

Thermodynamic contributions of 5'- and 3'-single strand dangling-ends to the stability of short duplex DNAs

Rebekah Dickman^{1,2}, Fidelis Manyanga¹, Greg P. Brewwood¹, Daniel J. Fish^{1*}, Cameron A. Fish^{1,3}, Charlie Summers^{1,4}, M. Todd Horne¹, Albert S. Benight^{1,2}

¹Portland Bioscience, Inc., Portland, USA; *Corresponding Author: djf@pdxbio.com

²Portland State University, Portland, USA

³Cleveland High School, Portland, USA

⁴LaSalle Catholic College Preparatory, Milwaukee, USA

Received 27 August 2011; revised 20 October 2011; accepted 11 November 2011

ABSTRACT

Differential scanning calorimetry (DSC) melting analysis was performed on 27 short double stranded DNA duplexes containing 15 to 25 base pairs. Experimental duplexes were divided into two categories containing either two 5' dangling-ends or one 5' and one 3' dangling-end. Duplex regions were incrementally reduced from 25 to 15 base pairs with a concurrent increase in length of dangling-ends from 1 to 10 bases. Blunt-ended duplexes from 15 to 25 base pairs served as controls. An additional set of molecules containing 21 base pair duplexes and a single four base dangling-end were also examined. DSC melting curves were measured in varying concentrations of sodium ion (Na^+). From these measurements, thermodynamic parameters for 5' and 3' dangling-ends were evaluated as a function of dangling end length. 5' ends were found to be slightly stabilizing but essentially constant while the 3' ends were destabilizing with increasing length of the dangling-end. 3' ends also display a stronger dependence on Na^+ concentration. In lower Na^+ environment, the 3' ends were more destabilizing than in higher salt environment suggesting a more significant electrostatic component of the destabilizing interactions. Analysis of thermodynamic parameters of dangling ended duplexes as a function of Na^+ concentration indicated the 3' dangling ends behave differently than 5' dangling ended and blunt-ended duplexes. Molecules with one 5' and one 3' dangling end showed variation in excess specific heat capacity (ΔC_p) when compared to the blunt-ended molecule, while the molecules with two 5' ends had ΔC_p

values that were essentially the same as blunt-ended duplexes. These observations suggested differences exist in duplexes with 3' and 5' dangling ends, which are interpreted in terms of composite differences in interactions with Na^+ , solvent, and terminal base pairs of the duplex.

Keywords: DNA Thermodynamics; Dangling-Ends; DNA Stability

1. INTRODUCTION

The generic scheme for the design of optimum probes for use in multiplex hybridization reactions involves two independent alignment and comparison steps [1-3]. In the first step, all possible probes of a given length covering the targeted region of interest are generated and compared with one another in a pairwise fashion. All possible pairs of strands are aligned in an antiparallel manner and the thermodynamic stability is determined for each alignment. In this process, the first alignment of two strands of the same length is the maximally overlapped one, with no single strand dangling ends. In subsequent alignments, one strand slides incrementally across the other, one base at a time, in the 5' direction. Stability of the duplex at each alignment is determined from the number of complementary base pairs and mismatches in the overlap (duplex) region, and the contribution of the two 5' dangling ends (both of the same length) on the ends of the duplex [4]. For each alignment, the length of the dangling ends increases with concomitant shortening of the overlap duplex region. Once all possible strands have been aligned and compared, and thermodynamic results processed, the optimum probe set can be determined.

During the second stage of probe design, the optimum probe set must be scanned against the entire genome in

which the target sequence resides [5-7]. In this step, the shorter (probe) strand is iteratively scanned against the much longer (target genome) strand, which offers a 5' dangling end and 3' dangling end at each alignment step. The primary aim of this work was evaluation of the thermodynamic contributions of the dangling single strand ends to duplex stability. If this is achieved, a more realistic thermodynamic stability can be determined at each alignment through inclusion of better thermodynamic parameters for the 5' and 3' dangling ends.

In this study, the relative contributions of 5' and 3' dangling ends to overall duplex stability as a function of end length were quantitatively evaluated. Results obtained in three Na⁺ environments indicate 5' ends were stabilizing while 3' ends were neutral or destabilizing. Use of how evaluated parameters augment duplex stability calculations is demonstrated.

2. MATERIALS AND METHODS

2.1. DNA Molecules

DNA strands for duplex molecules were purchased from Integrated DNA Technologies (IDT, Coralville, IA) and subjected to 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 [8]. Sequence composition for the duplex was maintained around 50% G-C base pairs and thermodynamic screening was used to remove any molecules with potential intramolecular structures above 20°C. DNA strands were received from the supplier as a dried powder.

2.2. Quantification of Samples

DNA solutions for DSC melting curves were solvated with NaCl (85, 300 or 1000 mM) and buffered in 10 mM Na₂PO₄ and 0.1 mM EDTA. Solution pH was adjusted to 7.3 ± 0.1 using an Orion 4-Star pH/Conductivity Meter (Thermo Electron Corporation, Beverly, MA). Further details of buffer preparation have been described elsewhere [9,10]. Samples were initially hydrated with 2.0 mL of 85 mM buffered sodium solution and allowed to stand for at least one hour at room temperature. Concentrations of diluted single strand DNAs were determined by absorbance at 260 nm (A260) using the molar extinction coefficient provided by the supplier [11,12]. DNA samples were diluted with buffer so that A260 values ranged from 0.2 - 0.9 optical density (OD) units. Absorbance measurements were made using a Hewlett-Packard 8452A Diode Array Spectrophotometer (Hewlett-Packard Corporation, Palo Alto, CA). Quartz cuvettes having 1.0 cm path lengths were used. After concentration determination, single strands were mixed in a

1:1 ratio and the final duplex concentration adjusted with buffer to around 1.0 mg/ml. Duplexes were annealed at room temperature for at least one hour prior to duplex characterization and melting experiments. After initial mixing, the A260 was noted and compared to the A260 value measured at higher temperatures well above the melting transition. This value was then used to estimate a suitable molar extinction coefficient for the duplex. Subsequently, duplex concentrations were determined from the low and high temperature A260 values.

2.3. Buffer Exchange

To change the ionic strength of the buffer, sample solutions were transferred to a DNA Centricon YM-3 centrifugal filter (Millipore, Bedford, MA) having a molecular weight cut-off of 3000 Daltons. Samples were washed with 2 ml of nanopure water and spun for approximately 90 minutes at 4000 - 4500 rpm. An additional 1 ml of nanopure water was added and the sample spun to dryness. Recovered samples were reconstituted in the desired buffer solution and stored at 4°C. Following buffer exchange, sample concentrations were determined using optical absorbance measurements at 260 nm. Generally sample recovery was greater than 95%. A comparison of pre- and post-transition duplex absorbance values was used to assess possible degradation. Sample quality was also determined by polyacrylamide gel electrophoresis. Electrophoresis was performed on 12% polyacrylamide gels using a Hoeffer MiniVE, vertical electrophoresis mini gel system. Other details of DNA quantification protocols have been described elsewhere [6,13,14].

2.4. Differential Scanning Calorimetry

Thermodynamic parameters, ΔH_{cal} and ΔS_{cal} , of the heat induced melting transitions of duplex DNAs were evaluated from measurements of the excess heat capacity, ΔC_p , versus temperature values obtained by DSC measurements over the range of 10°C to 120°C. ΔC_p values 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 measure excess heat capacity of a sample as it is heat denatured. Prior to loading, samples were degassed by bubbling with helium gas for 10 minutes. Experiments were performed under positive pressure (3 atm) with a heating rate of 2.0°C/min. Concentrations of melted DNA samples were between 0.5 and 1.2 mg/ml. Multiple heating and cooling curves were collected for each sample and analyzed individually. Three different DSC instruments were used to collect the data. Variations in the thermodynamic parameters resulting from small differences in different DSC cell

volumes were within experimental error (<10%). Regardless of the direction of transition (heating or cooling) or DSC employed, all melting curves were highly reproducible. Prior to measuring samples, a standard curve (*i.e.* buffer versus buffer curve) was taken at the same heating rate, over the same temperature range and at the same buffer concentration as sample melting curves. These curves were used to standardize subsequent DSC sample curves.

2.5. Data Analysis

Analysis of DSC data was performed using the CpCalc 2.1 (Applied Thermodynamics, Middlesex, NY) software package supplied by the manufacturer. After being standardized for machine noise by the buffer-buffer curve, ΔC_p versus temperature curves were normalized for total DNA strand concentration, molecular mass, cell volume and partial specific volume. A progressive polynomial baseline was fit to connect linear regions in the lowest and highest temperature portions of the curve to allow for curve integration. The integrated area under the curve provided a measurement of the calorimetric transition enthalpy, ΔH_{cal} , given by

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

For all DSC melting experiments performed, $\Delta C_p(T)$ values were measured from $T_1 = 10^\circ\text{C}$ to $T_2 = 120^\circ\text{C}$. The corresponding calorimetric entropy, ΔS_{cal} , was determined by dividing $\Delta C_p(T)$ by the temperature and integrating over the same temperature range.

