An Investigation of the Thermodynamics of Two Group I Single Nucleotide Bulges: The Effects of 2 versus 3 Base Pairs Outside the Bulges



Claire Crowther

Senior Comprehensive Project

Allegheny College

Meadville, PA 16335

An Investigation of the Thermodynamics of Two Group I Single Nucleotide Bulges: The Effects of 2 versus 3 Base Pairs Outside the Bulges

Claire V. Crowther

Senior Comprehensive Chemistry Project
Allegheny College
Meadvill, PA 16335

Submitted to the Allegheny College Department of Chemistry in partial fulfillment of the requirements for a degree of Bachelor of Science

Committee:
Dr. Martin Serra
Dr. Elizabeth Guldan
Dr. Mark Ams

April 12, 2013

I hereby recognize and pledge to fulfill my responsibilities, as defined in the Honor Code, and to maintain the integrity of both myself and the College as a whole

.

_____

Claire V. Crowther

Table of Contents

Table of Contents i

Acknowledgments iii

Abstract iv

Introduction 1
- RNA Structure 1
- Primary, Secondary, and Tertiary Structure 3
- Bulges 5
- Secondary Structure of Two Group I Bulges 8
- Thermodynamic Analysis 9
- Current study 13

Materials and Methods 16
- RNA Synthesis 16
- RNA Purification 17
- Quantification of RNA 20
- Preparation of the Duplex and Thermal Denaturation 20
- Thermodynamic Analysis 21

Results 22
- Thermodynamic Parameter 22
- The Effect on the Free Energy of the Duplex by the Bulges 25
- Overall Effect of the Two Group I Bulges 28

Discussion 30
- Duplex Stability 32
- Least Stable Stem Effect on Stability 34
- Predicted Bulge Stability 35
- Current Study 38
- Bulge Confirmation 38
- Apparent Trends 39

Conclusions 41

Further Studies 42

References 44

Appendices

A: *Meltwin 3.0* Equation 1 Derivation I
B: Derivation of T_M Dependence Equation for *Meltwin 3.0* Equation 2 IV
C: Detailed Purification Procedure VI
D: Raw Optical Melt Data XVII
E: Sample Free Energy Calculations XXVIII
F: Statistical Analysis via *GraphPad Prism* XXXI

Acknowledgements

Dr.Serra, thank you for convincing me to be a Chemistry Major, it was one of the best decisions of my life. Your guidance and help in getting into grad will always be something I will be grateful for. I am also appreciative to you for putting up with all my random questions as I was doing my research.

The Chemistry Department, each and every professor here has helped me at some point in learning something new that all culminated in this project being possible. Thank you for being some of the best faculty and always being there to just talk.

Fellow Chemistry Majors: It has been a blast sharing these past four years with you. All of the fun times we have had goofing off when none of us wanted to comp will truly be missed. I think I learned so many things from just not wanting to do my own work, and watching you all do yours. I will never forget the "unplanned departmental all-nighters". I will miss having you all around to help teach me things next year. I wish you all good luck with the next chapters of your lives.

The Lab Group: Abby, Jess, Ryan it has been a blast comping with all of you. Thanks for helping answer all of my random questions when I did not quite know what I was up to. I wish all the best for all of you.

Friends: Thank you for all of your support. This lovely comp would not be possible without your distractions to help me relieve some stress, and your motivation to keep going when I just did not want to care anymore. Caity, thanks for putting up with me these past four years, somehow we managed to survive. Emily, Nicole, Anna, Shelby, Lexie, Tyler, and Neil thanks for letting me rant to you about chemistry even though half the time I am pretty sure you had no idea what I was saying. Tucker, thanks for all the support and telling me that I could do this; you have been one amazing motivator. Your late night texts and listening to me rant about melts is greatly appreciated.

My Family, thank you so much for all of your support these past 21 years. I appreciate you always being therefore for me. None of this would have been possible without you. I cannot tell you how much I appreciate everything!

Abstract

Eleven RNA sequences with two Group I bulges were optically melted in 1 M NaCl to determine their thermodynamic parameters ($\Delta H°$, $\Delta S°$, $\Delta G°_{37}$, and T_m). The number of nucleotides outside the bulges was varied from two to three and the number of nucleotides between the two bulges was varied. The revised Blose model was used to analyze the data, stating that as the decreased stability of the least stable stem, the effect that a bulge has on the duplex is decreased. The data that was collected was combined with previous research by Jones and Mastrogiacomo and then compared to the revised Blose Model (Jones 2007, unpublished; Mastrogiacomo 2011, unpublished; Blose et al., 2007; McCann et al., 2011). The effect of the each of the two single nucleotide bulges on the thermodynamics were determined to be different for each the bulges. The revised Blose model does not account for the two bulges to have different effects on the duplex, because it does not account for the effect the bulges can have on each other and the least stable stem. The predicted value for the effect of the bulge on the free energy of the duplex and the measured value differ by 2-3 kcal/mol, where the predicted gives a less positive free energy. The revised Blose model agrees with the data that as the stability of the least stable stem decreases the effect of the bulge on the free energy of the duplex decreases. None of the trends were statistically significant, which may be the result of a small data set. With further experimentation a significant trend may be determined.

Introduction

Ribonucleic acid (RNA) performs many different functions necessary for life such as: storing information (Bloomfield et al., 2000), binding small molecules and proteins (Bloomfield et al., 2000; Wu and Uhlenbeck, 1987), gene expression and regulation (Bloomfield et al., 2000), chemical catalyst (Bloomfield et al., 2000), and the regulation and coding of proteins (Watson et al., 2007). The understanding of RNA is of great importance because RNA is present in HIV virus, AIDS (Hermann and Patel, 2000) and rhinovirus, which causes the flu (Price, 1956). RNA is also present in the R17 replicase gene used to bind the R17 coat protein, a virus that is present in *Escherichia coli* (Wu and Uhlenbeeck, 1987). There are three main forms of RNA: messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA); which all perform various functions. mRNA is the "genetic code" used for which assist in the process of transcribing DNA into specific protein using ribosomes (Bloomfield et al., 2000). rRNA complexes with ribosome proteins to make ribosomes, which are the organelles where protein synthesis occurs (Bloomfield et al., 2000). tRNA works to synthesize proteins by adapting mRNA sequences into the amino acid sequences in proteins (Bloomfield et al., 2000). Understanding how RNA is made will help develop a better understanding of how it works to help further the understanding of the physical properties of the molecule.

RNA Structure

Nucleotides are the elementary structures that make up RNA. Nucleotides are made up of a phosphate-sugar backbone attached to different nitrogenous bases as seen in Figure 1. In RNA there are four main nitrogenous bases: adenine (A), guanine (G), uracil (U), and cytosine (C). These nitrogenous bases can be separated into two categories based on their structure. Adenine and guanine are purines, which have fused bicyclic rings. Cytosine and uracil are pyrimidines which are single rings. The pyrimidines will pair up with the purines as is shown in Figure 1.

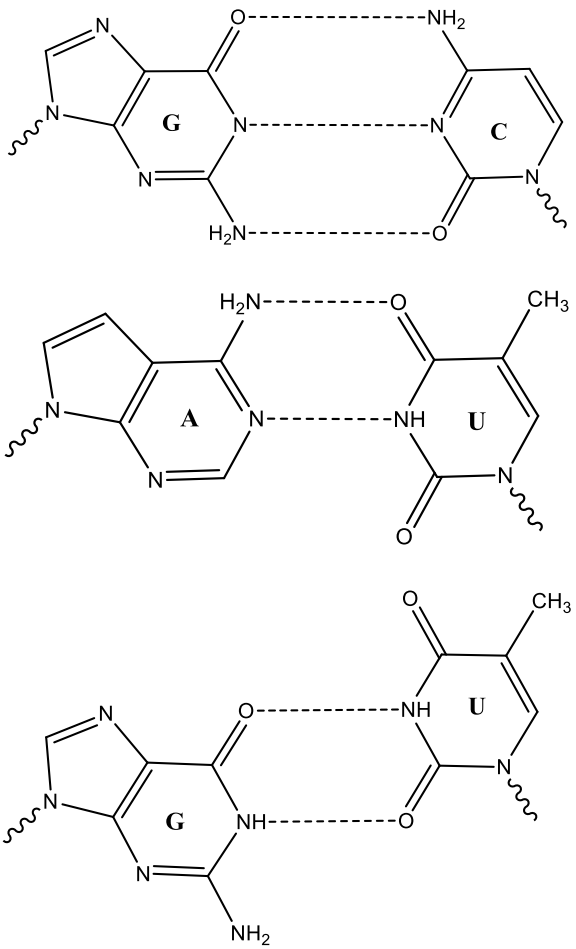


Figure 1. Normal Watson-Crick Base Pairing. G:C forms three hydrogen bonds, A:U and G:U form two hydrogen bonds (Watson and Crick, 1953, Crick, 1966; Bloomfield et al., 2000).

Normal Watson-Crick base pairing occurs in RNA to form a double stranded helix with the ribose sugar backbone; where G and C are paired and A and U are paired. The G-C pair is held together by three hydrogen bonds, while A-U pairs are held together with two hydrogen bonds (Figure 1) (Watson and Crick, 1953 ; Bloomfield et al., 2000). Wobble base pairs can also be formed between G and U (Figure 1) (Crick, 1966).

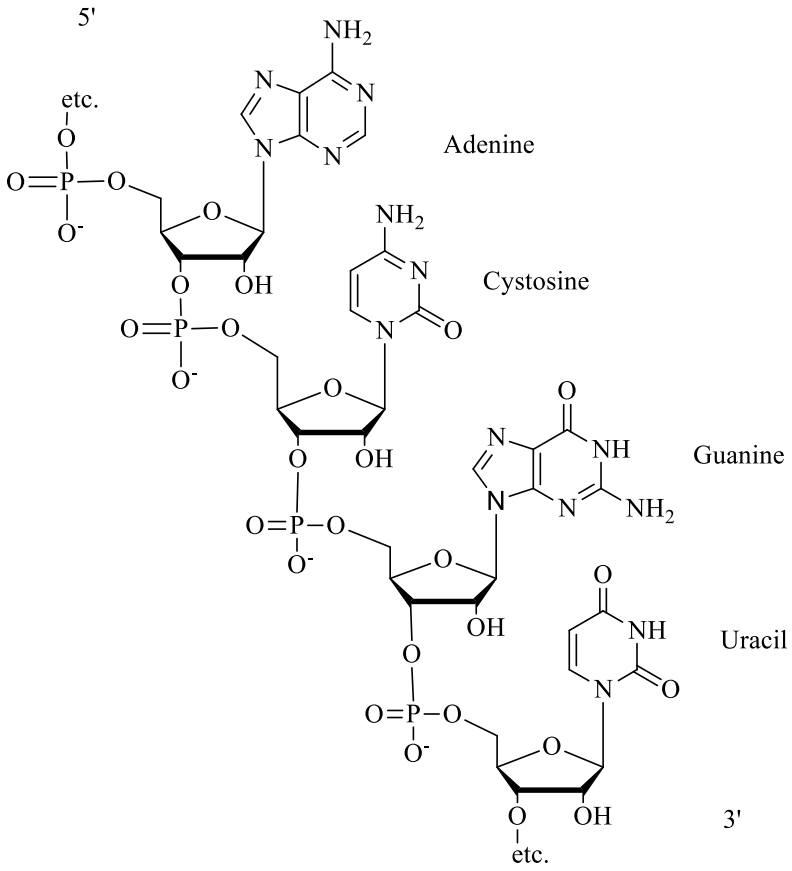


Figure 2. Structure of RNA. The RNA represents all the different nitrogenous bases found in RNA. By convention this strand would be written ACGU, as that is 5' to 3' (adapted from Garrett and Grisham, 2008).

Primary, Secondary, and Tertiary Structure

RNA has three structural levels: primary, secondary, and tertiary. Primary structure is the sequence of the nucleotides which are linked covalently by the phosphodiester bonds where the 5'phosphate group of one nucleotide is linked to the 3'alcohol group of another nucleotide (Figure 2). RNA has directionality in its primary structure; one end has a free phosphate group, this is known as the 5' end, while the other end has a free alcohol group, which is known as the 3' end (Figure 2) (Chen and Varani, 2010). Sequences of RNA are generally written left to right 5' to 3' by convention. For example in Figure 1 the sequence would be written ACGU.

