

## **Supporting Information:**

# **Folding Cooperativity in RNA and DNA is Dependent on Position in the Helix**

Nathan A. Siegfried, Shana L. Metzger, and Philip C. Bevilacqua\*  
*Department of Chemistry, The Pennsylvania State University, University Park,  
Pennsylvania 16802*

### **Contents:**

- S.1: Sample plots for UV-melt experiments**
- S.2: Analysis of oligonucleotides by PAGE and TOF-ES mass spectroscopy**
- S.3: Procedures for error propagation**

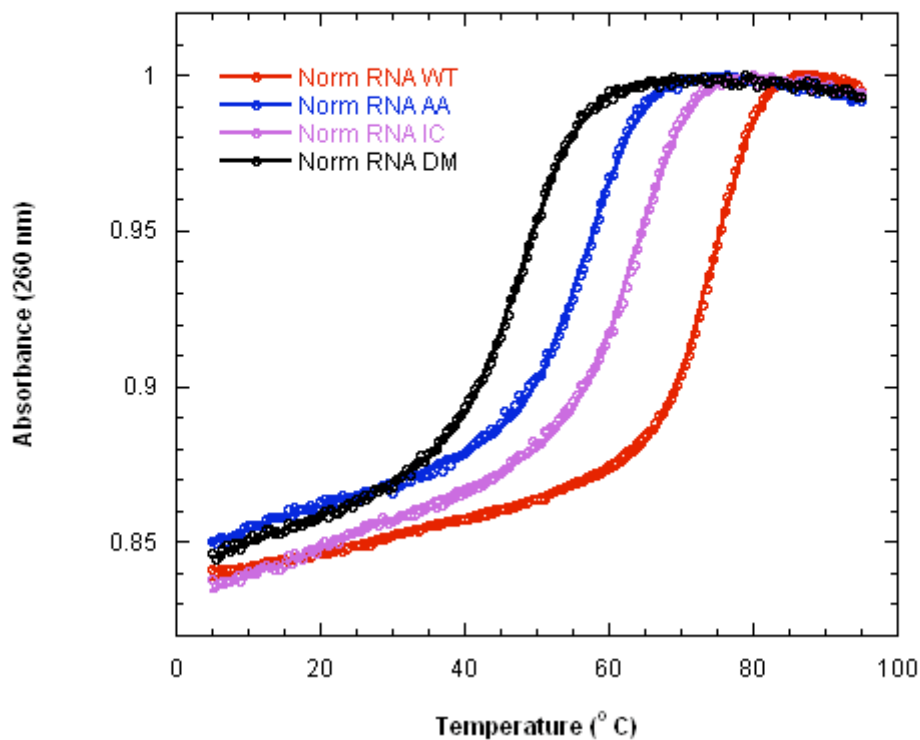
---

\* To whom correspondence should be addressed. Email: [pcb@chem.psu.edu](mailto:pcb@chem.psu.edu)

### *S.1: Sample plots from UV-melting experiments*

**Figure S.1:** Absorbance versus temperature plots resulting from UV-monitored thermal denaturation.

Low salt, internal registers for RNA are shown here as representative of typical data. All data are normalized by dividing absorbance values by the maximal absorbance. This was done so that percent hyperchromicity could be compared between sequences. Data were fit to a two-state model with sloping baselines and analyzed using a Marquadt algorithm for non-linear curve fitting in Kaleidagraph v3.5 (Synergy Software); see Materials and Methods. Fits are represented by solid lines; agreement between the fits and the experimental data is consistent with the accuracy of the model used. The drop in  $T_m$  caused by each mutation can clearly be seen from the mid-point of the transitions.