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

Reported values of the calorimetric free-energy, $\Delta G_{cal}(T)$, were then determined at $T = 298.15\text{ K}$ using the Gibb's relation,

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

In the analyses performed, it was assumed that the overall difference in excess heat capacity from the beginning to the end of the melting transition was negligibly small, (*i.e.* $\Delta C_p(T_2) - \Delta C_p(T_1) \sim 0$). Consequences of this assumption are that evaluated thermodynamic parameters are most accurate in the transition region and that the parameters are temperature independent.

2.6. Sequence Design

Short DNA molecules were used to evaluate the thermodynamic contributions of single strand dangling ends on overall duplex stability. Duplexes were grouped into three sets according to the orientation of their dangling ends with respect to the core region. By design it is im-

PLICITLY assumed the molecules studied can be partitioned into two separate regions: 1) a duplex region (of length n_D) and 2) single strand dangling ends (of length n_L). Explicit sequences are shown in **Table 1**. Blunt-ended control sequences, designated as set I, are shown in the column on the left. These core sequences range in length from 25 to 15 base pairs. Two types of single strand dangling ended molecules are also shown in **Table 1**. The set II molecules (center column) contain two 5' dangling ends (5'/5') and the set III molecules (right column) contain one 5' and one 3' dangling end (5'/3'). For all molecules of the same n_D , the duplex sequence was maintained to allow for relative thermodynamic calculations. In order to eliminate variations due to end stacking, the final (terminal) base pair at each end of the duplex was kept constant along with the initial base of the dangling end. Although core sequences were not identical, efforts were made to keep the composition of the duplex constant at 50% GC as the length of the central region was decreased.

Duplex molecules shown in **Table 1** are composed of two annealed single strands. For set II molecules the length of each strand was maintained at 25 base pairs. Here, the duplex region was decreased incrementally while the length of 5' single strands on both ends was increased. Ten different dangling-ended sequences were studied with n_D ranging from 24 to 15 base pairs and n_L ranging from 1 to 10 bases. Likewise, for set III molecules, as duplex length decreased, the length of 5' and 3' dangling ends increased. To accomplish this, one strand was lengthened while the other shortened. Six different dangling ended sequences were studied with n_D ranging from 24 to 17 base pairs and n_L ranging from 1 to 8 bases. An additional set of molecules was employed to assess the additive properties of thermodynamic parameters found for dangling ended sequences. They are referred to as set IV molecules and are shown in **Table 2**. Briefly, they are the half molecules of set II and set III DNAs with a duplex core of 21 base pairs and a four base dangling end from one of the strands. For completeness, all four possible orientations of placement of the dangling end were included.

3. RESULTS

3.1. Dangling Ends

Thermodynamic parameters measured as a function of sodium ion concentration ($[\text{Na}^+]$) are listed in **Table 3** for the 27 duplexes studied. Plots of ΔH versus $T\Delta S$ were linear with a mean correlation coefficient of $R^2 = 0.96$ for all molecules in the three Na^+ environments (see **Supplementary Figure 1**). Such a high linearity is typically observed for short duplex melting and indicates entropy-enthalpy compensation. Although not fully understood,

Table 1. DNA sequences and controls. Set I duplexes are blunt-ended control molecules. Set II molecules show the duplexes in the first column (*i.e.* set I) having 5' dangling ends on both strands of the duplex. For clarity, only sequences of the dangling ends are shown explicitly. The core duplex regions, matching the molecules of set I, are indicated by dashed lines. Sequences of set III molecules are given in the third column. For all molecules having a 5' overhang, the terminal 5' stack identity is retained (^{5'}AC/^{3'}G). For those having a 3' overhang, the terminal 3' stack is retained as (^{3'}C/^{5'}AG).

I		II	III
Control Sequences	n_D	5'/5' Sequences	n_L 5'/3' Sequences
5' CATCATCGAACTCAGGTCTCACTTG 3' GTAGTAGCTTGAGTCCAGAGTGAAC	25		
5' CATCATCGAACCAGGTCTCACTTG 3' GTAGTAGCTTGAGTCCAGAGTGAAC	24	5' A----- -----A	1 5' A-----A -----
5' CATCATCGACCAGGCTTCACTTG 3' GTAGTAGCTTGAGTCCGAAGTGAAC	23	5' GA----- -----AG	2 5' GA-----AG -----
5' CAAGTGATCGCTGAGAGAGTTG 3' GTTCACTAGCGACTCTCTCAAC	22	5' AGA----- -----AGA	3 5' AGA-----AGA -----
5' CATGTGATCGCTAGTGAGATG 3' GTACACTAGCGATCACTCATC	21	5' TAGA----- -----AGAT	4 5' TAGA-----AGAT -----
5' CATCTCACAGCGATCACTTG 3' GTAGAGTGTGCTAGTGAAC	20	5' CTAGA----- -----AGATC	5
5' CATCTCACAGCGTAACTTG 3' GTAGAGTGTGCGAATTGAAC	19	5' CCTAGA----- -----AGATCC	6 5' CCTAGA-----AGATCC -----
5' CATCACAGGCGTTACATG 3' GTAGTGTCCGCAATGTAC	18	5' TCCTAGA----- -----AGATCCT	7
5' CATCCGACTCTGCAATG 3' GTAGGCTGAGACGTTAC	17	5' ATCCTAGA----- -----AGATCCTA	8 5' ATCCTAGA-----AGATCCTA -----
5' CATTGGAAGTCCAGTG 3' GTAAGCTTCAGGTCAC	16	5' GATCCTAGA----- -----AGATCCTAG	9
5' CATAGCTGCACGTTG 3' GTATCGACGTGCAAC	15	5' TGATCCTAGA----- -----AGATCCTAGT	10

Table 2. Sequences for set IV molecules. Duplex sequences with either one 3' or 5' dangling end and one blunt end. The dashed duplex region is given by the duplex sequence shown in **Table 1** for $n_D = 21$ and $n_L = 4$.

	IV Sequences	n_D	n_L
a	5' TAGA----- -----	21	4
b	5' ----- -----AGAT	21	4
c	5' -----TAGA -----	21	4
d	5' ----- TAGA-----	21	4

such behavior is reportedly typical for melting of short duplex DNAs [1-3]. Apparently there is only a slight dependence of the slopes and intercepts on the concentration of sodium ions.

3.2. Thermodynamic Contributions to Duplex Stability

From the data of **Table 3**, quantitative estimates can be made on thermodynamic contributions of both 5' and 3'

dangling ends to duplex stability. The melting thermodynamics of short blunt-ended duplexes can be predicted with reasonable accuracy using the nearest-neighbor (n-n) model and corresponding sets of sequence dependent stability parameters [4,5]. The n-n model assumes the duplex region for a dangling ended molecule has the same calculated thermodynamic stability as the matching blunt ended duplex. Thus, the thermodynamic parameters of short duplex DNAs can be reliably parsed into two distinct contributions, those from the duplex region, and those from dangling ends. Following this principle, by preserving the duplex region and varying only the sequences of the dangling ends, it should be possible to determine individual contributions from dangling ends. Of course, reliability of this estimate assumes that interactions between the duplex and single strand regions are separable as described, and that the duplex region and inherent structure and stability remain unaffected (in a thermodynamic sense) by the presence of dangling ends.

3.3. Thermodynamic Characterization

Given the set of DNA molecules and thermodynamic parameters evaluated, and under the umbrella of the afo-

Table 3. Measured Thermodynamic Parameters for sets I, II and III molecules in 85 mM, 300 mM, and 1.0 M Na⁺ salt environments. Average errors for duplexes in all [Na⁺] were: $\Delta H = 4.2$ kcal/mol, $\Delta S = 9$ e.u., $\Delta G = 0.7$ kcal/mol.