Secondary structure is when strands of RNA come together and form Watson-Crick base pairs to form a double stranded helix. RNA is made as a single strand, which means that it has the ability to form several structural motifs. These motifs include: hairpins, bulged bases, and internal loops (Figure 3) (Bhattacharyya et al., 1990; Chen and Varani, 2010; Hermann and Patel, 2000).

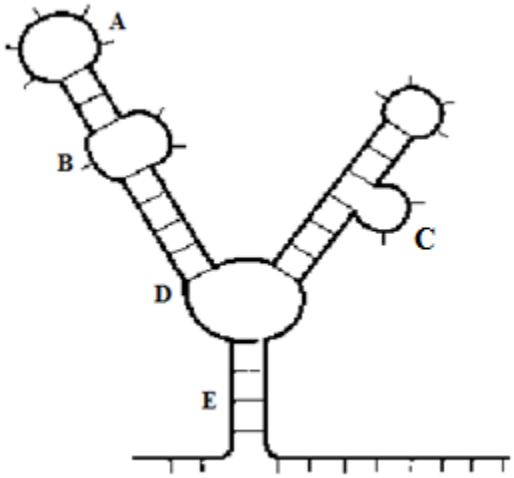


Figure 3. Secondary Motifs of RNA. The motifs represented above that RNA can form include A) hairpin loops B) internal loops C) bulge loops D) multi-branch loops and E) duplexes. Adapted from http://compbio.pbworks.com

The secondary structure motifs have varying effects on the stability of the RNA molecule as a whole, but are generally known to destabilize the RNA molecule (Freier et al., 1986).

Tertiary structure is bending and folding of the secondary structure, which includes base pairing and base stacking (Turner and Sugimoto, 1988). Tertiary structure is what makes RNA a three dimensional structure. An example of the tertiary structure is represented in Figure 4.

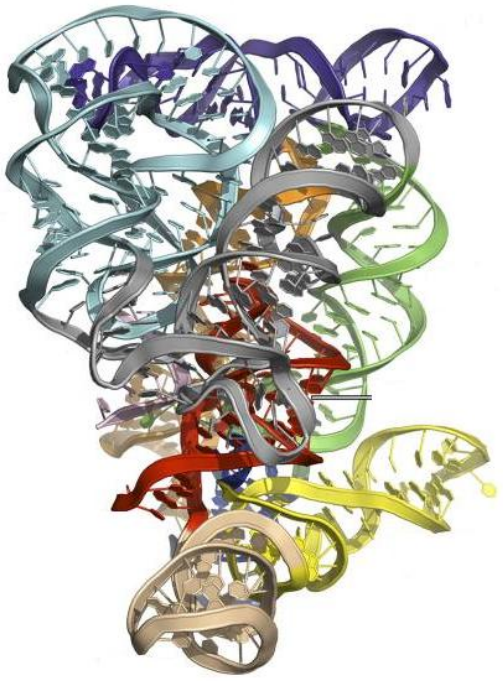


Figure 4. Example of Tertiary Structure of RNA. The different helixes of the RNA can be seen to interact (adapted from Toor et al., 2010).

Tertiary folding is important for many reasons including creating binding site for metals and small molecules (White and Draper, 1989). Tertiary structure is known to not be as stable as secondary structure, but both are important for the RNA to function (Turner et al., 1988; Tinoco and Bustamate, 1999). Because tertiary structure is so complex and involves so many different interactions it is very difficult to predict. In order to understand the folding of RNA the primary structure and secondary structure, of base pairing and base stacking, must first be understood. By understanding how different motifs in the secondary structure affect the stability of RNA, tertiary structure can hopefully be predicted to help further the understanding of how RNA molecules function (Turner et al., 1988).

Bulges

In secondary structure bulges are places where one of the strands of RNA has nucleotides that are unpaired, which can be seen as C in Figure 3. Bulges occur most frequently as a motif in

secondary structure (Xiong et al., 2001). There are several different types of bulges that are classified into four groups. Group I bulges are nucleotides on one strand that are unpaired, but different from their nearest-neighbor nucleotides (Hermann et al., 2000). Group II bulges are when the bulged nucleotide is identical to one of its nearest-neighbor nucleotides creating ambiguity as to the location of the bulge (Hermann et al., 2000). Group III bulges are ambiguous in both the identity of the nucleotide that is bulged location of the bulge (McCann et al., 2011). For Group IV bulges it is a combination of Group II and III bulges (McCann et al., 2011). Examples of the groups are in Table 1.

Bulge Group	Bulge Location 1	Bulge Location 2	Bulge Location 3
I	GC**A**GC CG CG	-	-
II	AU**C**<u>**C**</u>AU UAG UA	AU<u>**C**</u>**C**AU UA GUA	-
III	C<u>**G**</u>**A**CG G UGC	C**G**<u>**A**</u>CG GU GC	-
IV	C**A**<u>**A**</u>**G**G GUU A	C**A**<u>**A**</u>**G**G GU UA	C<u>**A**</u>**AG**G G UUA

Table 1. Example of group I and II bulges. The bold nucleotide(s) are the possibly bulged nucleotide/bulged nucleotide. The underlined nucleotide is what would actually be bulged. All of the different possible bulges are represented for each sequence (McCann, 2011).

When determining the effect of a bulge the nearest neighbor model is one method that can be used. In the nearest neighbor model, the nucleotides that are right beside the bulge are the only ones that affect the bulge. For example in Table 1 for the Group I bulge the nearest neighbors are the CG base pair to the left of the bulge, and the GC pair to the right of the bulge. While non-nearest neighbor refers to the nucleotide not next to a nucleotide affecting the structure. For example in Table 1 for the Group I bulge, the non-nearest neighbors would be the GC pair listed first and the CG pair that is listed last.

For RNA the identities of the nearest neighbor nucleotides do not affect stability of the bulge, but rather is affected by the least stable stem (McCann et al., 2011). In nature the most commonly bulged nucleotide is adenosine (Gutell et al., 2000). The different group bulges are thought to have different effects on the stability of the RNA molecule, so investigating how they affect the thermodynamic stability will reveal more about the secondary and tertiary structure of the molecule. It has been seen that there is a correlation between the destabilization of the duplex and the least stable stem for a single nucleotide bulge (McCann et al., 2011).

Bulges are most commonly seen in RNA structure for two functions. Bulges can act as hinge sites for the RNA backbone, allowing it to kink and twist. This happens because the bulged nucleotide is not constrained by a pairing with another nucleotide. Another function of bulge loops is that they can act as wedges, where the bulged nucleotide can point into a deep grove and interact with base pairs that are otherwise inaccessible (Hermann et al., 2000; Xiong et al., 2000).

Bulges can fall into different families: extrahelical, intercalated, and side-by-side (Figure 5).

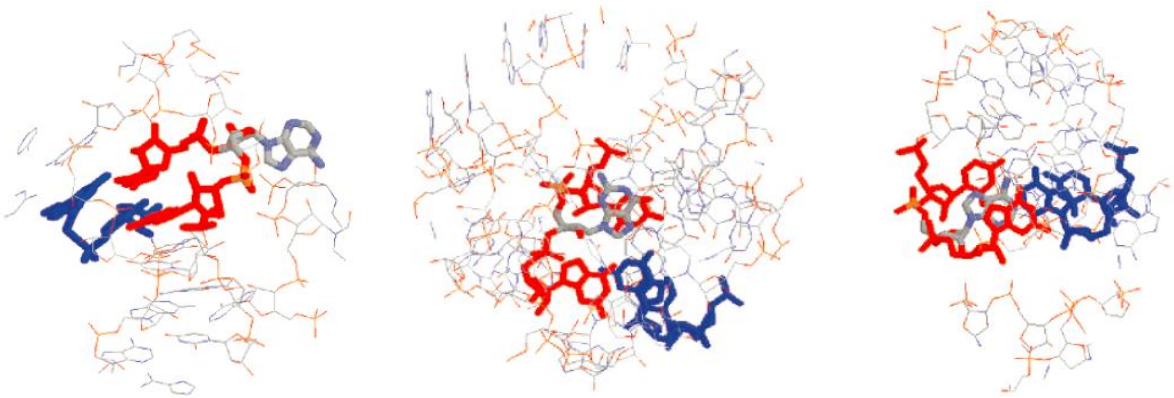


Figure 5. Three Families of Bulge Loops. The bulges as they are observed in the large subunit rRNA. The different bulge families are shown in CPK colors where the neighboring base pairs are in red and blue (Znosko et al., 2002).

Bulges that are directed away from the helix are considered extrahelical. Intercalated bulges are positioned between adjacent base pairs. Side-by-side bulge loops are coplanar to either the base that is to the 5’ or 3’ side of the loop (Hermann et al., 2000). Side-by-side bulges can form two different families. Either the bulged nucleotide is side-by-side with the nucleotide to the 5’ side or the nucleotide to the 3’ side. The side-by-side bulge has been seen to increase the twist of the helix, causing the stacking interactions of the side of the duplex with the bulge to be broken (Znosko et al., 2002). Nucleic acids duplexes are normally stacked end-to-end in a crystal, so there is continuous stacking of the neighboring regions or have bulges that are stick out from the helix (Hermann et al., 2000). This can stabilize the helix through lateral intermolecular interactions (Hermann et al., 2000). About two-thirds of bulges are extrahelical, half of all bulges interact with neighboring nucleotide either by base pair formation, π stacking, or hydrogen bonding (Shen et al., 1995). The bulges are able to interact locally with the strand they are on as well as with other helixes through tertiary structure. Reasons the bulged nucleotide would not interact with its surrounding would be because it is in a pocket, so there would be no interactions. Intercalated bulges are not seen very often, but they bend the helix (Znosko et al., 2002). Intercalate bulges maybe a way that the RNA forms a recognition site to help the RNA bind other molecules (Hermann et al., 2000).

Secondary Structure of Two Group I Bulges

RNA structures can have more than one bulge located within a duplex. When working with RNA that has two Group I bulges, they can be arranged three different ways in the molecule (Figure 6).

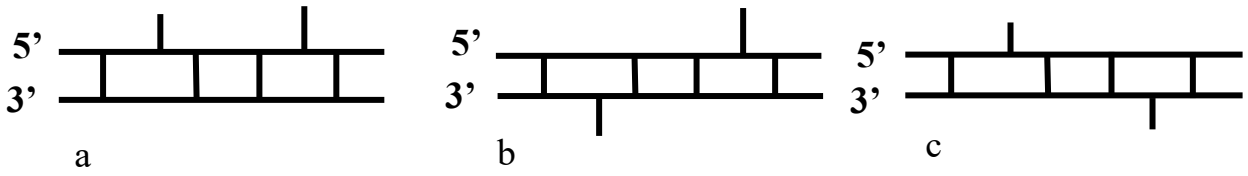


Figure 6. Different possible arrangements of two group I bulges. The different bulges has a) both bulges on one strand, which can be either the 5’ or 3’ strand (1 base pair outside the bulges and 2 base pairs between the bulges) b) the bulges with the first one on the 3’ strand (1 base pair outside the bulges and 2 base pairs between the bulges) and c) the first bulge on the 5’ strand(1 base pair outside the bulges and 2 base pairs between the bulges).

For Figure 6 all the presented examples have 1 base pair outside the bulges and 2 base pairs between the bulges. The different placements of the Group I nucleotide bulges may have different effects on the stability of RNA. The bulges can cause many different changes in structure such as: folding, base stacking, and base pairings (Blose et al., 2007; Turner et al., 1988; Wu and Uhlenbeck, 1987).

Thermodynamic Analysis

Tertiary structure is what determines the function of a given sequence of RNA. Using information of a RNA molecules’ sequence and the secondary structure, the tertiary structure can be predicted. To determine the thermodynamic properties of RNA, optical melting, also known as ultra-violet (UV) melting is utilized. The absorbance of the RNA is taken for several concentrations of the RNA, having the temperature increase slowly to cause the RNA transition from double to single stranded (Blose et al., 2007; Freier et al., 1986; Longfellow et al., 1990; Lu et al., 2006; McCann et al., 2011; Peritz et al., 1991; Sugimoto et al., 1986; Turner et al., 1992; Xia et al., 1998; Znosko et al., 2002). Optical melting is used in this experiment to determine the effect that the two Group I bulges have on the stability of the RNA sequence.