## S.2: Analysis of oligonucleotides by PAGE and TOF-ES mass spectroscopy

Purity of all oligonucleotides was confirmed by electrophoresis, LC-mass spectroscopy, or both. For electrophoresis, oligonucleotides were 5' end-labeled using [ $\gamma$ - $^{32}\text{P}$ ]ATP and polynucleotide kinase (PNK) (New England Biolabs). Samples were mixed with an equal volume of 95% formamide loading buffer/20 mM EDTA and loaded onto a 10% acrylamide/7 M urea gel, which revealed a single band for all RNA oligonucleotides. The electrophoretic mobility of these bands was in approximate correlation with the thermodynamic stability of the oligonucleotide, *i.e.* the least stable (double mutant) oligonucleotide migrated the shortest distance, supporting folding events on the urea gel. The presence of a single band supported a homogenous molecular population. The DNA samples did not resolve into highly distinct populations under these conditions, but rather migrated in broad bands with uneven intensity. We suspected that different extents of denaturation were occurring for DNA on this denaturing gel, suggesting secondary structure was partially disrupted. Complete disruption of secondary structure was unsuccessfully attempted through the use of glyoxal modifications prior to gel electrophoresis. Next, an ultra-denaturing 40% formamide/7 M urea/6% acrylamide gel was used, which caused the radiolabeled DNA samples to resolve into single bands of equivalent migration consistent with a homogenous molecular population.

For LC-mass spectroscopy, we used a Waters Micromass LCT Premier time-of-flight mass spectrometer equipped with a Waters Alliance 2695 Separations Module and utilizing electrospray ionization (The Huck Institute of Life Sciences Proteomics and Mass Spectrometry Core Facility, The Pennsylvania State University).

**Table S.1:** Summary of Mass Spectroscopy Data

Sequence	Calculated MW (amu)	Measured MW (amu)
<b>DNA</b>		
WT	6099.0	6100
Ext AA	6108.0	6109
Ext IC	6084.0	6081
Ext DM	6093.0	6094
Int AA	6108.0	6109
Int IC	6084.0	6085
Int DM	6093.0	6094
<b>RNA</b>		
WT	6320.8	6320
Ext AA	6343.8	6343
Ext GG	6375.8	6376
Ext IC	6305.8	6305
Ext DM	6328.8	6328
Int AA	6343.8	6343
Int IC	6305.8	6305
Int UU	6282.7	6283
Int DM	6328.8	6328

### S.3: Procedures for error propagation

Error for  $\Delta H$  and  $T_m$  were obtained from non-linear curve fits of UV-melt data to a two-state model as described in S.1 and Materials and Methods. Error in a value  $x$  is represented as  $\sigma_x$ . Propagation of this error into  $\Delta S$ ,  $\Delta G$ , and the averages of these errors were carried out according to Bevington for a general function  $x=f(u,v)$ :(1)

$$(\sigma_x)^2 = (\sigma_u)^2 (\partial x / \partial u)^2 + (\sigma_v)^2 (\partial x / \partial v)^2 + 2(\sigma_{uv})^2 (\partial x / \partial u)(\partial x / \partial v) \quad \text{Eq 1}$$

Application of this form to the standard equation  $\Delta G = \Delta H - T\Delta S$  gives:

$$\sigma_{\Delta G}^2 = (\partial \Delta G / \partial \Delta H)^2 \sigma_{\Delta H}^2 + (\partial \Delta G / \partial \Delta S)^2 \sigma_{\Delta S}^2 + 2(\sigma_{\Delta H \Delta S})^2 (\partial \Delta G / \partial \Delta H)(\partial \Delta G / \partial \Delta S) \quad \text{Eq 2}$$

$$(\sigma_{\Delta G})^2 = (\sigma_{\Delta H^\circ})^2 + T^2 (\sigma_{\Delta S^\circ})^2 - 2T(\sigma_{\Delta H^\circ \Delta S^\circ})^2 \quad \text{Eq 3}$$

The square of the covariance between  $\Delta H^\circ$  and  $\Delta S^\circ$ ,  $(\sigma_{\Delta H \Delta S})^2$ , can be expressed in terms of the correlation coefficient ( $R_{\Delta H \Delta S}$ ) of a plot of  $\Delta H^\circ$  vs  $\Delta S^\circ$  or  $\sigma_{\Delta H}$  vs  $\sigma_{\Delta S}$ .(2)

$$(\sigma_{\Delta G})^2 = (\sigma_{\Delta H^\circ})^2 + T^2 (\sigma_{\Delta S^\circ})^2 - 2T(R_{\Delta H^\circ \Delta S^\circ}) \sigma_{\Delta H^\circ} \sigma_{\Delta S^\circ} \quad \text{Eq 4}$$