	85 mM			300 mM			1.0 M				
	n_L	n_D	ΔH kcal/mol	ΔS e.u.	ΔG_{25} kcal/mol	ΔH kcal/mol	ΔS e.u.	ΔG_{25} kcal/mol	ΔH kcal/mol	ΔS e.u.	ΔG_{25} kcal/mol
I Controls	0	25	-186.9	-530.8	-28.6	-192.3	-544.3	-30.0	-205.2	-578.9	-32.6
	0	24	-171.3	-500.0	-22.2	-169.6	-481.8	-25.9	-181.6	-508.3	-30.0
	0	23	-159.2	-462.5	-21.3	-146.5	-414.8	-22.8	-157.5	-440.3	-26.2
	0	22	-151.9	-444.5	-19.4	-149.4	-426.8	-22.1	-153.9	-433.3	-24.8
	0	21	-142.3	-410.9	-19.8	-152.5	-440.0	-21.3	-156.3	-444.0	-23.9
	0	20	-138.7	-406.0	-17.7	-137.3	-394.3	-19.8	-137.2	-389.3	-21.1
	0	19	-130.2	-381.9	-16.4	-129.2	-372.2	-18.2	-129.5	-368.9	-19.5
	0	18	-121.8	-357.9	-15.1	-121.1	-350.2	-16.7	-121.8	-348.4	-18.0
	0	17	-109.7	-322.3	-13.6	-122.2	-353.3	-16.8	-121.0	-346.8	-17.6
	0	16	-103.3	-305.8	-12.1	-91.9	-269.5	-11.6	-98.8	-286.8	-13.3
II 5'/5' Ends	0	25	-186.9	-530.8	-28.6	-192.3	-544.3	-30.0	-205.2	-578.9	-32.6
	1	24	-185.0	-525.8	-28.2	-191.1	-542.5	-29.3	-190.2	-535.0	-30.7
	2	23	-179.3	-510.8	-27.0	-183.0	-519.3	-28.2	-184.6	-517.0	-30.4
	3	22	-171.9	-491.5	-25.4	-180.5	-515.5	-26.8	-182.8	-513.8	-29.6
	4	21	-161.5	-459.1	-24.6	-163.2	-457.9	-26.7	-169.6	-469.8	-29.5
	5	20	-161.5	-464.8	-22.9	-148.7	-429.5	-20.7	-162.8	-460.0	-25.7
	6	19	-151.3	-444.8	-18.7	-141.0	-407.3	-19.5	-153.0	-440.0	-21.8
	7	18	-141.1	-414.3	-17.6	-137.4	-396.5	-19.2	-144.7	-412.5	-21.7
	8	17	-135.2	-397.3	-16.7	-135.8	-392.8	-18.7	-137.0	-390.0	-20.7
	9	16	-120.3	-356.5	-14.0	-113.5	-331.3	-14.8	-129.2	-371.3	-18.5
III 5'/3' Ends	10	15	-117.6	-347.3	-14.1	-113.3	-329.3	-15.2	-121.5	-348.3	-17.7
	0	25	-186.9	-530.8	-28.6	-192.3	-544.3	-30.0	-205.2	-578.9	-32.6
	1	24	-180.0	-520.0	-24.9	-182.3	-516.7	-28.2	-181.8	-510.0	-29.7
	2	23	-157.2	-457.0	-21.0	-160.2	-454.3	-24.7	-175.8	-490.5	-29.6
	3	22	-155.6	-454.8	-20.0	-161.9	-460.0	-24.7	-174.8	-490.3	-28.6
	4	21	-140.8	-416.7	-16.5	-152.7	-434.1	-23.2	-156.8	-447.2	-23.5
	6	19	-117.0	-345.3	-14.1	-117.6	-339.5	-16.3	-131.5	-380.0	-18.2
	8	17	-102.9	-304.0	-12.2	-117.1	-340.0	-15.8	-125.2	-360.0	-17.8

rementioned assumptions, the following descriptions apply to all experimental DNA molecules studied. Dangling ended molecules contained n_D double strand pairs with n_L single strand dangling ends. In general, experimentally determined values of thermodynamic parameters from the DSC melting transition (*i.e.* enthalpy, entropy and free energy at 25°C) can be represented as $\Delta X_{cal} \equiv \Delta H_{cal}$, ΔS_{cal} or ΔG_{cal} , respectively. For instance, if the measured parameters for set II molecules, $\Delta X_{cal}^{II}(n_D, n_L)$, arise from individual energies of the duplex region and two dangling end regions, the generic molecule with two 5'-dangling ends can be characterized as

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

where $\Delta X_{cal}^I(n_D)$ is the measured thermodynamic parameter for the blunt-ended duplex molecule having the

same core sequence, 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 added twice because two 5' dangling ends are present in each set II molecule. Although explicit sequence dependence of the ends is not considered, utilization of the n-n approximation requires that the terminal stack, that is, the duplex terminal base pair and adjoining single strand base, be the same in all molecules.

Set III molecules differ from those in set II in that they have one 5' and one 3' dangling end. Modifying **Eq.4** above, the measured parameters for set III, $\Delta X_{cal}^{III}(n_D, n_L)$, can be expressed as

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

where $\Delta X_{cal}^I(n_D)$ is again the measured thermodynamic energy for the blunt-ended duplex, $\delta X_{ss}^5(n_L)$ is

the thermodynamic contribution of a 5' single strand dangling-end composed of L bases and $\delta X_{ss}^3(n_L)$ is the thermodynamic contribution of a 3' single strand dangling-end composed of L bases.

From the above expressions and experimentally measured parameters in **Table 3**, individual thermodynamic contributions of $\delta X_{ss}^5(n_L)$ and $\delta X_{ss}^3(n_L)$ were determined using **Eq.4** and **Eq.5**, viz.

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

and

$$\delta X_{ss}^3(n_L) = \Delta X_{cal}^{III}(n_D, n_L) - \Delta X_{cal}^I(n_D) - \delta X_{ss}^5(n_L). \quad (7)$$

The values of $\delta X_{ss}^5(n_L)$ and $\delta X_{ss}^3(n_L)$ evaluated in this manner are plotted versus dangling end length in **Figure 1** for $X = G$ (similar plots for $X = H$ and S are shown in **Supplementary Figure 2**). This plot was generated using the thermodynamic data measured for sets I, II and III in 85 mM, 300 mM and 1.0 M $[Na^+]$ shown in **Table 3**. Solid symbols represent computed values for $\delta X_{ss}^5(n_L)$ and open symbols represent values for $\delta X_{ss}^3(n_L)$.

3.4. 5' Dangling Ends

When comparing the data of **Figure 1**, several trends are clearly seen. The 5' ends are stabilizing and show a very weak dependence on $[Na^+]$. The values of $\delta H_{ss}^5(n_L)$ and $\delta S_{ss}^5(n_L)$ display similar trends as a

function of n_L .

The behavior of $\delta G_{ss}^5(n_L)$ as a function of n_L (**Figure 1**) indicates that, despite variations in $\delta H_{ss}^5(n_L)$ and $\delta S_{ss}^5(n_L)$ for different molecules, the parameters are apparently compensatory in such a way as to render $\delta G_{ss}^5(n_L)$ essentially linear. That is, values of $\delta G_{ss}^5(n_L)$ vary only slightly with increasing length of the dangling end, and the dependence on $[Na^+]$ appears to be minimal. Thermodynamic contributions from the 5' dangling ends were determined by taking the average of $\delta G_{ss}^5(n_L)$ over various values of n_L (in each Na^+ environment). These values are summarized in **Table 4**.

3.5. 3' Dangling Ends

In contrast to 5' ends in **Figure 1**, the 3' ends ($\delta X_{ss}^3(n_L)$ -open symbols), appear to be primarily destabilizing, and more strongly affected by $[Na^+]$. Interestingly, they become less destabilizing with higher $[Na^+]$. The free energy imparted by the 3' end increases to greater positive values with increasing end length up to $n_L = 4$ and remains essentially constant thereafter. The salt dependent stability is also more pronounced and therefore presumably much stronger than for the set I molecules. Apparently, lower sodium ion concentrations (e.g. 85 mM) have an immediate destabilizing influence on duplex stability that manifests in the corresponding effect is observed until $n_L = 4$ at which point $\delta G_{ss}^3(n_L)$ becomes significantly destabilizing. In the 1.0 M salt