Optical melting is a method to determine when a RNA molecule goes from double stranded to single stranded (Figure 7).

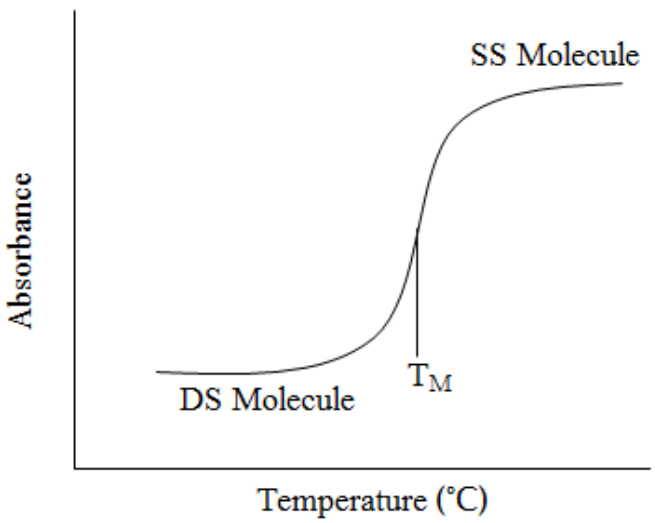


Figure 7. Example Optical Melting Curve. At low temperatures RNA is in a double stranded (DS) form and at high temperatures the molecule is single stranded (SS). The graph is of temperature versus absorbance.

When the temperature rises the hydrogen bonds between the base pairs weaken, which cause the RNA molecule to denature into two separate single stranded molecules. Single stranded RNA molecules are known to absorb more light at high temperatures, as the molecules are no longer bonded together (Turner et al., 1988). The absorbance of RNA is generally monitored at either 260 or 280 nm (Xia et al., 1998; Znosko et al., 2002).

Thermodynamic parameters (ΔS°, ΔH°, ΔG^o_{37}, and T_M) can be determined from optical melt experiments, using a two state model that assumes that the RNA molecule is either double stranded or single stranded with no intermediates (Turner et al., 1988; Schroeder and Turner, 2009). The optical melts are done over a temperature range of 5°C to 85°C. There are two methods that are used to analyze the melt curves. One method uses a non-linear least squares regression to fit the melt curve using the *MeltWin3.0* program, which uses the two-state model to fit self-complimentary and non-self-complimentary RNA strands. Using Equation 1, a derivation

of the Beer Lambert Law, the thermodynamic parameters are found (Turner et al., 1988; SantaLucia and Turner, 1998; Schroeder and Turner, 2009).

$$A_T = \left(\frac{\{\left[e^{(T\Delta S°-\Delta H°)/RT)2} + 2\left(e^{(T\Delta S°-\Delta H°)/RT)}C_T\right] - e^{(T\Delta S°-\Delta H°)/RT)}\}}{2}\right) *$$

$$(2M_{SS} + 2b_{SS} - M_{DS}T - b_{DS}) + (\frac{C_T}{2})(M_{DS}T + b_{DS}))$$

(Equation 1)

*derivation of Equation 1 in Appendix A

Where:

A_T = total absorbance of sample
M_{SS} = slope of linear baseline of random coil form of single-stranded RNA
M_{DS} = slope of linear baseline of RNA-RNA duplex
b_{SS} = the y-intercept of the linear plot of the random coil form of single-stranded RNA
b_{DS} = the y-intercept of the linear plot for the RNA-RNA duplex
C_T = total concentration of the single-stranded RNA
T = temperature in Kelvin
R = 10986 cal/mol*K

The other method to determine the thermodynamic parameters is by graphing the inverse of the melting temperature (T_M^{-1}) versus the natural log of the total concentration(C_T) of the single-stranded RNA. The relationship between concentration, temperature, and the thermodynamic properties can be seen in Equation 2.

$$T_M{}^{-1} = \left(\frac{R}{\Delta H°}\right) * \ln(C_T) + (\frac{\Delta S°}{\Delta H°}) \quad \text{(Equation 2)}$$

*derivation of Equation 2 in Appendix B

If the strand is non-self-complimentary C_T is replaced by $\frac{C_T}{4}$ (Turner et al., 1988; SantaLucia and Turner, 1998; Schroeder and Turner, 2009). The slope of the lease squares regression line is $\frac{R}{\Delta H°}$ and the y-intercept is $\frac{\Delta S°}{\Delta H°}$. Using Equation 3 the ΔG_{37}^o can be determined.

$$\Delta G_{37}^o = \Delta H° - T\Delta S° \quad \text{(Equation 3)}$$

These two methods of determination of the thermodynamic parameters must agree within 15% for the data to be consistent with the two state model (Freier et al,. 1986). The free energy, ΔG^o_{37}, is important in understanding the structure of RNA, because if the structure has a greater negative free energy the molecule is more stable. This means that the structural motif is stabilizing the duplex.. Whereas if ΔG^o_{37} is positive, it means that the structural motif is causing a destabilization in the duplex. Depending on the value of free energy it can be determined how likely a duplex is to form and when the RNA will denature. Higher melting temperatures will be seen from sequences that are more stable.

The nearest-neighbor model is used to determine the effect of the single nucleotide bulge, as a bulge does not affect the interactions of the neighboring base pairs (Longfellow et al., 1990; Znosko et al., 2002). To determine the effect of a bulge the free energy of the bulge is determined at 37°C, ΔG^o_{37}, using Equation 4 (Znosko et al., 2002).

$$\Delta G^o_{37,bulge} = \Delta G^o_{37,measured} - \Delta G^o_{37duplex} \quad \text{(Equation 4)}$$

$\Delta G^o_{37,measured}$ is the measured free energy of the RNA sequence, which is found experimentally. $\Delta G^o_{37duplex}$ is the free energy of the parent strand, which is the sequence excluding the bulged nucleotide(s). The $\Delta G^o_{37duplex}$ is calculated using the nearest neighbor model for the RNA duplexes, which accounts for the nucleotides interactions and the effect of initiating and terminating the sequence. Using Equation 4, the effect of bulge can be seen by removing the effect of the parent strand. This equation can be used for a single nucleotide bulge or having several single nucleotide bulges.

Current study

The current study is looking at the effect of two group I bulges in the different conformations in Figure 6. The following sequences are being analyzed for their thermodynamic properties (Table 2).

BK 9 BK10	GC ACU**A**CG CG**A**UGA GC
BN 1 BO 10	GC**A**UGA GC CG ACU**A**CG
BN 2 BL 5	GUC ACU**A**CUG CAG**A**UGA GAC
BI 42	GC GGCC**A**GC CG**A**CCGG CG

BL 7 BL 8/ BN 5	GUC**A**UCA CUG CAG AGU**A**GAC
BL 23	GC**A**UGCGCA GC CG ACGCGU**A**CG
BL 3 BL 17	GC**A**UGCGU**A**CG CG ACGCA GC
BN 3 BO 11	CG GAUGCA GC GC**A**CUACGU**A**CG
BM 1 BM 2	CAG G**A**CUG GUC**A**C GAC
BL 11 BN 4	GC**A**UGCGA GC CG ACGCU**A**CG

Table 2. Sequences used in this study. The RNA sequences that are non-self-complimentary are made with two strands the normal and prime strand. The top strand is 5’ to 3’, while the bottom strand is 3’ to 5’.The sequences without a second sequence are self-complimentary. The bulged nucleotide is bolded and underlined.

Optical melts of all these sequences will be performed to determine the structural effects of the location of the bulged nucleotide. The bulged nucleotide is adenine as it is most prevalent in nature, and the identity of the bulge has not been seen to change the effect on the duplex (McCann et al., 2011; Gutell et al., 2000). Group I bulges are being studied as they are known to destabilize the RNA molecule, and have essentially the same effect as the Group II bulges (McCann et al., 2011). The optical melts of the sequences in Table 2 are to be added to previous experimentation (Table 3).

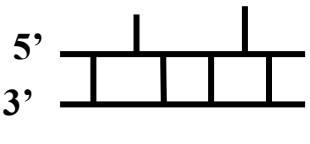
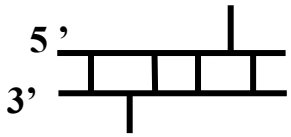
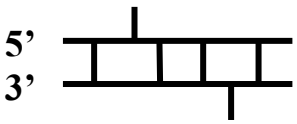
Base Pairs Outside Bulges	**Base Pairs Between Bulges**			
		ΔG^o_{37} (kcal/mol)	ΔG^o_{37} (kcal/mol)	ΔG^o_{37} (kcal/mol)
1	4	2.1, 2.0	2.6, 3.7, 2.6, 3.8	-
1	5	-	3.5	1.5,1.6, 2.2
1	6	-	1.2, 7.6, 2.5,1.9, 1.5, 3.1	1.1, 1.8, 5.7
2	2	-	4.0, 3.9, 3.9	4.0
2	**3**	7.0	*	*
2	**4**	6.3, 7.1	6.7	6.4
2	**5**	*	8.8	*
2	**6**	*	8.5	*
3	**1**	6.8, 6.3	*	*
3	**2**	9.1, 7.8	4.7, 8.3	2.3, 8.0, 11.0
3	**3**	6.9	*	*
3	4	9.5, 10.9, 10.9	8.1	9.1
3	5	-	10.9	-
3	6	-	11.0, 11.4	-
4	3	8.6	-	-
4	4	-	-	11.3

Table 3. Previous work done with these bulge motifs. ΔG^o_{37}, the free energy at 37°C, previously determined (Jones 2007, Morelli 2009, Mastrogiacomo 2011, Unpublished). Locations designated with a * are samples that have been made to complete this study.

The samples selected to optically melt in Table 3 were picked to study the effect of having two or three base pairs outside the bulges. From the data that has been collected currently it seems that as the number of base pairs outside the bulged nucleotide increases, the bulge

destabilizes the duplex more. It is expected that the data collected in this study will continue to follow this trend.

Research by Blose suggests that for single Group I nucleotides that are bulged that non-nearest neighbors affect the stability of the stem of the helix (Blose et al., 2007). As the stems adjacent to the bulge become more stable, the destabilizing of effect of the bulge increases. The destabilizing effect of one bulge is not dependent on the bulge's nearest neighbors as long as the bulge is surrounded by Watson-Crick base pairs (Blose et al., 2007). The current study is varying the number of nucleotide both outside the bulges and between the bulges to determine the effects of the two bulges. The aim is to develop the current models so they are able to incorporate for multiple bulges in sequences. By furthering the collected experimental data of the thermodynamic parameters of RNA, a better understanding of secondary RNA structure can be determined, and therefore a better understanding of RNA structure and folding.

Materials and Methods

*more detailed procedure included in Appendix C

RNA Synthesis:

Some of the sequences used in this study were purchased from Integrated DNA Technologies (Coralville, Iowa). These sequences were dissolved in water and then checked for purity. Based on their purity, either further purification was necessary, or the sequence could be used for melting immediately.

All of the other RNA oligomers were synthesized using the phosphoramide approach at the University of Rochester on an automated RNA synthesizer (Usman et al., 1987). The nucleotides were added in the desired sequential order to a controlled pore glass support (CPG). To prevent side-reactions from happening during the synthesis protecting groups were added. To the 2'oxygen of the ribose sugar a terbutylmethylsiyl group was attached and to the amine end of each base a benzoyl group was added (Figure 8).

