-5353.

37. Chalikian, T. V.; Volker, J.; Plum, G. E.; Breslauer, K. J., A more unified picture for the thermodynamics of nucleic acid duplex melting: A characterization by calorimetric and volumetric techniques. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, 96 (14), 7853-7858.

38. Riccelli, P. V.; Mandell, K. E.; Benight, A. S., Melting studies of dangling-ended DNA hairpins: effects of end length, loop sequence and biotinylation of loop bases. *Nucleic Acids Research* **2002**, 30 (18), 4088-4093.

39. Quartin, R. S.; Wetmur, J. G., EFFECT OF IONIC-STRENGTH ON THE HYBRIDIZATION OF OLIGODEOXYNUCLEOTIDES WITH REDUCED CHARGE DUE TO METHYLPHOSPHONATE LINKAGES TO UNMODIFIED OLIGODEOXYNUCLEOTIDES CONTAINING THE COMPLEMENTARY SEQUENCE. *Biochemistry* **1989**, 28 (3), 1040-1047.

40. Vesnaver, G.; Breslauer, K. J., THE CONTRIBUTION OF DNA SINGLE-STRANDED ORDER TO THE THERMODYNAMICS OF DUPLEX FORMATION. *Proceedings of the National Academy of Sciences of the United States of America* **1991**, 88 (9), 3569-3573.

41. Isaksson, J.; Acharya, S.; Barman, J.; Cheruku, P.; Chattopadhyaya, J., Single-stranded adenine-rich DNA and RNA retain structural characteristics of their respective double-stranded conformations and show directional differences in stacking pattern. *Biochemistry* **2004**, 43 (51), 15996-16010.

42. Manning, G. S.; Mohanty, U., Counterion condensation on ionic oligomers. *Physica A* **1997**, 247 (1-4), 196-204.

43. Sugimoto, N.; Nakano, S.; Yoneyama, M.; Honda, K., Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Research* **1996**, 24 (22), 4501-4505.

44. Manyanga, F.; Horne, M. T.; Brewood, G. P.; Fish, D. J.; Dickman, R.; Benight, A. S., Origins of the "Nucleation" Free Energy in the Hybridization Thermodynamics of Short Duplex DNA. *Journal of Physical Chemistry B* **2009**, 113 (9), 2556-2563.
45. Zuker, M., Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **2003**, 31 (13), 3406-3415; mfold, <http://dinamelt.bioinfo.rpi.edu/twostate.php>.