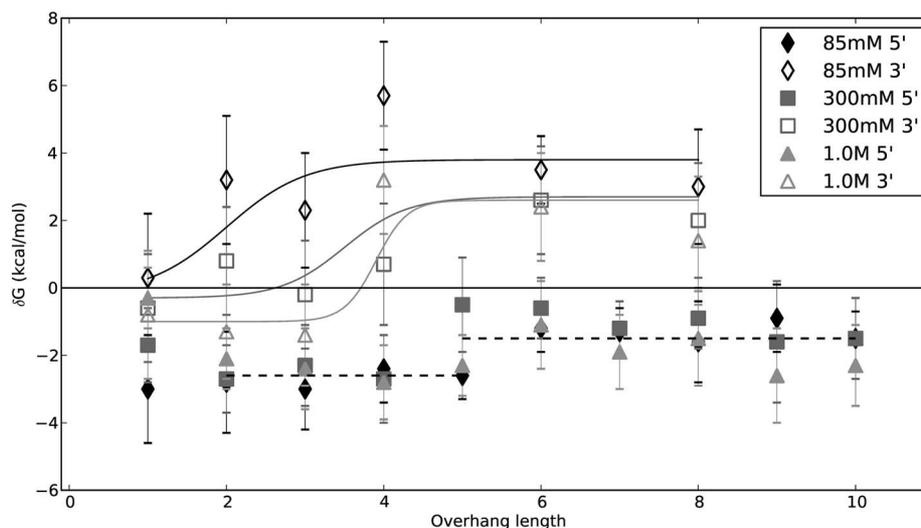


Figure 1. Dangling end parameters calculated for δG^5 (solid markers) and δG^3 (open faced markers) as a function of overhang length (n_L). Comparison of the evaluated δG thermodynamic parameters are shown for the 5' and 3' dangling ends in sets II and III in all three Na^+ environments. This plot shows the relative free-energy contribution (δG) for dangling ends plotted versus overhang length (n_L). Solid curves are drawn through the data to guide the eye. Set II molecules are averaged over all salts. Set III data are shown along with fits of logistical curve **Eq.8**. Equation parameters used to fit the data are: (85 mM) $A = 4$, $r = 2$, $c = -0.2$, $k = 1$, (300 mM) $A = 3$, $r = 2.5$, $c = -0.3$, $k = 2$, (1.0 M) $A = 3.6$, $r = 5$, $c = -1$, $k = 3.5$. The value of $\delta G^3(n_L = 1)$ for 1.0 M was adjusted within error parameters to improve the fit.

Table 4. Thermodynamic contributions of 3' and 5' ends. Length dependent contributions calculated for dangling ends. Negative values are stabilizing and positive values are destabilizing for the duplex.

		$\delta G(n_L)$		
		85 mM	300 mM	1.0 M
5' ends	$n_L = 1$	-3	-1.7	-0.3
	$n_L = 2-4$	-2.6	-2.6	-2.6
	$n_L = 5$	-2.6	-1.5	-1.5
	$n_L > 5$	-1.5	-1.5	-1.5
3' ends	$n_L = 1$	1.8	-0.1	-1.0
	$n_L = 2$	3.3	1.2	-1.0
	$n_L = 3$	3.7	2.5	-0.7
	$n_L = 4$	3.8	2.7	2.3
	$n_L = 6$	3.8	2.7	2.6
	$n_L = 8$	3.8	2.7	2.6

environment, $\delta G_{ss}^3(n_L)$ is initially stabilizing before it becomes increasingly destabilizing at $n_L = 4$. Clearly the values display a more complicated dependence on n_L and $[Na^+]$ than their 5' analogues.

Behavior of the 3' data was modeled with a generalized sigmoidal growth function given by the parameterized logistic equation

$$y = \frac{A}{[1 + \exp(-r(x-k))] + C}. \quad (8)$$

Best fits to the 3' data are shown in **Figure 1**. Parameter values determined for c , k , r and A in each Na^+ environment are listed in the **Figure 1** caption. These expressions were used to predict thermodynamic contributions of 3' dangling ends to the stability of duplex DNA. Specific values computed are summarized in **Table 4**.

3.6. Case Study: $n_D = 21$, $n_L = 4$

The values given in **Table 4** for $\delta X_{ss}^5(n_L)$ and $\delta X_{ss}^3(n_L)$ offer potential for improving thermodynamic predictions for short duplex DNAs containing single strand dangling ends. However, the utility of these values ultimately relies on the validity of assumptions under which the parameters were evaluated. As stated, evaluations of $\delta X_{ss}^5(n_L)$ and $\delta X_{ss}^3(n_L)$ are founded on the n-n model in which individual contributions of the ends are considered separately from those of the duplex region. To test validity of this assumption further, an additional set of DNA molecules were prepared and examined. Set IV molecules provide a secondary method for evaluating thermodynamic contributions of dangling ends where $n_L = 4$, i.e. $\delta X_{ss}^5(n_L = 4)$ and $\delta X_{ss}^3(n_L = 4)$. Molecules of set IV are in a sense "half-molecules" of those in sets II and III containing a duplex region of 21 base pairs and a single dangling end. Sequences IVa and IVb have a

single 5' dangling end of $n_L = 4$, while IVc and IVd contain one 3' dangling end of $n_L = 4$. Duplex and single strand end sequences are the same as those of sets II and III where $n_D = 21$ and $n_L = 4$. Altering **Eq. 4** and **Eq. 5** specifically for set IV molecules with one 5' dangling end and one blunt end leads to the following equation for sequences IVa and IVb,

$$\Delta X_{cal}^{IVa/b}(n_D = 21, n_L = 4) = \Delta X_{cal}^I(n_D = 21) + \delta X_{ss}^5(n_L = 4), \quad (9a)$$

and the equation

$$\Delta X_{cal}^{IVc/d}(n_D = 21, n_L = 4) = \Delta X_{cal}^I(n_D = 21) + \delta X_{ss}^3(n_L = 4). \quad (9b)$$

for sequences IVc and IVd in which molecules have one 3' dangling end and one blunt end. $\Delta X_{cal}^I(n_D = 21)$ is the measured thermodynamic parameter for the blunt-ended duplex with 21 base pairs.

Due to the similarities and differences between the sets, a number of relationships can be found to obtain thermodynamic parameters of the dangling ends. Using **Eq. 9a** and **Eq. 9b** and pertinent results for sets I, II and III molecules along with those from set IV summarized in **Table 5**, estimates on $\delta X_{ss}^5(n_L = 4)$ and $\delta X_{ss}^3(n_L = 4)$ were calculated. Consider the following for $\delta X_{ss}^5(n_L = 4)$:

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

or

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

or

$$\delta X_{ss}^5(n_L = 4) = \Delta X_{cal}^{III}(n_D = 21, n_L = 4) - \Delta X_{cal}^{IVc}. \quad (10c)$$

Table 5. Thermodynamic data for set IV molecules.

$[Na^+]$	Seq.	ΔH	σ	ΔS	σ	ΔG	σ
85 mM	IVa	-153.4	11.9	-436.1	18.7	-23.4	0.5
	IVb	-136.3	3.2	-389.0	13.0	-20.3	0.9
	IVc	-154.5	4.7	-440.2	19.4	-23.3	0.7
	IVd	-137.7	15.2	-392.1	13.6	-20.8	3.1
300 mM	IVa	-160.5	9.2	-456.3	6.0	-24.5	0.2
	IVb	-140.2	7.9	-396.4	10.0	-22.0	0.7
	IVc	-161.5	5.3	-460.4	8.9	-24.2	0.2
	IVd	-143.1	13.0	-407.9	18.3	-21.5	1.2
1.0 M	IVa	-164.8	3.9	-468.5	10.1	-25.1	0.4
	IVb	-151.4	11.3	-430.2	13.9	-23.1	1.8
	IVc	-165.4	1.2	-471.1	7.3	-24.9	1.2
	IVd	-154.5	7.2	-440.1	14.2	-23.3	2.9

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

$$\delta X_{ss}^5(n_L = 4) = \Delta X_{cal}^{III}(n_D = 21, n_L = 4) - \Delta X_{cal}^{IVc}. \quad (11)$$

A total of five equations were generated to determine $\delta X_{ss}^5(n_L = 4)$ and three equations to determine $\delta X_{ss}^3(n_L = 4)$. Specific values in each salt environment are displayed in **Figure 2**. The asterisk on each plot denotes values taken from **Figure 1** where $n_L = 4$. δX_{ss}^5 (lighter bars) is shown on the left of each plot and δX_{ss}^3 (darker bars) is shown on the right. Examination of the histograms in **Figure 2** reveals several interesting observations.