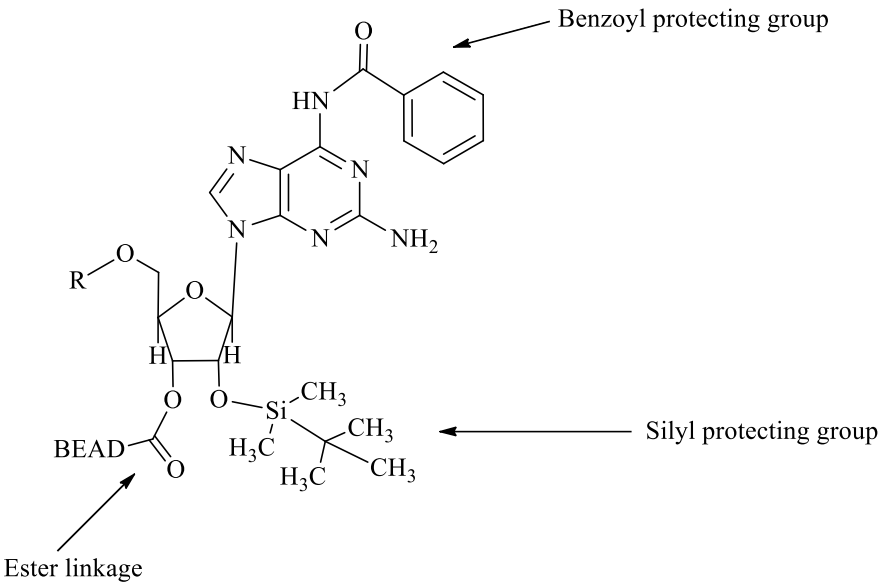


Figure 8. Protecting Groups on the Ribonucleic Acids from Synthesis. Example of an adenine nucleotide attached to a glass bead (CPG) by an ester linkage, with the silyl and benzoyl protecting groups. These groups are used to prevent the adenosine from reacting during synthesis.

RNA Purification

To purify the RNA, the ester linkage between the RNA ogligomer and the CPG had to be cleaved. This was done using a 3:1 solution of cold, fresh, concentrated ammonium hydroxide and 100% ethanol. The ammonium removes the benzoyl protecting group and the ethanol dissolves the non-polar protecting groups (Figure 10).

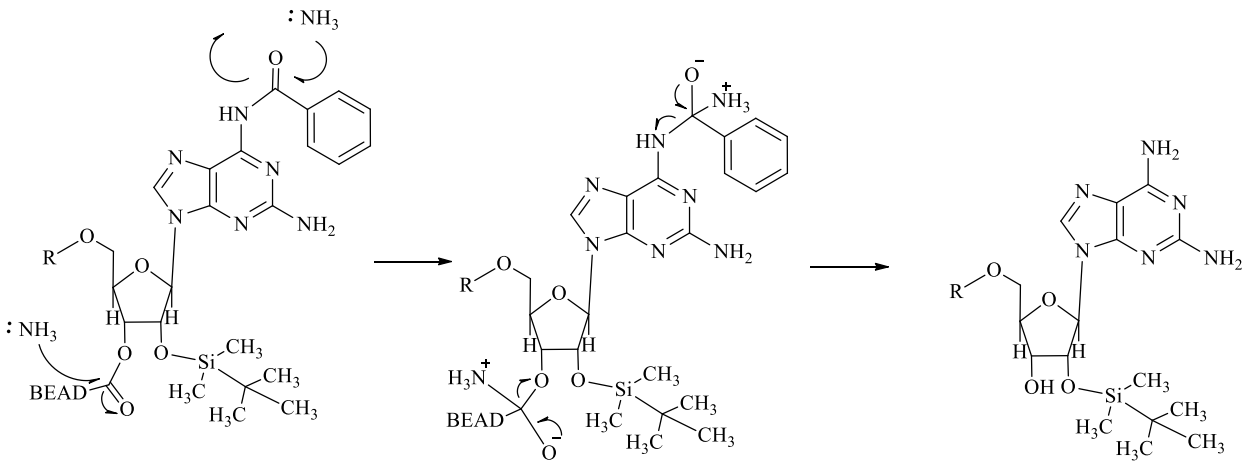


Figure 10. Mechanism for the Removal of the Benzoyl Protecting Group from the RNA oligomers. The ammonium hydroxide cleaves the bond between the benzoyl group and the amide of the nucleotide and the ester linkage between the adenosine and the glass bead.

The solution was placed in a water bath for 18 hours at 55°C, and then cooled and filtered through a disposable chromatography column to separate the glass beads from the RNA. The column was washed with solutions of decreasing polarities to ensure the RNA passed through the column. The vials used for incubation were rinsed with deionized water, a 50% acetonitrile and water solution, and then a 100% acetonitrile solution. All of these rinses were added to the column. The oligomers may have different polarities if some of the protecting groups were not removed, which is why all the washes are necessary. The collected filtrate was dried for an hour using a Savant Speed-Vac concentrator centrifuge. A water aspirator was attached to the Speed-Vac to remove the corrosive ammonium hydroxide gas. The sample was then separated into three microcentrifuge tubes; ensuring that they all had equal volumes. The tubes were dried overnight in a Speed-Vac attached to a vacuum pump.

To each dried sample a 1:9 solution of dimethylformamid (DMF) and triethylamine hydrogen fluoride (TEAHF) was added, and was then placed in a 55°C water bath for two hours. The TEAHF during the incubation serves to cleave the silyl protecting group, and DMF serves as

the solvent. The vials were then removed from the water bath and 1.2 mL of n-butanol was added to each vial, which was then placed in the freezer (-20°C) for at least an hour. The vials were then centrifuged at 12K revolutions per minute for 20 minutes. This allowed the RNA to precipitate. Once a precipitate was observed in the vials the supernatant was decante and the precipitate air dried.

To remove the inorganic salts from the RNA sample a Waters Sep Pak C-18 chromatography column was used. The precipitate, the RNA oligomers, were dissolved in 8 mL of 2X Sep Pak buffer (5 mM ammonium bicarbonate, pH 4.5) and then added to the column. As the buffer is polar it allowed the RNA to bind to the column. The absorbance of different washes at 280 nm using a UV-visible spectrophotometer was used to check that the RNA was bound to the column. To flush the RNA off the column a rinse of 30% acetonitrile and water by volume was used which was collected and then dried using the Speed-Vac attached to the vacuum pump. The sample was resuspended in 150 µL of deionized water. Thin Layer Chromatography (TLC) was used to separate the desired RNA from sequence failures. The sample was spread across a preparative Baker Si500 TLC Plate-Silica Gel (20X20 cm, 500 micrometer thickness). The plate was run in chamber using a mobile phase of 55:35:10 n-propanol: ammonium hydroxide; water with a piece of filter paper to increase the gas-liquid equilibrium. The plate was dried and visualized with a UV lamp. The desired band is the darkest band lowest down on the plate. The larger the RNA oligomers, the slower they travel up the plate. The bands further up are shorter oligomers which indicate failures in the synthesis or a cleavage reaction, so they are ignored. The desired band of RNA was scrapped off and suspended in 2 mL of deionized water. To this sample 2 mL of 20X Sep Pac buffer was added. The sample was then run through another

Waters Sep Pack C-18 chromatography column to remove any remaining inorganic salts. The final sample was dried and then resuspended in 125 µL of deionized water.

RNA Purity Check

An analytical TLC was used to check the purity of all the samples. On a Baker Si500 TLC plate (3x10cm) a single spot of RNA was spotted. The plate was run using 55:35:10 n-propanol: ammonium hydroxide; water as the mobile phase. The appearance of single spot when visualized with a UV light, "proved" the samples to be pure.

Quantification of RNA

The concentration of the pure RNA was determined using Beer's Law: $A = \varepsilon l C$. The sample was diluted by a factor of 100. At 80°C the absorbance (A) of the sample was taken at 260 nm and 280 nm was taken at 80°C. For each sequence at 280nm at 260nm the molar absorptivity (ε) using the *RNACalc version 1.1* program. The program uses the nearest neighbor model to calculate the molar absorptivity. The pathlength (l) was 1 cm. Using this information and the dilution factor the concentration of RNA was determined.

Preparation of the Duplex and Thermal Denaturation

For non-self complementary RNA duplexes equal molar amounts of the two complementary RNA strands were combined for melting. The sample was then dried and resuspended in melt buffer (1M NaCl, 0.01 M cacodylic acid, and 0.001M ethylenediaminetetracetic acid).

The duplexes were melted at ten different concentrations. The optical melts were done on a Beckman DU 800 Spectrophotomer with High Performance Temperature Controller. The absorbances were measured at a rate of 1°C/min ranging the temperature from 5°C to 90°C. Measurements of absorbance were taken for each cell every half degree. The absorbances were read at either 260nm or 280 nm. This is the point where the transition from single to double stranded RNA is seen. 280 nm gives better data, but requires twice the RNA, which is not always possible.

Thermodynamic Analysis

The data was fit two ways using *Meltwin 3.0* as described in the introduction. From the *Meltwin 3.0* least squares regression program and the plot of temperature versus concentration the average entropy, enthalpy, melting point, and free energy are determined.

Results

The current study investigated two Group I adenosine bulges with varying numbers of nucleotides outside the bulges, and between the bulges. The positions of the two bulges was also varied as exhibited in Figure 6. The sequences were made up of only Watson-Crick base pairs for the stems. The purpose was to determine what parameter has the greatest effect on the stability of the duplex. Previously it has been seen that the nearest neighbors to a bulge have no effect, but this was done on single bulges, with the nearest neighbors being Watson-Crick base pairs (Blose et al., 2007).

Thermodynamic Parameters

The revised Blose model, Equation 5, is used to calculate the Gibb's free energy contribution of one Group I bulge (Blose et al, 2007; McCann et al., 2011).

$$\Delta G^0_{37(bulge\ loop)} = \left(-0.53 * \Delta G^0_{37(less\ stable\ stem)}\right) + 0.7\ kcal.mol \qquad \text{(Equation 5)}$$

Using Equation 5 the effect of a single nucleotide bulge can be calculated. Equation 5 can also be used to find the effect of one bulge when there are two bulges in a duplex. A sample calculation of this can be found in Appendix E. When working with multiple single nucleotide bulges it is difficult to calculate the free energy for only one of the bulges. This comes from the fact that the least stable stem affects how a bulge loop changes the stability of the duplex (McCann et al., 2011). This raises problems as each bulge may have a different least stable stem. All off the calculations are thus based off one definition of what the least stable stem. Three values were determined: $\Delta G^o_{37,\ 2nd\ Bulge}$, $\Delta G^o_{37,\ Predicted\ Bulge}$, and $\Delta G^o_{37,\ Average\ Bulge}$. Example calculations can be seen in Appendix E.

When calculating the effect of the second bulge a least stable stem is determined. For example using the sequence, $^{5'}GC\ \ ACUACG^{3'}$ $^{3'}CGAUGA\ \ GC^{5'}$, For the different stems: $\frac{CG}{GC}$, in green has a free energy of -2.36 kcal/mol; the stem $\frac{ACU}{UGA}$, in purple has a free energy of -4.32 kcal/mol; and $\frac{GC}{CG}$, in blue has a free energy of -3.42 kcal/mol. All of these values were obtained using the nearest neighbor model and omitting the penalty of the initiation, as initiation will start with the more stable of the two end stems, so it will not affect the free energy of least stable stem.. From this it can be seen that the $\frac{CG}{GC}$ stem in green is the least stable as it has the least negative free energy, so when doing the calculations the bulge in red was considered the second bulge. Using Equation 5, the effect of the bulge can be found (Appendix E). When predicting the effect of the bulges Equation 5 is again used, but this time the two bulges are calculated as though they have the same effect (Appendix E). The least stable stem in determined in the same manner as previously described. With all of these calculations if the least stable stem were defined differently altered results would be obtained. The calculations were done in this manner so that the data could be compared with previous data.

When working with two single bulged nucleotides it has previously been observed that multiply bulges do not have the same effect on the duplex (Jones 2007, Morelli 2009, Mastrogiacomo 2011, unpublished). However, there has not been enough data collected to be certain of this trend. To continue collecting data to explain these results, eleven sequences were studied, with only seven of them producing useful results. Each of these sequences was studied using optical melting. Using optical melting the thermodynamic parameters of duplex formation of sequences that contain multiple single adenosine bulge loops. Table 4 lists the ΔH°, ΔS°,

ΔG^{o}_{37}, and T_m results which were calculated by *Meltwin3.0* by the least-square regression program of the melt curve and the T_m dependence plot.