Typical correlation values of  $\Delta H^\circ$  and  $\Delta S^\circ$  are  $\sim 0.99$ ,(2) and similar values were obtained for our sequences (data not shown). This allowed the R value to be approximated as unity. To determine the value of  $\sigma_{\Delta S}$ , the equation  $\Delta S = \Delta H/T_m$  was treated according to Eq 1:

$$\sigma_{\Delta S} = \frac{1}{T_m} \sqrt{\sigma_{\Delta H^\circ}^2 + \Delta S^2 \sigma_{T_m}^2 - 2\Delta S \sigma_{\Delta H^\circ} \sigma_{\Delta T_m}} \quad \text{Eq 5}$$

The error in  $T_m$  was typically quite low. This led to the second and third terms under the radical in Eq. 5 contributing only 0.007% and 1% on average to the radical near a temperature of 310.15 K, where  $\Delta G_{37}^\circ$  is calculated, indicating that they make a negligible contribution to  $\sigma_{\Delta G}$ . As a result, the following simplification of Equation 4 is possible:

$$(\sigma_{\Delta G})^2 = (\sigma_{\Delta H^\circ})^2 + T^2 \frac{\sigma_{\Delta H^\circ}^2}{T_m^2} - \frac{2T\sigma_{\Delta H^\circ}^2}{T_m} \quad \text{Eq 6}$$

$$\sigma_{\Delta G} = \sigma_{\Delta H^\circ} \times \frac{|T - T_m|}{T_m} \quad \text{Eq 7}$$

Eq 7 holds true for temperature values away from the  $T_m$ . Values reported in the paper are at 310.15 K, and typical  $T_m$ s are in the 338.15-353.15 K range. Eq 7 represents the error in a single  $\Delta G$  value resulting from a UV melt. A given sequence was melted three or more times. To evaluate the final error in the average  $\Delta G$ , Eq 8 was treated according to Eq 1 to give:

$$\Delta G_{Avg} = \frac{\Delta G_1 + \Delta G_2 + \dots + \Delta G_n}{n} \quad \text{Eq 8}$$

$$\sigma_{\Delta G_{Avg}}^2 = \left( \frac{\partial \Delta G_{Avg}}{\partial \Delta G_1} \right)^2 \sigma_{\Delta G_1}^2 + \dots + \left( \frac{\partial \Delta G_{Avg}}{\partial \Delta G_n} \right)^2 \sigma_{\Delta G_n}^2 \quad \text{Eq 9}$$

$$\sigma_{\Delta G_{Avg}}^2 = \frac{1}{n^2} \times [\sigma_{\Delta G_1}^2 + \dots + \sigma_{\Delta G_n}^2] \quad \text{Eq 10}$$

$$\sigma_{\Delta G_{Avg}} = \frac{1}{n} \times \sqrt{\sigma_{\Delta G_1}^2 + \dots + \sigma_{\Delta G_n}^2} \quad \text{Eq 11}$$

In the limit that all  $\Delta G$  determinations are approximately the same for a given sequence, Eq 11 reduces to the familiar expression:

$$\sigma_{\Delta G_{Avg}} = \frac{\sigma_{\Delta G}}{\sqrt{N}} \quad \text{Eq 12}$$

In Eq 12,  $\sigma_{\Delta G}$  is the error in a given melt, and N is the number of melts. Having determined the  $\Delta G$  values and their respective errors for individual sequences allowed propagation of error into the thermodynamic boxes.

All mathematical operations carried out beyond this point consisted of addition and subtraction of non-correlated terms (*e.g.*  $\Delta G$ s for isolated species), therefore all the remaining error functions (error in  $\Delta G_A$ ,  $\delta$ , etc.) reduced down to:

$$\sigma = \sqrt{x^2 + y^2 \dots n^2} \quad \text{Eq 13}$$

Errors in  $\delta$  values were calculated according to Eq 1a rather than Eq 1b in the paper to avoid overcounting the error.(3)

### References:

- (1) Bevington, P. R. (1969) *Data reduction and error analysis for the physical sciences*, McGraw-Hill, New York,.
- (2) SantaLucia, J., Jr., and Turner, D. H. (1997) Measuring the thermodynamics of RNA secondary structure formation. *Biopolymers* 44, 309-319.
- (3) Moody, E. M., and Bevilacqua, P. C. (2003) Folding of a stable DNA motif involves a highly cooperative network of interactions. *J. Am. Chem. Soc.* 125, 16285-16293.