To a first approximation, trends in **Figure 2** are consistent with those in **Figure 1**. Averages of the results from the different calculation schemes provide values in agreement with $\delta X_{ss}^3(n_L)$ and $\delta X_{ss}^5(n_L)$. The 5' dangling ends appear to be stabilizing across all plots, which is in agreement with previous observations. Conversely, the 3' dangling ends appear to be near zero or destabilizing in all but a few cases. These exceptions are free energy values (δG_{ss}^3) at 85 mM $[Na^+]$ and 300 mM $[Na^+]$, where somewhat contradictory results are observed. Further, the plot for δG_{ss}^5 in 85 mM $[Na^+]$ shows significant variability. Apparently, at least in some cases, specific values obtained depend on the particular calculation method employed. The thermodynamic parameters $\delta X_{ss}^3(n_L = 4)$ and $\delta X_{ss}^5(n_L = 4)$ depicted in **Figure 2**, were determined by utilizing variations of the dangling ended molecules and subtracting to determine individual end contributions. Calculations can be grouped together into five different schemes by the generic type of molecules used, *i.e.* blunt ended, single dangling end, sym-

metric double dangling end or non-symmetric double dangling end, and the specific method in which they were used. The five different schemes are depicted in **Table 6**. Under the specific assumptions of the n-n model, the resulting values should reasonably agree regardless of the particular calculation scheme used. For comparison, histograms of values calculated using the same general scheme are paired in **Figure 2**. This comparison reveals the different schemes provide semi-quantitative results in the higher Na^+ environments. In 85 mM, the calculation scheme involving the set III molecules appears to provide a different result than the other schemes using set II or set IV molecules. Although the origins of this observation are not known, results suggest a significant electrostatic effect associated with the 3' dangling end in the set III molecules.

3.7. Counterion Binding

The melting data and corresponding thermodynamic parameters evaluated as a function of $[Na^+]$ provide a means of quantitatively estimating the net Na^+ released upon melting of the short duplex DNAs (as function of both duplex and dangling end length). The release of Na^+ upon melting or the number of ions lost, represented as Δn , can be estimated assuming a simple binding equilibrium and evaluated according to [6],

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

where ΔH^o is the standard state enthalpy of dissociation of the duplex, R is the ideal gas constant and α is a correction term for the sodium ion activity coefficient. A standard value of $\alpha = 0.92$ was assumed throughout

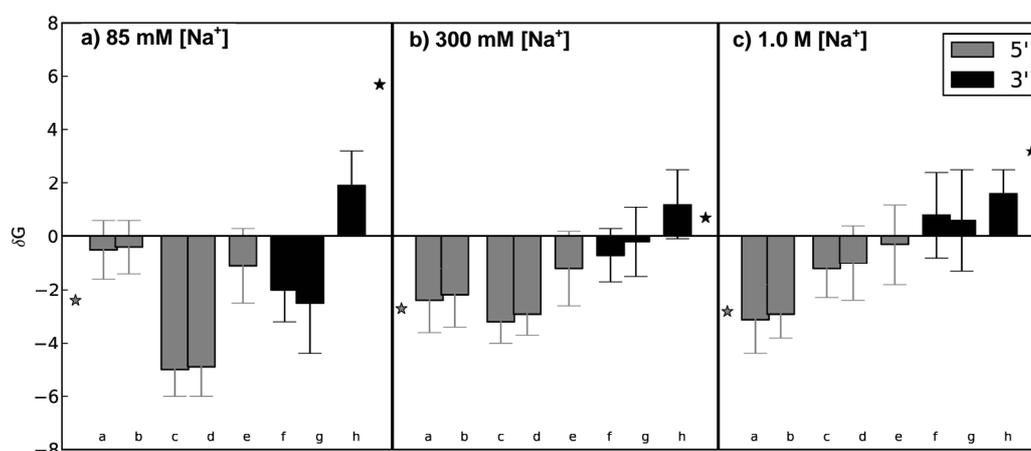


Figure 2. Comparative histograms of δG^5 and δG^3 for $n_L = 4$. Histograms show values of δG^5 and δG^3 determined using thermodynamic data from set IV molecules and relevant molecules from sets II and III. Results from the different calculation schemes are shown in **Table 5**. They are designated as follows: δG^5 (lighter bars) (a) II-IVa, (b) II-IVc, (c) IVa-I, (d) IVc-I, (e) III-IVb, (*) (II-I)/2, δG^3 (darker bars), (f) IVb-I, (g) IVd-I, (h) III-IVa and (*) III-I- δG^5 . Stars (or asterisks) depict values determined from data shown in **Figure 1**.

3.8. Heat Capacities of Dangling Ends

In DSC experiments, the thermodynamic parameters of the melting transition, ΔH_{cal} and ΔS_{cal} are evaluated from the area under the melting curve, $\Delta C_p(T)$ versus Temperature. The experimental transition enthalpy is given by,

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

where T^o is an arbitrary reference temperature and δH is a correction term for small variations in different salt environments. Generally for short duplex DNA the assumption is made that differences between T^o and T are small in the transition region where the thermodynamic parameters are evaluated, thus $\Delta C_p = 0$ [1]. As a matter of practice, calculated values of ΔH_{cal} and ΔS_{cal} are routinely used to predict thermodynamic stabilities of duplexes at temperatures far below (37°C) the actual transition region ($55^\circ\text{C} - 75^\circ\text{C}$), where there is more biological relevance and where practical applications occur. The accuracy of such predictions relies on the validity of the assumption that evaluated parameters are temperature independent and the overall change in excess heat capacity (ΔC_p). 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.

As the technology to measure these transitions has grown in sophistication and precision over the past 15 years, some studies have reported the existence of a relative standard transition heat capacity for all DNA duplexes. Estimates for this value are as high as $100 \text{ cal}\cdot\text{deg}^{-1}\cdot(\text{mol}\cdot\text{base}\cdot\text{pair})^{-1}$ and vary slightly with sequence and salt [2,3,7]. Recently, the average value of $\Delta C_p = 64.6 \pm 21.4 \text{ cal}\cdot\text{deg}^{-1}\cdot(\text{mol}\cdot\text{base}\cdot\text{pair})^{-1}$ has been reported as a good approximation [8]. Inclusion of the ΔC_p parameter allows for more accurate predictions of the transition enthalpy and entropy at temperatures below the transition region [6,7].

For dangling ended molecules, a question arises regarding their effect on, and contributions to ΔC_p . To address this question, plots of ΔH_{cal} versus T_m in all Na^+ environments for sets I, II and III were constructed and are shown in **Figure 4**. Using all of the data, each set of molecules displays a different linear slope. Values of ΔC_p determined from slopes are $31.2 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$ for set I, $25.1 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$ for set II, and $68.9 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$ for set III. These values for sets I and II differ somewhat from the average reported ΔC_p value of $64.6 \pm 21.4 \text{ cal}\cdot\text{deg}^{-1}\cdot(\text{mol}\cdot\text{base}\cdot\text{pair})^{-1}$. However, if values from the lowest $[\text{Na}^+]$ are omitted, ΔC_p values of 54.1, 54.6, $85.2 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$ are found for sets I, II and III respectively, which is in reasonable agreement with the reported best value.

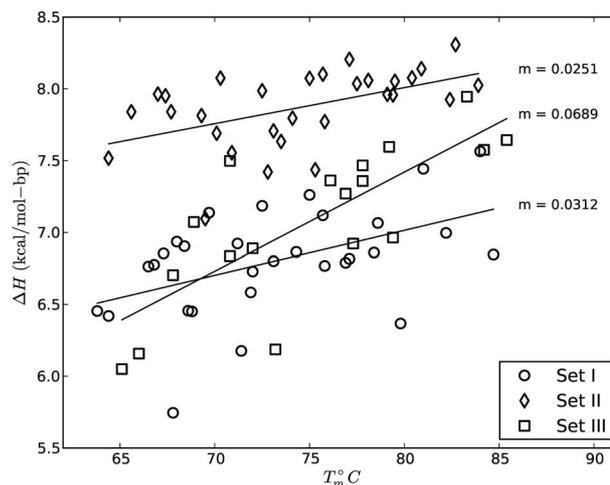


Figure 4. $\Delta H(T_m)$ vs T_m for each type of molecule in all salts. Set II molecules display a ΔC_p value of $25.1 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$, set III molecules have $68.9 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$ and ΔC_p for set I molecules is $31.2 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$.