Base Pairs Outside Bulges	Base Pairs Between Bulges	Duplex	Analysis of Melt Curve Fits				Analysis of T_m Dependence			
			$-\Delta G^{\circ}_{37}$ (kcal/ mol)	$-\Delta H^{\circ}$ (kcal/ mol)	$-\Delta S^{\circ}$ (cal/K *mol)	T_m (°C)	$-\Delta G^{\circ}_{37}$ (kcal/ mol)	$-\Delta H^{\circ}$ (kcal/ mol)	$-\Delta S^{\circ}$ (cal/K *mol)	T_m (°C)
2	3	GC ACU**A**CG CG**A**UGA GC	5.2	50	146	29	5.0	52	153	28.4
2	3	GC**A**UGA GC CG ACU**A**CG	5.9	53	153	34	6.0	47	133	33.8
2	4	GC GGCC**A**GC CG**A**CCGG CG	9.8	50	131	65	10.2	59	158	62.5
2	5	GC**A**UGCGU**A**CG CG ACGCA GC	8.1	34	83.29	52	8.0	42	109	48.5
2	5	GC**A**UGCGA GC CG ACGCU**A**CG	Needs to be Melted Again							
2	6	GC**A**UGCGCA GC CG ACGCGU**A**CG	9.8	54	141	63	9.8	52	136	63.4
2	6	CG GAUGCA GC GC**A**CUACGU**A**CG	10	48	122	62	11	74	201	58.8
3	1	CAG G**A**CUG GUC**A**C GAC	7.0	22	48.3	44	7.0	26	61.8	42.1
3	3	GUC ACU**A**CUG CAG**A**UGA GAC	6.6	41	111	38	6.6	39	104	37.8
3	3	GUC**A**UCA CUG CAG AGU**A**GAC	6.7	56	160	38	6.8	49	137	38.7
3	3	GUC**A**UCA CUG CAG AGU**A**GAC	6.7	57	163	38	6.9	44	119	39.3

Table 4. Thermodynamic Data for the Thermal Denaturation[a] of two Group I Bulge Loops. The data for all of the sequences that were melted to further understand the effect of two Group I bulges.

[a]All measurements were taken in 1M NaCl, 0.01M cacodylic acid, and 0.001M Na_2EDTA, pH 7.0.

*Sequences with one asterisk are duplexes that are not two state.

The melts were done at either 260 nm or 280 nm, based on how much RNA was available. The analyses of the melts revealed that some of the sequences have enthalpies that do

not agree within fifteen percent, which is the limit for the two state model (Freier et al., 1984). One duplex was determined to not be two states, because the ΔH° values with not within fifteen percent, and the melts were well defined, so the sequence cannot be represented by a two state model. The duplex is $\begin{matrix} CG\ \ GAUGCA\ \ GC \\ GCACUACGUACG \end{matrix}$. Two more sequences were not within fifteen present although because of difficulties with the melts and data analyzation further experimentation may achieve better results. The sequences are: $\begin{matrix} GUCAUCA\ \ CUG \\ CAG\ \ AGUAGAC \end{matrix}$ and $\begin{matrix} GCAUGCGUACG \\ CG\ \ ACGCA\ \ GC \end{matrix}$. The data for these three sequences listed above were omitted from the rest of the data analysis.

The Effect on the Free Energy of the Duplex by the Bulges

The effect of the bulges on the stability of the duplex was found using the nearest neighbor model. This was done using Equation 4. An example calculation is as follows:

$\begin{matrix} GUCAUCA\ \ CUG \\ CAG\ \ AGUAGAC \end{matrix}$ $\Delta G^o_{37,measured}$= -6.80 kcal/mol

$\begin{matrix} GUCUCACUG \\ CAGAGUGAC \end{matrix}$ $\Delta G^o_{37duplex}$ = -13.5kcal/mol

$\Delta G^o_{37,bulges} = \Delta G^o_{37,measured} - \Delta G^o_{37duplex}$ = -6.80 kcal/mol – (-13.5kcal/mol) = 6.7 kcal/mol

In this case the $\Delta G^o_{37,bulge}$, is representative of the effect that both bulges have on the duplex. By dividing this value by two the $\Delta G^o_{37,\ Average\ Bulge}$ can be determined (Appendix E). The two bulges have been previously shown to not have the same effect on the sequence (Jones 2007, M Mastrogiacomo 2011, Unpublished). The following data was collected for the effect of the bulge, Table 5.

Duplex	$-\Delta G^{\circ}_{37}$ (kcal/ mol)
GC ACU**A**CG CG**A**UGA GC	5.4
GC**A**UGA GC CG ACU**A**CG	4.4
GUC ACU**A**CUG CAG**A**UGA GAC	6.8
GC GGCC**A**GC CG**A**CCGG CG	6.0
GUC**A**UCA CUG CAG AGU**A**GAC	6.7
GC**A**UGCGCA GC CG ACGCGU**A**CG	10
CAG G**A**CUG GUC**A**C GAC	4.0

Table 5. Experimentally Determined Values for the Both Group I Nucleotides[a].
[a]The $\Delta G^{o}_{37,bulges}$ were calculated using: $\Delta G^{o}_{37,bulges} = \Delta G^{o}_{37,measured} - \Delta G^{o}_{37duplex}$

From Table 5 it can be seen that all of the bulges are destabilizing to the overall duplex.

As previously described there are many different way to calculate the effect of the bulge. The free energies that were calculated for the second bulge are given in Table 6.

Duplex	Least Stable Stem	$\Delta G°_{37,\ 2nd\ Bulge}$ (kcal/ mol)
GC ACU**A**CG	CG	3.5
CG**A**UGA GC	GC	
GC**A**UGA GC	GC	1.9
CG ACU**A**CG	CG	
GUC ACU**A**CUG	CUG	3.9
CAG**A**UGA GAC	GAC	
GC GGCC**A**GC	GC	3.5
CG**A**CCGG CG	CG	
GUC**A**UCA CUG	CUG	3.8
CAG AGU**A**GAC	GAC	
GC**A**UGCGCA GC	GC	7.6
CG ACGCGU**A**CG	CG	
CAG G**A**CUG	G	3.3
GUC**A**C GAC	C	

Table 6. Free Energy Contribution of the 2nd Bulge. The values are calculated separate of the other bulged nucleotides in the sequence.

For Table 6, using the current model, the method to predict the free energy of the bulge is only depend on the least stable stem that is adjacent to the bulge. Comparing the values in Table 5 and Table 6 the bulges are seen to have different effects on the duplex. This is because in all cases the values for Table 6 are not half of those in Table 5.

To further confirm that the bulges do not have the identical effect a comparison between the predicted free energy of the duplex with the experimentally determined free energy of the duplex. The differences can be seen in Table 7.

Duplex	Predicted -ΔG°37 (kcal/mol)	Measured -ΔG°_{37} (kcal/ mol)	Difference (kcal/mol)
GC ACU**A**CG CG**A**UGA GC	6.6	5.0	1.5
GC**A**UGA GC CG ACU**A**CG	5.5	6.0	-0.6
GUC ACU**A**CUG CAG**A**UGA GAC	7.6	6.6	1.0
GC GGCC**A**GC CG**A**CCGG CG	11	10	1.0
GUC**A**UCA CUG CAG AGU**A**GAC	7.6	6.8	0.8
GC**A**UGCGCA GC CG ACGCGU**A**CG	14.9	9.8	5.2
CAG G**A**CUG GUC**A**C GAC	9.6	7.0	2.6

Table 7. The Measured Free Energy versus the Predicted Free Energy The effect of the bulges from the predicted model where both bulges have the same effect versus the experimentally obtain data for each sequence.

The predicted values and the experimentally obtained values do no show agreement. Good agreement is desired because the data is used to develop a model of RNA structure, which is desired to have high accuracy. This suggests that the two bulges do not have the same effect on stability of the duplex. The prediction both over and under predicts the free energy. Suggesting that other factors will need to be considered when developing the revised Blose model further to account for multiple bulge loops not having the same effect.

Overall Effect of the Two Group I Bulges

To determine a trend in the effect based on the location, this research was combined with other work (Jones 2007, Morelli 2009, Mastrogiacomo 2011, Unpublished). The effects of the bulge can be seen in Table 8.

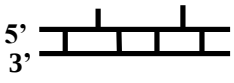
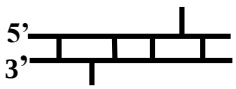
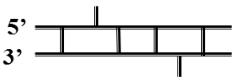
Base Pairs Outside Bulges	**Base Pairs Between Bulges**	 **$\Delta G°_{37}$ (kcal/mol)**	 **$\Delta G°_{37}$ (kcal/mol)**	 **$\Delta G°_{37}$ (kcal/mol)**
1	4	2.1, 2.0	2.6, 3.7, 2.6, 3.8	-
1	5	-	3.5	1.5,1.6, 2.2
1	6	-	1.2, 7.6, 2.5,1.9, 1.5, 3.1	1.1, 1.8, 5.7
2	2	-	4.0, 3.9, 3.9	4
2	**3**	7.0	**5.4**	**4.4**
2	**4**	6.3, 7.1	6.7	6.4
2	**5**	-	8.8	-
2	**6**	-	8.5	**10**
3	**1**	6.8, 6.3	**4.0**	-
3	**2**	9.1, 7.8	4.7, 8.3	2.3, 8.0, 11.0
3	**3**	6.9	**6.8**	**6.7**
3	4	9.5, 10.9, 10.9	8.1	9.1
3	5	-	10.9	-
3	6	-	11.0, 11.4	-
4	3	8.6	-	-
4	4	-	-	11.3

Table 8. The Effect of the Two Group I Bulges on the Stability of the Duplex. Previously collected data combine with data from the current study on how the arrangement of the bulges affects the stability of the duplex (Jones 2007, Morelli 2009, Mastrogiacomo 2011, Unpublished)

From the table the effect of the two Group I bulges can be seen. All the values give positive free energies so all of the bulges have a destabilizing effect.

Discussion

Working with the secondary structure motifs of RNA and the thermodynamic parameters, is how models are developed to predict and help to further understand the secondary structure of RNA. Using the secondary structure of RNA the tertiary structure of RNA can be predicted. The revised Blose model currently predicts that with a single bulged nucleotide the effect on the free energy of the duplex is such that the more stable the least stable stem the greater the bulges effect on the stability of the duplex (McCann et al., 2011). However this model does not account for multiple bulge loops and how they can affect the stability of the duplex and of each other.

Comparing the value for the second bulge versus the effect of the both bulges divided by two (Table 5) gives further insight as to whether the bulges have the same or different effects on the duplex (Table 9). Examples of these calculations are seen in Appendix E.

Duplex	Bulge Average ΔG°37 (kcal/mol)	2nd Bulge ΔG°_{37} (kcal/ mol)
GC ACU**A**CG CG**A**UGA GC	2.72	3.49
GC**A**UGA GC CG ACU**A**CG	2.21	1.92
GUC ACU**A**CUG CAG**A**UGA GAC	3.42	3.92
GC GGCC**A**GC CG**A**CCGG CG	3.01	3.51
GUC**A**UCA CUG CAG AGU**A**GAC	3.33	3.75
GC**A**UGCGCA GC CG ACGCGU**A**CG	5.08	7.64
CAG G**A**CUG GUC**A**C GAC	2.01	3.32

Table 9. Comparison of Average Free Energy versus Free Energy of 2nd Bulge The calculated values for the effect of each bulge either by determining the bulges to have the same effect or calculating the second bulge as related to, but not exactly the same as the first bulge.

From Table 9 it can be seen that the two bulges do not have the same effect on the duplex as was previously observed (Jones 2007, Morelli 2009, Mastrogiacomo 2011, Unpublished). This may because each of the bulges may change which stem is the least stable, and the bulges may interact. In Table 9 depicts that the two bulges do not have the same effect as the values for the second bulge and average of the two bulges are not the same. The fact that both bulges do not have the same effect is not accounted for the in the revised Blose model.