Differences in ΔC_p values for each type of duplex were estimated using **Eq.14**. For each dangling-ended duplex, the difference between transition enthalpies of the dangling ended molecule of set II and corresponding blunt-ended molecule of set I, evaluated at the transition temperature of the corresponding dangling-ended duplex, T_m^{II} , is

$$\Delta\Delta H(T = T_m^{\text{II}}) = \Delta\Delta H^o + \delta\delta H + \Delta\Delta C_p^{\text{II-I}}(T_m^{\text{I}} - T_m^{\text{II}}) \quad (15)$$

where $\Delta\Delta C_p^{\text{II-I}} = \Delta C_p^{\text{II}} - \Delta C_p^{\text{I}}$ is the difference in the heat capacities of the dangling-ended duplexes of set II and the corresponding blunt-ended duplexes of set I. Clearly a plot of $\Delta\Delta H(T = T_m^{\text{II}})$ versus T_m^{II} should have a slope of $\Delta\Delta C_p^{\text{II-I}}$, viz.

$$d\Delta\Delta H(T = T_m^{\text{II}})/dT_m^{\text{II}} = \Delta\Delta C_p^{\text{II-I}}. \quad (16)$$

An analogous expression for $\Delta\Delta C_p^{\text{III-I}}$ can also be found with $T = T_m^{\text{III}}$. If there were no difference in ΔC_p for set II and set I molecules, a plot of $\Delta\Delta C_p^{\text{II-I}}$ versus T_m^{II} would produce a line having zero slope. Such a line is seen in **Figure 5** for the dangling ended molecules of set II. Conversely, if a plot of $\Delta\Delta C_p^{\text{III-I}}$ versus T_m^{III} produces a line having some non-zero slope, as observed in **Figure 5** for the set III dangling ended molecules, additional factors related to the dangling ends must contribute to their ΔC_p . The slope for set III molecules from the plot in **Figure 5** provides an estimated value of $\Delta\Delta C_p^{\text{III-I}} = 52.5 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$, which is in excess to ΔC_p of the blunt ended molecule. Thus, if the reported average value of $\Delta C_p = 64.6 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$ for the blunt-ended molecules is used, an estimated $\cong 117 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$ is found for the set III molecules. Inclusion of this difference with the noted

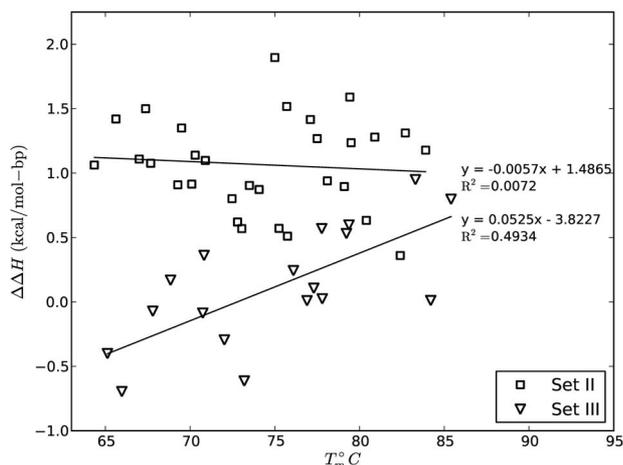


Figure 5. $\Delta\Delta H(T_m)$ vs T_m for each type of molecule in all salts. As compared to the set I molecules, set II molecules display a $\Delta\Delta C_p = -5.8 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$ and set III molecules have $\Delta\Delta C_p = 52.5 \text{ cal}\cdot\text{deg}^{-1}\cdot\text{base}\cdot\text{pair}^{-1}$.

difference in sodium release per phosphate further suggests changes in the set III duplexes. These will be discussed in more detail below.

4. DISCUSSION

4.1. 5' Dangling Ends

A number of factors are involved in determining the relative influence a dangling single strand end will have on the adjoining duplex. These include: the type of molecule (*i.e.* whether it is DNA or RNA), identities of the terminal duplex base pair and initial dangling-end base, and relative orientation of the first dangling base with respect to the terminal duplex base pair. Here, the duplex terminal base pair and adjoining single strand base were held constant in order to focus on relative effects of the end position (5' versus 3') as a function of increasing length. Previous studies have examined thermodynamic contributions of various permutations and combinations of dangling end length, end sequence and sequence of terminating duplex base pair of DNA dangling ends to duplex stability [9,10,13-16]. None, however, have examined effects of end length in an incremental fashion for ends ranging in length from $n_L = 1$ to 10 bases, nor have they compared 5' and 3' ends in different Na^+ environments. However, there remain several places where these results can be compared with published work.

Thermodynamic effects of all possible single base dangling ends were studied by Santa Lucia and coworkers through UV melting analysis [13,17]. In their systematic study, molecules were designed to have a duplex fixed at eight base pairs and single dangling bases attached to the 5' or 3' ends. Relevant sequence dependent

interactions (e.g. those of the terminal $5' \text{AC}/3' \text{G}$ stack) determined in their study should be comparable with our results in 1.0 M $[\text{Na}^+]$. After conversion to 25°C, they report an observed stabilization of $\Delta G_{25} = -1.21 \text{ kcal/mol}$. It appears that our value of $\delta G_{25}(n_L = 1)$ at 25°C for the $5' \text{AC}/3' \text{G}$ stack of -0.3 kcal/mol is not nearly as stabilizing as reported. However, for longer ends where n_L is greater than 1, an average value of $\delta G_{25} = -2.6 \text{ kcal/mol}$ was obtained, which is in better agreement with published results.

Over 20 years ago, Doktycz *et al.* published results from melting studies of four-base dangling-ended DNA hairpin molecules [14]. The general sequence, $5'(\text{XY})_2(\text{GGATAC})_2(\text{T})_4$, naturally folds to form a hairpin with a six base duplex stem, a T_4 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 $5' \text{AC}/3' \text{G}$), the reported δG_{25} had an average value of -1.17 kcal/mol in 115 mM $[\text{Na}^+]$. In this salt environment, for $n_L = 4$ our data has coalesced to an average value of -1.5 kcal/mol , in reasonable agreement with published results.

Ohmichi published results of melting studies using an eight base pair duplex molecule with 5' dangling ends varying from one to four bases [9]. The terminating sequence in their molecules was slightly different, $5' \text{AG}/3' \text{C}$ instead of $5' \text{AC}/3' \text{G}$, which prohibits a direct comparison. Nonetheless, trends reported as a function of increasing length can be considered. For a single base 5' dangling end in 1.0 M $[\text{Na}^+]$, a stabilization of $\delta G_{25} = -0.3 \text{ kcal/mol}$ was reported. Addition of a second dangling end base increased stabilization to -0.4 kcal/mol . This trend seemed to reach a constant value at three bases with $\delta G_{25} = -0.6 \text{ kcal/mol}$. This increase in stability coupled to the lengthening of the single strand dangling end is qualitatively consistent with our observations.

4.2. 3' Dangling Ends

Compared to 5' dangling ends, fewer studies of 3' dangling ends have been performed. Overall, reports have found that 3' dangling ends make favorable contributions to duplex stability, and are thus stabilizing, but less so than their 5' counterparts [4,9]. In contrast we found 3' ends to be generally destabilizing. Although unexpected, observations of destabilizing dangling ends are not unprecedented. The terminal stacks, $5' \text{GT}/3' \text{A}$, $5' \text{T}/3' \text{AC}$ and $5' \text{T}/3' \text{TA}$ were found by Santa Lucia to be mildly destabilizing [13]. Additionally comparison of the published temperature corrected value for the $5' \text{C}/3' \text{AG}$ stack, $\delta G_{25} = -1.0 \text{ kcal/mol}$ [13], is in good agreement with our results in 1.0 M $[\text{Na}^+]$, where $\delta G_{25} = -0.8 \text{ kcal/mol}$. Perhaps by focusing only on single dangling bases in

high salt, the case found here to have the greatest stabilizing effect, the appearance of 3' dangling end destabilization was overlooked.

4.3. Origins of Stabilization

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. In the duplex state, DNA adopts the preferable B-form, which persists to some degree in the single strand state. NMR studies of single strand DNA hexamers with multiple A-A base stacks showed that in DNA, the imidazole stacks above the pyrimidine in the 5' to 3' direction [18].

A systematic review of crystal structures from the protein database demonstrated for DNA that addition of a single strand base on the 5' end is positioned in such a way that it can freely interact with the hydrogen bonds of the terminal base pair. In contrast, a 3' base end is positioned away from the same hydrogen bonds, and therefore is less likely to experience such stabilization [19]. Thus, placement of the dangling base is optimal for terminal base pair interactions in DNA on the 5' end but when an additional base is added to the 3' end of DNA minimal overlap occurs which apparently translates to fewer stabilizing interactions.

4.4. Structural Perturbations

Our results indicate a 3' dangling end is generally destabilizing to a DNA duplex. Previous studies of counterion binding to duplex DNA suggests fewer Na⁺ ions bind near the ends compared to in the middle [6,20]. This suggests differences between the dangling ended and blunt molecules should result in negligible changes in counterion binding if the duplex region of the dangling-ended duplex is not affected by the ends. Comparison of the blunt ended duplexes and set II molecules support this assumption. The net counterion release per phosphate is the same for the two sets (I and II) of molecules and the plots of $\Delta\Psi$ vs. n_L (in **Figure 3**) are the same. Conversely, for the set III molecules having 3'/5' dangling ends, the plot of $\Delta\Psi$ vs. n_L is approximately 15% lower than for the set I and set II molecules and indicates a net lower Na⁺ release during melting suggesting the duplex region for set III is perturbed in some way that results in slight differences of the associated counterion binding properties.