When looking at Table 7 it is also possible to determine if the revised Blose model is accurate or not for two single nucleotide bulges. The predicted values for the effect the bulges have on the free energy of the duplex are both over and underestimations. All of the predicted values are at least 0.5 kcal/mol off from the measured free energy. This is greater than the experimental error for most of the duplexes studied. This can be seen with the sequence $\begin{matrix} GC\ \ ACU\mathbf{A}CG \\ CG\mathbf{A}UGA\ \ GC \end{matrix}$, where the predicted effect of the bulge on the free energy is 6.6 kcal/mol, while the measured free energy is 5.0 kcal/mol. The difference between the two is 1.5 kcal/mol. This suggests that the revised Blose model does not account for more than one single nucleotide bulge. The reason for the variation may have to do with the sequences and the placement of different base pairs. Some things that might change how the bulges interact with the sequences are: the number of base pairs outside the bulges and between the bulges, the placement of the bulges as depicted in Table 8, the overall sequence, and the sequence of the three different stems. When working with the least stable stem in all but one of the sequences it was one of the end stems. The one sequence where the least stable stem was the inner duplex is $\begin{matrix} CAG\ \ G\mathbf{A}CUG \\ GUC\mathbf{A}C\ \ GAC \end{matrix}$, where the G-C pair of the inner duplex is the least stable stem rather than the 3 base pair end stems. Not having the least stable stem on the end may be a factor that needs to be

looked into further. Also, the effects of having G-C rich sequences versus A-U rich sequences should be looked at further, because they may affect how the bulge influences the stability. The duplexes in this study are also seen to always be less stable than predicted, which has been observed previously (Jones 2007, unpublished). Prior work determined that the bulge has a lesser destabilizing effect if the adjacent stem is less stable (Blose et al., 2007). The revised Blose model does not account for two bulges and the fact that the first bulge will change the stability of the stems for the second bulge. This could mean that the least stable stem changes for the second bulge. To understand the effect of the different bulges several comparisons have to determine what factors are important for bulge loops stability.

Duplex Stability

A comparison of the free energy of the parent duplex, the sequence without the bulges, and the effect of the bulge demonstrates the effect of the bulges on the stability of the sequence. The parent duplex includes the effects of all three of the different stems that are present in the duplex. This may give a correlation between the parent sequence stability and the bulge stability. However, it must be considered that this allows for the stems to interact; an accurate portrayal of the effect of the bulge may not be possible. It has been stated that the nearest neighbors to the bulge do not affect the stability of the bulge (McCann et al., 2011). When the parent duplex is calculated there are nearest neighbor interactions occurring that are not present when the bulge is present, because now all three stems are interacting with each other. This may cause the effect of the bulge on the free energy to not be accurately calculated. The interactions of the ends of the stems may result in factors combining or cancelling one another out, leading the model to be off. The comparison can be seen in Figure 11.

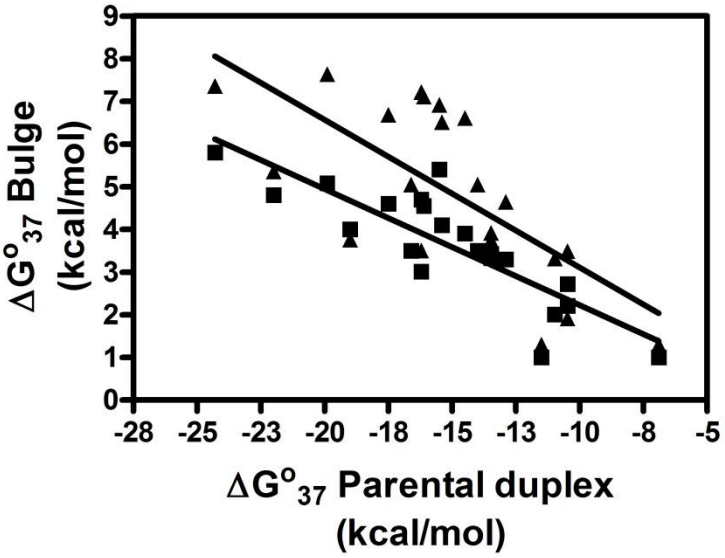


Figure 11. Free Energy of Parent Duplex versus the Bulge Free Energy (kcal/mol) The triangles represent the $\Delta G^{\circ}_{37,\ 2nd\ Bulge}$ (kcal/ mol), while the squares represent the $\Delta G^{\circ}_{37,\ Average\ Bulge}$ (kcal/ mol). The top linear regression is for the 2nd bulge, while the bottom linear regression is representative of the average bulge. This data was compiled with previously collected data (Jones 2007, Mastrogiacomo 2011, unpublished).

From Figure 11 as both linear regressions have a negative slope representing the fact that the bulges are less destabilizing with duplexes that are already less stable. Using *GraphPad Prism* the slope of both lines was determined to be significant (Appendix F). Looking at Figure 11 it can also be seen that the linear regressions are representative of the data and there does not appear to be any outliers. The slope of the line representing the $\Delta G^{\circ}_{37,\ 2nd\ Bulge}$ is -0.35 while for $\Delta G^{\circ}_{37,\ Average\ Bulge}$ the slope is -0.27 (Appendix F). Given that both linear regressions are significant there is a correlation between the parent duplex and the effect that the bulge has on the sequence. The more stable the parent duplex is the greater effect the bulge seems to have. Also, the free energy of the second bulge compared with the parent duplex has a greater slope, which suggests that it has more of an effect on the stability of the duplex. More data would help to make these trends more defined.

Least Stable Stem Effects on Stability

Using the revised Blose model the effect of a single nucleotide bulge is determined based on the stability of the least stable stem in the duplex (McCann et al., 2011). To see if the least stable stem is what effects the stability of multiple single nucleotide bulges the free energy of the least stable stem was graphed against the free energy of the average bulge and the second bulge (Figure 12). The least stable stem is calculated as described in the results.

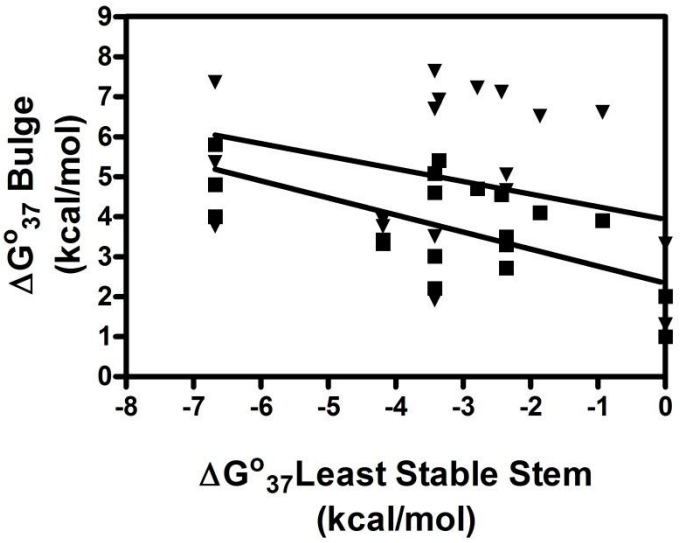


Figure 12. Free Energy of Least Stable Stem versus the Free Energy of the 2nd Bulge and the Average Bulge (kcal/mol) The triangles represent the $\Delta G^{\circ}_{37,\text{ 2nd Bulge}}$ (kcal/ mol), while the squares represent the $\Delta G^{\circ}_{37,\text{ Average Bulge}}$ (kcal/ mol). The top linear regression is for the 2nd bulge, while the bottom linear regression is representative of the average bulge. This data was compiled with previously collected data (Jones 2007, Mastrogiacomo 2011, unpublished).

Using *GraphPad Prism* to preform statistical analysis the slope of the line for the free energy of the least stable stem versus the free energy of the average bulge is significant (Appendix F). While for the linear regression representing the free energy of the least stable stem versus the free energy of the second bulge is not significant. This makes sense because the majority of the points (triangles) representing the free energy seem to be higher on the graph than the linear regression as there are several points scattered all over the graph. With more data points it may

help to clarify this trend. The slope of the free energy of the average bulge and the lease stable stem has negative slope of -0.43, suggesting a correlation between the stem becoming less stable and the bulge loop destabilized the duplex less. The free energy of the second stem does not have a significant slope, but the data still seems to have a negative trend, which suggests a correlation between the stability of the least stable stem and the bulge. This data supports the revised Blose model, as with the increased stability of the least stable stem the greater the effect the bulge has on the free energy of the duplex (McCann et al., 2011).

Predicted Bulge Stability:

To determine if the revised Blose model if representative of the effect of the two single nucleotide bulges, a comparison of the predicted free energy is graphed versus the second bulge (Figure 13).

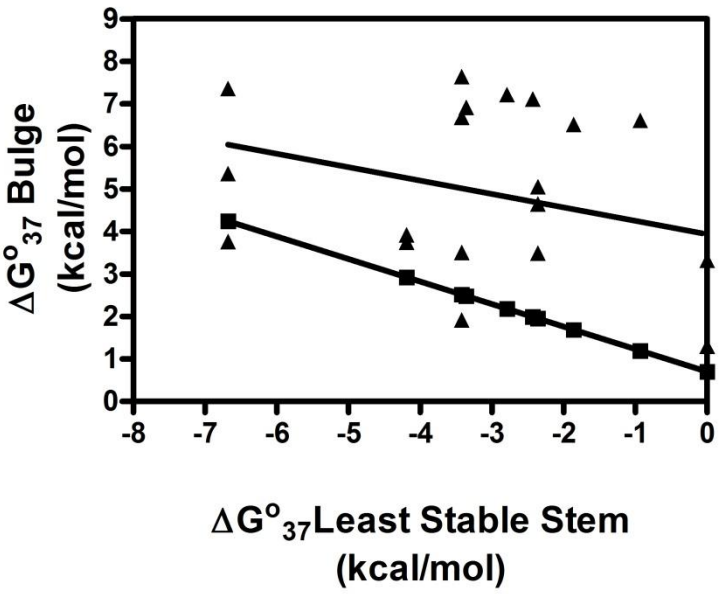


Figure 13. Free Energy of Least Stable Stem versus the Free Energy of the Predicted Bulge and the Second Bulge (kcal/mol) The triangles represent the $\Delta G^{\circ}_{37,\ 2nd\ Bulge}$ (kcal/ mol), while the squares represent the $\Delta G^{\circ}_{37,\ Predicted\ Bulge}$ (kcal/ mol). The top linear regression is for the 2nd bulge, while the bottom linear regression is representative of the average bulge. This data was compiled with previously collected data (Jones 2007, Mastrogiacomo 2011, unpublished).

The downward trend is still present for both lines, which agree with the revised Blose model that with decreasing the stability of the least stable stem, the less the effect is from the bulge. As was previously mentioned the slope between the free energy of the least stable stem and the second bulge is not significant using *GraphPad Prism* (Appendix F). Using *GraphPad Prism* the slope of the line for the free energy versus the free energy of the predicted bulge is significant, however this is expected (Appendix F). The reason this is expected is because the predicted value is obtained using Equation 5, so it is calculated using the least stable stem and no other variables so they are directly correlated. From the graph about a 3 kcal/mol difference is observed between the predicted effect of the bulge and the effect of the second bulge. This is consistent with what was previously determined by Jones, although she had less data (Jones 2007, unpublished). This difference between the predicted and the second bulge suggests that there are factors affecting the stability of the duplex that the revised Blose model does include. There is some factor that causes the second bulge to destabilize the duplex more than the first that is not being accounted for. This may be because the first bulge destabilizes the duplex, changing the least stable stem, so then the second bulge can have a greater effect than just a single nucleotide bulge on the duplex. Determining this factor would help to make the revised Blose model more accurate, because the effect of more than one bulge loop would be incorporated.

A graph was also made of the free energy of the least stable stem versus the free energy of the average free energy of the bulge (Figure 14).

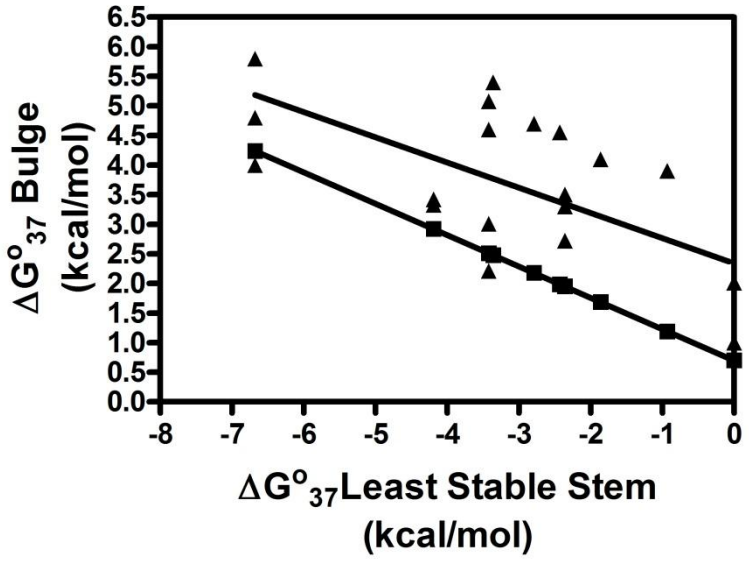


Figure 14. Free Energy of Least Stable Stem versus the Free Energy of the Average Bulge Predicted Bulge (kcal/mol) Bulge Free Energy (kcal/mol) The triangles represent the $\Delta G°_{37,\text{ Average Bulge}}$ (kcal/ mol), while the squares represent the $\Delta G°_{37,\text{ Predicted Bulge}}$ (kcal/ mol). The top linear regression is for the average bulge, while the bottom linear regression is representative of the average bulge. This data was compiled with previously collected data (Jones 2007, Mastrogiacomo 2011, unpublished).

A downward trend is again seen for both data sets, which is consistent with the revised Blose model. As was previously mentioned the slope between the free energy of the least stable stem and the second bulge is significant using *GraphPad Prism* (Appendix F).The slope of the line for the free energy versus the free energy of the predicted bulge is significant, although this should be expected since the predicted value is calculated based on the least stable stem using *GraphPad Prism* (Appendix F). From the graph it can be seen that there is about a 2 kcal/mol difference between the predicted effect of the bulge and the effect of the second bulge. This would suggest that the data and the predicted are closer in value that is seen previously with Figure 13, which compares the free energy of the predicted to the second bulge. It has been demonstrated that the bulges do not have the same effect on the stability of the duplex. While the method of determining the free energy of the second bulge is one of many ways to make the

calculations. This makes it difficult to determine inaccuracies that are present in the revised Blose model. Although both methods of determining the effect of the bulge show that the model gives a free energy for the bulges that is less destabilizing than what is experimentally found. There is still some factor that must be determined and added to the revised Blose model.

Current Study

For the current study, the purpose was to compare the effect of the two versus three base pairs outside the bulge to see the effect of the stability of the duplex containing two single-adenosine bulges. In this study the least stable stem was one of the outside stems, except in once case (Table 6). The one exception is $\begin{matrix} CAG & G\mathbf{A}CUG \\ GUC\mathbf{A}C & GAC \end{matrix}$, where the G-C base pair of the inner duplex is the least stable. For six sequences where the end stem was least stable two base pairs outside the bulge had a less stable stem, than that of the sequences with three base pairs outside the bulge. This excludes the sequence where the inner duplex (that between the bulges) was one. So in following the revised Blose model the sequences with two base pairs outside the bulge would see less of an effect from the bulges than those with three base pairs outside the bulge. This does not appear to be the case, but more data would be necessary to make a conclusion. From the compiled data it appears that the both the number of base pairs outside the bulges and between the bulges have an effect on the stability of the duplex (Table 8). There may also be a correlation between the locations of the bulges, but more data would be need to make a conclusion on this (Table 8).

Bulge Confirmation

The placement of the bulges must also be considered, because the interactions that are possible depending on the placement of the bulge must also be considered. Depending on the location of the bulges, as depicted in Table 8, they may have the ability to interact with each

other once the duplex forms. On paper it would appear the bulges on the same side of the helix would interact, but in three dimensions any the different arrangements of the bulges would be able to interact.

The different families of bulges will also have different effects. As previously discussed a bulge is in one of the following families: intercalated, side-by-side, or extrahelical (Hermann et al., 2000; Znosko, 2002). These different families are expected to give different thermodynamic parameters. One potential reason that the different bulge families would give different thermodynamic data would be because of the loss of base stacking interactions (Barthel et al., 2006). However, base stacking interactions have not been observed to occur with single nucleotide bulges (Longfellow, 1990). With multiple single nucleotide bulged bases this may be a factor that is not considered in the revised Blose model. More research needs to be done to look into the thermodynamic effects of these different families of bulges, so see if they result in changes to the thermodynamics of the structure (Barthel et al., 2006). A better understanding of the families will allow for the bulge families to either be incorporated into future models or disregarded as it is not a factor that affect stability.

Apparent Trends

All of the comparisons have shown that there is a correlation between the least stable stem and the effect of the bulges in a manner that agrees with the revised Blose model. It is seen that as the stability of the least stable stem decreases that the destabilization caused by the bulge decreases. This may be because the more stable the least stable stem, a bulge has the ability to destabilize the duplex more. This is because something that is already not stable won’t change much with another factor, but something that is very stable changing it confirmation will have much greater effect. The least stable stem has fewer interactions and bonds that are disrupted,

which would make it more logical that the more stable least stable stems are affected more by the bulge. This study again goes to show that average of the bulges is not the same as the effect of the second bulge, meaning that the two bulges do not have the same effect (Jones 2007, Morelli 2009, Mastrogiacomo 2011, unpublished).

Conclusions

It was determined that two Group I bulged adenosines have a different effect on the stability of a duplex than just combining the effect of the two single nucleotide bulges. The two bulges destabilize the duplex more than is predicted by the revised Blose model. The effect of the bulges is has a greater positive free energy than what was predicted by about 2-3 kcal/mol. A trend is seen between the average effect of the bulges and the least stable stem. It shows that the less stable the least stable stem, the less of an affect the bulges have. This agree with the revise Blose model. The working with the location of the bulges in the duplex and the number outside the bulge (Table 8), there does not seem to be a trend. That could be a result of a small data set, so continuing the research may reveal more about these potential relationships.

Further Studies

Further work could be down working with comparing thermodynamic data of the two single nucleotide bulges helix, with the sequences that are the same but only have one single nucleotide bulge. All of this data could be compiled with the predicted effect for each of the three sequences being compared. This would help to see how the two bulges affect the sequence differently as compared with a duplex with both single nucleotide bulges. From this it would be possible to look at how the predicted values differ from the measured, and potentially determine some reasons for the difference.

Further work can be done to complete Table 8, to see if there is a correlation between the number of base pairs either outside the bulges or between the bulges. When working with this it would also be valuable to look further into how changing the sequence of the inner duplex changes the stability as was looked at by Jones in 2007 (Jones 2007, unpublished). More sequences must also be melted for Table 8 to determine if there is a consistency with each combination of base pairs outside the bulges and between the bulges, because currently there are some outliers. If there are several outliers further research would need to be done into why there are those abnormalities.

When working with the data it was seen when graphing the $\Delta G^{\circ}_{37,\text{Least Stable Stem}}$ versus both the $\Delta G^{\circ}_{37,\text{ Average Bulge}}$ and $\Delta G^{\circ}_{37,\text{ 2nd Bulge}}$ there almost seen to be two trend lines forming. Different possibilities reason why this was occurring was looked into including: the placement of the bulges on the duplex, the sequence of the three stems of the duplex, and the number of base pairs outside the bulges and between the bulges. The only trend that was observed was that in general the lower potential trend line had an inner duplex that was generally guanine and

cytosine rich. While the potential upper trend line was uracil and adenine rich. How this was not true for every sequence, there are some that do not follow the trend. Some potential sequences that could studied to further analyze this are:

1. *CG UUAAU***AGC**
   *GC***A***AUUA CG*
2. *CAG AUUA***A***GUC*
   *GUC***A***UAAU CAG*
3. *CG***A**GCCGC *GC*
   *GC GGGCG***A***CG*
4. *CAUG***A**GCG *GUAC*
   *GUAC CGC***A***CAUG*

References

Barthel, A.; Zacharias M. Conformational Transitions in RNA Single Uridine and Adenosine Bulge Structures: A Molecular Dynamics Free Energy Simulation Study. *Biophys. J.* **2006**, *90*, 2450–2462.

Batey, R. T.; Rambo, R. P.; Doudna, J.A. Tertiary Motifs in RNA Structure and Folding. *Agnew. Chem. Int. Ed.* **1999**, *38*, 2326-2343.

Bhattacharyya, A.; Murchie, A. I .H.; Lilley, D. M. J. RNA bulges and the helical periodicity of double-stranded RNA. *Nature*, **1990**, *343*, 484-487.

Bloomfield, V. A.; Crothers, D. M. Tinoco, I.; Nucleic Acids; Structures, Properties, and Functions. **2000**. Sausalito, California: University Science Books.

Blose, J. M.; Manni, M. L.; Klapec, K. A.; Stranger-Jones, Y.; Zyra, A. C.; Sim, V.; Griffith C. A.; Long, J. D.; Serra, M. J. Non-Nearest-Neighbor Dependence of the Stability for RNA Bulge Loops Based on the Complete Set of Group I Single-Nucleotide Bulge Loops. *Biochemistry*, **2007**, *46*, 15123-15135.

Chen, Y.; Varani, G. RNA Structure, **2010**, eLS.

Crick, F. H. C. Codon-Anticodon Pairing: The Wobble Hypotheis. *J. Mol. Biol.* **1966**, *19*, 548-555.

Frier, S. M.; Kierzek, R.; Jaeger, J.A.; Sugimoto, N.; Caruthers, M. H.; Neilson, T.; Turner, D. H. Improved free-energy parameters for predictions of RNA duplex stability. *Biochemistry*, **1986**, *83*, 9373-9377.

Garrett, R.H., Grisham, C. M. Biochemistry. 4th ed. **2008**, Boston, MA: Brooks/Cole.

Gesteland, R. F.; Cech, T. R.; Atkins, J. F. The RNA World, Second Edition. **1999**, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Gutell, R. R.; Cannone, J. J.; Shang, Z.; Du, Y.; Serra, M. J. A Story: Unpaired Adenosine Bases in Ribosomal RNAS, *J. Mol. Biol.*, **2000**, *304*, 335-354.

Herman, T.; Patel, D. J.; RNA bulges as architectural and recognition motifs. *Structure*, **2000**, *8*, R47- R54.

Jones, L. E. A Thermodynamic Study of Multiple Single-Adenosine Bulge Loops in RNA Duplexes. *Allegheny College Senior Comprehensive Project*. **2007**. Unpublished.

Ke, S.; Wartell, R. M.; Influence of Neighboring Base Pairs on the Stability of Single Base Bulges and Base Pairs in a DNA Fragment. *Biochemistry*, **1995**, *34*, 4593-4600.

Longfellow, C. E.; Kierzek, R.; Turner, D.H. Thermodynamic and Spectroscopic Study of Bulge Loops in Oligoribonucleotides. *Biochemistry*, **1990**, *29*, 278-285.

Lu. Z. J.; Turner D. H.; Mathews, D. H. A set of nearest neighbor parameters for predicting the enthalpy change of RNA secondary structure formation. *Nucleic Acids Research*, **2006**, *17*, 4912-4924.

Mastrogiacomo E. M. A Thermodynamic Study of Multiple Single-Nucleotide Bulge Loops in RNA Duplexes. *Allegheny College Senior Comprehensive Project*. **2012**. Unpublished.

McCann, M. D.; Lim, G. F. S.; Manni, M. L., Estes, J.; Klapec, K. A.; Frattini, G. D.; Knarr, R. J.; Gratton, J. L.; Serra, M. J. Non-nearest-neighbor dependence of the stability for RNA group II single-nucleotide bulge loops. *RNA*, **2011**, *17*, 108-119.

Morelli J. N. A Thermodynamic Study of Multiple Single-Base Bulge Loops in RNA Duplexes. *Allegheny College Senior Comprehensive Project*. **2009**. Unpublished.

Peritz, A. E.; Kierzek, R.; Sugimoto, N.; Turner, D. H. Thermodynamic Study of Internal Loops in Oligoribonucleotides: Symmetric Loops Are More Stable Than Asymmetric Loops. *Biochemistry*, **1991**, *30*, 6428-6436.

Price, W. H. The Isolation of a New Virus associated with Respiratory Clinical Disease in Humans. *Bacteriology*, **1956**, *42*, 892-896.