When DNA molecules anneal from their single strand state to form a duplex state, a net change in solvent exposed surface occurs. This change is accompanied by the burying of hydrophobic residues, which contributes to ΔC_p . Differences in the estimated ΔC_p values for set II, when compared to set III (**Figures 4** and **5**), further

support the idea of subtle differences between the duplex regions. Comparison of ΔC_p values for the blunt-ended set I molecules and the set II molecules revealed an essentially constant $\Delta\Delta C_p$ with a $\Delta\Delta H_{cal}$ difference of about 1 cal·deg⁻¹·base·pair⁻¹. There was also no marked length dependence, indicating that most of the stability came from interactions of the first base with the terminal base pair, presumably due to the favorable stacking interactions. Conversely, comparison of ΔC_p values for set III and set I molecules revealed a significant difference, $\Delta\Delta C_p = 52.5$ cal·deg⁻¹·base·pair⁻¹. Here $\Delta\Delta H_{cal}$ is initially around 0.5 and decreases to -0.5 as the length of the dangling end increases to $n_L = 10$. This suggests an initial smaller buried surface area as compared to the 5' dangling end, which is lost as the dangling end increases in length. The continued loss of enthalpy suggest that with longer dangling ends, the duplex region itself may be perturbed in this particular molecular environment.

4.5. Predictive Ability and Applications to Probe Design

Ascertaining specific thermodynamics involved in probe/target alignment and being able to further predict energies of all possible alignments is key to optimal probe design. The ability to design probes with exquisite accuracy is imperative to successfully locate target sequences differing by as little as a single nucleotide. In fact, molecules used in this study were designed to mimic those that might occur in a multiplex hybridization reaction (such as on a DNA microarray) where dangling ends presumably occur with a moderate to high frequency. The more specific the predictive ability of thermodynamic binding properties, the more effective probe design can be achieved.

In the n-n model the free energy of melting a duplex molecule is given by

$$\Delta G^o = \Delta G_{initiation}^o + \sum \Delta G_{stack}^o + \Delta G_{additional}^o \quad (17)$$

For calculations involving the duplex region, this model uses combinations of the 10 possible n-n values (ΔG_{stack}^o) experimentally determined by a number of independent labs that are generally in good agreement with one another [17]. The first term, $\Delta G_{initiation}^o$, is the cost required to begin the annealing process and therefore has a positive free energy contribution. This value has recently been determined by our group, as well as other investigators [21,22]. The last term, $\Delta G_{additional}^o$, encompasses any extra terms such as those arising from symmetry considerations, a terminating A·T base pair or single strand dangling ends. The sum of terms estimates the total free energy.

Current prediction programs relying on the n-n model to calculate thermodynamic properties are limited by the

Table 7. ΔG values predicted using length dependent overhang parameters. Sequences used to predict ΔG_{25} kcal/mol are shown in column one. The dashed duplex region is given by the duplex sequence shown in **Table 1** for $n_D = 21$ and $n_L = 4$. Unified N-N values were used in calculations [13]. All predicted values were corrected with an averaged difference for set I (*i.e.* $\Delta G_{\text{exp}} - \Delta G_{\text{mfold}}$, 85 mM = +6.2, 300 mM = +6.7), and (1000 mM = +7.8).

		85 mM [Na ⁺]			300 mM [Na ⁺]			1.0 M [Na ⁺]		
Sequence		Exp	Mfold	NN	Exp	Mfold	NN	Exp	Mfold	NN
5'	-----									
3'	-----	-19.8	-18.6	-19.2	-21.3	-20.9	-21.5	-23.9	-22.4	-23.0
5'	TAGA-----									
3'	-----AGAT	-24.6	-20.7	-24.4	-26.7	-23.1	-26.7	-29.5	-25.8	-28.2
5'	TAGGA-----	*			*			*		
3'	-----AAGTCGAT		-20.7	-23.5		-23.1	-24.9		-25.8	-26.4
5'	TAGA-----AGAT									
3'	-----	-16.5	-20.5	-18.1	-23.2	-22.9	-21.4	-23.5	-24.6	-23.3
5'	ATCGA-----AGAATCT	*			*			*		
3'	-----		-20.5	-18.0		-22.9	-20.5		-24.6	-22.1

quality of parameters used in the calculation. Mfold is one such program readily available via the Internet [23]. In this program, a computational algorithm searches for the most stable structure formed from the sequence of two DNA strands, through a calculation of the thermodynamic stabilities using tabulated n-n parameter values. To include effects of dangling ends, the specific value of -1.18 kcal/mol for the 5' dangling end (^{5'}AC^βG) and -1.05 kcal/mol for the 3' dangling end (^{3'}C^βAG) are added to n-n calculations [13]. Potential stability differences due to dangling ends longer than one base are not considered.

To test the applicability and utility of dangling end parameters evaluated here, five sequences having a 21 base pair duplex region were designed with variable end lengths. Calculated thermodynamic parameters were generated using two methods, Mfold and the n-n model. Sequences differed in the placement and length of dangling ends. The sequences were: 1) a 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 having 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. Where available, experimental data has been included as a comparison. Results of computations are shown in **Table 7**.

Upon closer examination, the n-n model and Mfold stability predictions are quite comparable for 5' ended molecules. However, Mfold predicts a greater stability for the 3' ended molecules than that predicted using the end parameters described here. Since no reported measurements have been made of these specific 3' ends destabilizing the duplex, it is not surprising the standard program (Mfold) overestimates the stability of those molecules. These comparisons suggest the dangling ended parameters in **Table 4** may provide more accurate quan-

titative predictions of the stability of dangling ended molecules. Confirmation of the ultimate practical utility of the evaluated dangling end parameters must be demonstrated through more accurate probe design and improved quantitative performance of multiplex hybridization reactions.

5. ACKNOWLEDGEMENTS

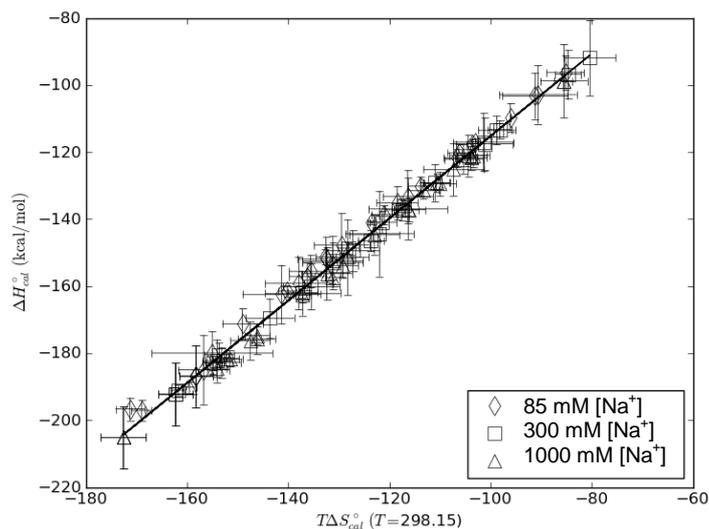
Portions of this work appeared in a thesis by Rebekah Dickman submitted in partial fulfillment for requirements of the Master of Science in Chemistry at Portland State University, December 2010. This work was supported by grants GM080904 and GM084603 from the National Institutes of Health.

REFERENCES

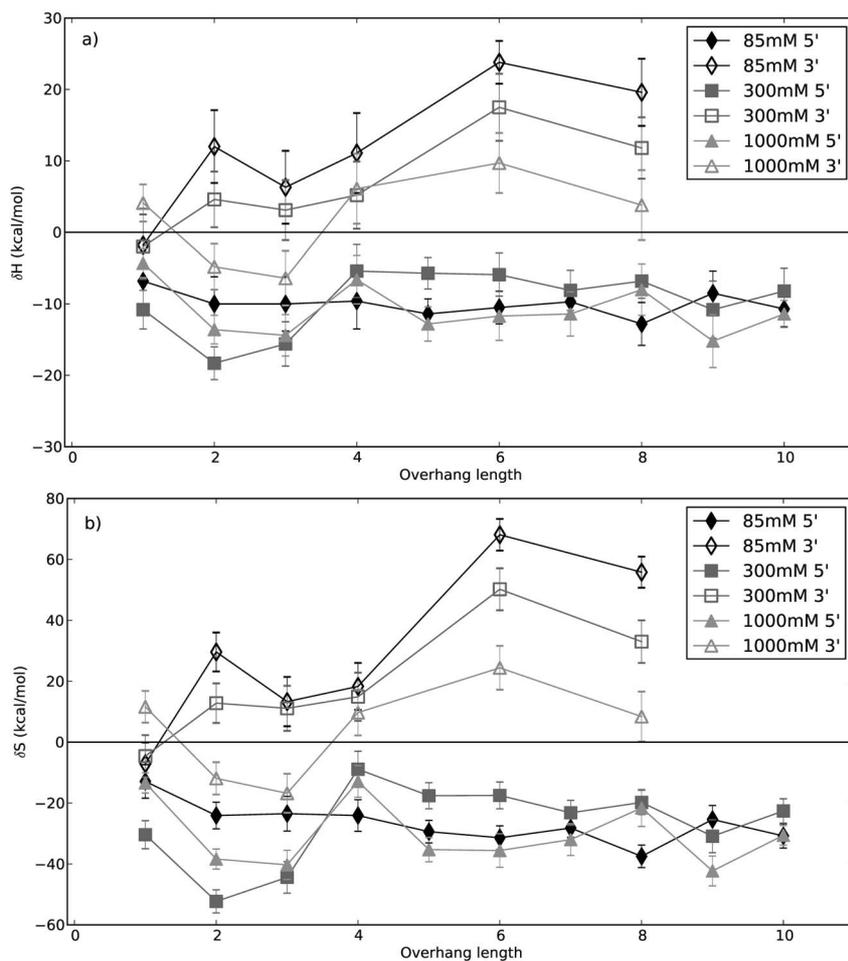
- [1] Petersheim, M. and Turner, D.H. (1983) Base-stacking and base-pairing contributions to helix stability—Thermodynamics of double-helix formation with CCGG, CCGGP, CCGGAP, ACCGGP, CCGGUP, and ACCGGUP. *Biochemistry*, **22**, 256-263. doi:10.1021/bi00271a004
- [2] Rouzina, I. and Bloomfield, V.A. (1999) Heat capacity effects on the melting of DNA. 1. General aspects. *Biophysical Journal*, **77**, 3242-3251. doi:10.1016/S0006-3495(99)77155-9
- [3] Wu, P., Nakano, S. and Sugimoto, N. (2002) Temperature dependence of thermodynamic properties for DNA/DNA and RNA/DNA duplex formation. *European Journal of Biochemistry*, **269**, 2821-2830. doi:10.1046/j.1432-1033.2002.02970.x
- [4] Freier, S.M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M.H. and Turner, D.H. (1986) Stability of XGCGP, GCGCYP, and XGCGCYP helices—An empirical estimate of the energetics of hydrogen-bonds in nucleic-acids. *Biochemistry*, **25**, 3214-3219. doi:10.1021/bi00359a020
- [5] SantaLucia, J. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodyna-

- mics. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 1460-1465. [doi:10.1073/pnas.95.4.1460](https://doi.org/10.1073/pnas.95.4.1460)
- [6] Owczarzy, R., You, Y., Moreira, B.G., Manthey, J.A., Huang, L.Y., Behlke, M.A. and Walder, J.A. (2004) Effects of sodium ions on DNA duplex oligomers: Improved predictions of melting temperatures. *Biochemistry*, **43**, 3537-3554. [doi:10.1021/bi034621r](https://doi.org/10.1021/bi034621r)
- [7] Holbrook, J.A., Capp, M.W., Saecker, R.M. and Record, M.T. (1999) 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*, **38**, 8409-8422. [doi:10.1021/bi990043w](https://doi.org/10.1021/bi990043w)
- [8] Chalikian, T.V., Volker, J., Plum, G.E. and Breslauer, K.J. (1999) 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*, **96**, 7853-7858. [doi:10.1073/pnas.96.14.7853](https://doi.org/10.1073/pnas.96.14.7853)
- [9] Ohmichi, T., Nakano, S., Miyoshi, D. and Sugimoto, N. (2002) Long RNA dangling end has large energetic contribution to duplex stability. *Journal of the American Chemical Society*, **124**, 10367-10372. [doi:10.1021/ja0255406](https://doi.org/10.1021/ja0255406)
- [10] Senior, M., Jones, R.A. and Breslauer, K.J. (1988) Influence of dangling thymidine residues on the stability and structure of 2 dna duplexes. *Biochemistry*, **27**, 3879-3885. [doi:10.1021/bi00410a053](https://doi.org/10.1021/bi00410a053)
- [11] Cantor, C.R., Warshaw, M.M. and Shapiro, H. (1970) Oligonucleotide Interactions III. *Circular Dichroism Studies of the Conformation of Deoxyoligonucleotides*. *Biopolymers*, **9**, 1059-1077. [doi:10.1002/bip.1970.360090909](https://doi.org/10.1002/bip.1970.360090909)
- [12] Cavaluzzi, M.J. and Borer, P.N. (2004) Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic Acids Research*, **32**, e13. [doi:10.1093/nar/gnh015](https://doi.org/10.1093/nar/gnh015)
- [13] Bommarito, S., Peyret, N. and SantaLucia, J. (2000) Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Research*, **28**, 1929-1934. [doi:10.1093/nar/28.9.1929](https://doi.org/10.1093/nar/28.9.1929)
- [14] Doktycz, M.J., Paner, T.M., Amaratunga, M. and Benight, A.S. (1990) Thermodynamic stability of the 5'dangling-ended dna hairpins formed from sequences 5'-(xy)2ggatc(t)4gtatcc-3', where x, y = A, T, G, C. *Biopolymers*, **30**, 829-845. [doi:10.1002/bip.360300718](https://doi.org/10.1002/bip.360300718)
- [15] Quartin, R.S. and Wetmur, J.G. (1989) Effect of ionic-strength on the hybridization of oligodeoxynucleotides with reduced charge due to methylphosphonate linkages to unmodified oligodeoxynucleotides containing the complementary sequence. *Biochemistry*, **28**, 1040-1047. [doi:10.1021/bi00429a018](https://doi.org/10.1021/bi00429a018)
- [16] Riccelli, P.V., Mandell, K.E. and Benight, A.S. (2002) Melting studies of dangling-ended DNA hairpins: Effects of end length, loop sequence and biotinylation of loop bases. *Nucleic Acids Research*, **30**, 4088-4093. [doi:10.1093/nar/gkf514](https://doi.org/10.1093/nar/gkf514)
- [17] SantaLucia, J. and Hicks, D. (2004) The thermodynamics of DNA structural motifs. *Annual Review of Biophysics and Biomolecular Structure*, **33**, 415-440. [doi:10.1146/annurev.biophys.32.110601.141800](https://doi.org/10.1146/annurev.biophys.32.110601.141800)
- [18] Isaksson, J., Acharya, S., Barman, J., Cheruku, P. and Chattopadhyaya, J. (2004) Single-stranded adenine-rich DNA and RNA retain structural characteristics of their respective double-stranded conformations and show directional differences in stacking pattern. *Biochemistry*, **43**, 15996-16010. [doi:10.1021/bi048221v](https://doi.org/10.1021/bi048221v)
- [19] Isaksson, J. and Chattopadhyaya, J. (2005) A uniform mechanism correlating dangling-end stabilization and stacking geometry. *Biochemistry*, **44**, 5390-5401. [doi:10.1021/bi047414f](https://doi.org/10.1021/bi047414f)
- [20] Manning, G.S. and Mohanty, U. (1997) Counterion condensation on ionic oligomers. *Physica A*, **247**, 196-204. [doi:10.1016/S0378-4371\(97\)00413-5](https://doi.org/10.1016/S0378-4371(97)00413-5)
- [21] Sugimoto, N., Nakano, S., Yoneyama, M. and Honda, K. (1996) Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Research*, **24**, 4501-4505. [doi:10.1093/nar/24.22.4501](https://doi.org/10.1093/nar/24.22.4501)
- [22] Manyanga, F., Horne, M.T., Brewood, G.P., Fish, D.J., Dickman, R. and Benight, A.S. (2009) Origins of the "nucleation" free energy in the hybridization thermodynamics of short duplex DNA. *Journal of Physical Chemistry B*, **113**, 2556-2563. [doi:10.1021/jp809541m](https://doi.org/10.1021/jp809541m)
- [23] Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, **31**, 3406-3415. [doi:10.1093/nar/gkg595](https://doi.org/10.1093/nar/gkg595)

Supplementary Figures



Supplementary Figure 1. ΔH_{cal}° versus $T\Delta S_{cal}^{\circ}$ ($T = 298.15$ K) for the 27 duplex DNAs in 85 mM, 300 mM and 1.0 M $[\text{Na}^+]$.



Supplementary Figure 2. Comparison of the evaluated thermodynamic parameters for the 5' and 3' dangling ends in sets II and III ((a) δH and (b) δS), in the three Na^+ environments plotted versus overhang length, n_L .