SantaLucia, J.; Turner, D. H. Measuring the Thermodynamics of RNA Secondary Structure Formation. *Biopoly.*, **1998**, *44*, 309-319.

Schroeder, S. J.; Turner, D. H. Optical Melting Measurements of Nucleic Acid Thermodynamics. *Methods in Enzymology*. **2009**, *468*, 371-387.

Shen, L X.; Cai Z.; Tinoco I. RNA Structure at High Resolution. *FASEB*. **1995**. *9*, 1023-1035.

Sugimoto. N.; Kierzek, R.; Freier, S. M.; Turner, D. H. Energetics of Internal GU Mismatches in Ribooligonucleotide Helixes. *Biochemistry*, **1986**, *25*, 5755-5759.

Tinoco, I; Bustamante, C. How RNA Folds. *J. Mol. Biol.*, **1999**, *293*, 271-281.

Toor, N.; Keating, K. S.; Fedorova, O.; Rajashankar, K.; Wang, J; Pyle, A. M. Tertiary architecture of the *Oceanobacillus iheyensis* group II intron. *RNA*, **2010**, *16*, 57-59.

Turner, D. H. Bulges in nucleic acids, *Current Opinions in Structural Biology*, **1992**, *2*, 334-337.

Turner, D.; Sugimoto, N.; Freier, S. M. RNA structure prediction. *Annu. Rev. Biophys. Chem.*, **1988**, 17, 167-192.

Usman, N.; Ogilvie, K. K.; Jiang, M. Y.; Cedergren, R. J. Automated Chemical Synthesis of Long Oligoribonucleotides using 2'-O-Silylated Ribonucleoside 3'-O-Phosphoramidites on a Controlled-Pore Glass Support: Synthesis of a 43-Nucleotide Sequence Similar to the 3'-Half Molecule of an *Escherichia coli* Formylmethionine in tRNA. *J. Am. Chem. Soc.*, **1987**, *109*, 7845-7854.

von Ahsen, U.; Schroeder R. RNA as a catalyst: natural and designed ribozymes. *BioEssays*. **1993**. *15*, 299-307.

Watson, J. D.; Baker, T. A.; Bell, S. P.; Gann, A.; Levine, M. Molecular Biology of the Gene, Sixth Edition. **2007**. Benjamin Cummings.

Watson, J. D.; Crick, F. H. Molecular Structure of Nucleic Acids: A structure of Deoxyribose Nucleic Acid. *Nature*. **1953**, *171*, 737-738.

Wu, J. C.; Gardner, D. P.; Ozer, S.; Gutell, R. R.; Ren, P. Correlation of RNA Secondary Structure Statistics with Thermodynamic Stability and Applications to Folding. *J.Mol.Biol.*¸ **2009**, *391*, 769-783.

Wu, H.; Uhlenbeck, O.C. Role of a Bulged Residue in a Specific RNA-Protein Interaction. *Biochemistry*, **1987**, *26*, 8221-8227.

Xia,T.; SantaLucia, J.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X.; Cox, C.; Turner D. H. Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson-Crick Base Pairs. *Biochemistry*, **1998**, *37*, 14719-14735.

Xiong Y.; Deng, J.; Sudarsanakumar C.; Sundaralingam M. Crystal Structure of an RNA duplex r(gugucgcac)$_2$with uridine bulges. *J. Mol. Biol.* **2001**. *313*, 573-582.

Xiong, J.; Sundaralingam M. Two Crystal Form of Helix II of Xenopus Laevis 5S rRNA with a Cytosidne Bulge. *RNA*, **2000**, *16*, 1316-1324.

Znosko, B. M.; Silvestri, S. B.; Volkman, H.; Boswell, B.; Serra, M. J. Thermodynamic Parameters for and Expanded Nearest-Neighbor Model for the Formation of RNA Duplexes with Single Nucleotide Bulges. *Biochemistry*, **2002**, *41*, 10406-10417.

Appendix A: *Meltwin 3.0* Equation 1 Derivation

$$A_T = (\{[e^{(T\Delta S° - \Delta H°)/RT)2} + 2\left(e^{(T\Delta S° - \Delta H°)/RT)} C_T\right] - e^{(T\Delta S° - \Delta H°)/RT)}\}/2) \ * \\ (2M_{SS} + 2b_{SS} - M_{DS}T - b_{DS}) + (\frac{C_T}{2})(M_{DS}T + b_{DS}))$$

Where:

A_T = total absorbance of sample
M_{SS} = slope of linear baseline of random coil form of single-stranded RNA
M_{DS} = slope of linear baseline of RNA-RNA duplex
b_{SS} = the y-intercept of the linear plot of the random coil form of single-stranded RNA
b_{DS} = the y-intercept of the linear plot for the RNA-RNA duplex
C_T = total concentration of the single-stranded RNA
T = temperature in Kelvin
R = 10986 cal/mol*K

Beer's Law

$$A = \varepsilon l C \quad \text{(A1)}$$

A = Absorbacne
ε = extinction coefficient
$l =$ cell pathlength
$\mathrm{C} =$ concentration

Using the known equilibrium of the two state model, the following expression is determined:

$$[\mathrm{DS}] \leftrightarrow [\mathrm{SS}] \quad \text{(A2)}$$

DS = double stranded RNA
SS= single stranded RNA

$$A_T = A_{SS1} + A_{SS2} + A_{DS} \quad \text{(A3)}$$

A_T = total absorbance of the sample
A_{SS1} = absorbance of single- stranded RNA
A_{DS} = absorbance of the double-stranded RNA

Combing A1 and A3, the following is obtained:

$$A_T = \varepsilon_{SS1} l C_{SS1} + \varepsilon_{SS2} l C_{SS2} + \varepsilon_{DS} l C_{DS} \quad \text{(A4)}$$

A_T = total absorbance of sample
ε_{SS} = extinction coefficients of the single-stranded RNA
ε_{DS} = extinction coefficients of the double-stranded RNA
l = cell pathlength
C_{SS} = cocentrations of the single-stranded RNA
C_{DS} = concentrations of the double-stranded RNA

When the melt starts there are equal amounts of single-stranded RNA:

$$C_{SS1} = C_{SS2} = C_{SS} \quad \text{(A5)}$$

Substituting A5 into A4 and simplifying using ε_{SS} = average of ε_{SS1} and ε_{SS2} gives:

$$A_T = 2\varepsilon_{SS} l C_{SS} + \varepsilon_{DS} l C_{DS} \quad \text{(A6)}$$

Combining the concentrations of RNA gives:

$$C_T = 2C_{SS} + 2C_{DS} \quad \text{(A7)}$$

C_T = total concentration of the RNA
C_{SS} = concentrations of the single-stranded RNA
C_{DS} = concentrations of the double-stranded RNA

Solving A7 for C_{DS}:

$$C_{DS} = \left(\frac{C_T}{2}\right) - C_{SS} \quad \text{(A8)}$$

Substituting A8 in A6, and assuming that the cell pathlength is 1 cm gives:

$$A_T = 2\varepsilon_{SS} C_{SS} + \varepsilon_{DS}\left(\left(\frac{C_T}{2}\right) - C_{SS}\right) \quad \text{(A9)}$$

The extinction coefficients can be expressed as a linear function in terms of temperature:

$$\varepsilon_{SS} = m_{SS} T + b_{SS} \quad \text{(A10)}$$

m_{SS} = slope of the linear sloping baseline of single-stranded RNA0
T = temperature in Kelvin
b_{SS} = y-intercept of the linear sloping baseline of single-stranded RNA

$$\varepsilon_{SS} = m_{DS} T + b_{DS} \quad \text{(A11)}$$

m_{DS} = slope of the linear sloping baseline of double-stranded RNA
T = temperature in Kelvin
b_{DS} = y-intercept of the linear sloping baseline of double-stranded RNA

Substituting A10 and A11 into A9 gives:

$$A_T = 2(m_{SS} T + b_{SS}) C_{SS} + m_{DS} T + b_{DS}\left(\left(\frac{C_T}{2}\right) - C_{SS}\right) \quad \text{(A12)}$$

A12 can be rearrange to give:

$$A_T = C_{SS}(2m_{SS} T + 2b_{SS} - m_{DS} T - b_{DS}) + \left(\frac{C_T}{2}\right)(m_{DS} T + b_{DS}) \quad \text{(A13)}$$

Two known equations are used that involve Gibbs free energy:

$$\Delta G^0 = \Delta H^0 - T\Delta S^0 \quad \text{(A14)}$$

$$\Delta G^0 = -RTlnK \quad \text{(A15)}$$

R = 1.986 cal/(mol*K)

Combining A14 and A15 gives:

$$\Delta H^0 - T\Delta S^0 = -RTlnK \quad \text{(A16)}$$

A16 can be simplified:

$$K = e^{\frac{T\Delta S^0 - \Delta H^0}{RT}} \quad \text{(A17)}$$

The equilibrium constant of A2 can be written as:

$$K = \frac{[(C_{SS1})(C_{SS2})]}{(C_{DS})} \quad \text{(A18)}$$

A18 is substituted into A5 and rearranged to get the following:

$$C_{SS}{}^2 = KC_{DS} \quad \text{(A19)}$$

A8 can then be substituted into A19 and rearranged:

$$C_{SS}{}^2 = K\left(\frac{C_T}{2}\right) - KC_{SS}\,; 0 = 2C_{SS}{}^2 - KC_T + 2C_{SS} \quad \text{(A20)}$$

Using the quadratic Equation A20 is solved to be:

$$C_{SS} = \frac{\left[(K^2 + 2KC_T)^{\frac{1}{2}} - K\right]}{2} \quad \text{(A21)}$$

A17 can be substituted into A21:

$$C_{SS} = \frac{\left\{\left[e^{2\frac{T\Delta S^\circ - \Delta H^\circ}{RT}} + 2Ke^{\frac{T\Delta S^\circ - \Delta H^\circ}{RT}}C_T\right]^{\frac{1}{2}} - e^{(T\Delta S^\circ - \Delta H^\circ)/RT)}\right\}}{2} \quad \text{(A22)}$$

Substituting A22 into 12 give Equation 2:

$$A_T = \left(\frac{\left\{\left[e^{(T\Delta S^\circ - \Delta H^\circ)/RT)2} + 2\left(e^{(T\Delta S^\circ - \Delta H^\circ)/RT)}C_T\right] - e^{(T\Delta S^\circ - \Delta H^\circ)/RT)}\right\}}{2}\right) *$$
$$(2M_{SS} + 2b_{SS} - M_{DS}T - b_{DS}) + (\frac{C_T}{2})(M_{DS}T + b_{DS}))$$

Appendix B: Derivation of T_M Dependence Equation for Meltwin3.0 Equation 2

$$T_M{}^{-1} = \left(\frac{R}{\Delta H°}\right) * \ln(C_T) + \left(\frac{\Delta S°}{\Delta H°}\right)$$

The equilibrium between double-stranded and single-stranded gives the following expression:

$$[DS] \leftrightarrow [SS]_1 + [SS]_2 \quad \text{(B1)}$$

The equilibrium equation for B1 is:

$$K = \frac{(C_{SS1} C_{SS2})}{C_{DS}} \quad \text{(B2)}$$

The initial concentrations of the single-stranded RNA are equal

$$C_{ss1} = C_{ss2} = C_{SS} \quad \text{(B3)}$$

Substituting B3 into B2 gives:

$$K \frac{C_{DS}{}^2}{C_{DS}} \quad \text{(B4)}$$

Two equations commonly used for Gibbs free energy are:

$$\Delta G^0 = \Delta H^0 - T\Delta S^0 \quad \text{(B5)}$$

$$\Delta G^0 = -RTlnK \quad \text{(B6)}$$

B5 and B6 can be set equal to each other:

$$\Delta H^0 - T\Delta S^0 = -RTlnK \quad \text{(B7)}$$

B7 can be arranged to give:

$$K = e^{\frac{T\Delta S^0 - \Delta H^0}{RT}} \quad \text{(B8)}$$

B8 and B4 can be combined as they are equal to the same thing, and simplified:

$$\frac{C_{SS}{}^2}{C_{DS}} = e^{\frac{T\Delta S^0 - \Delta H^0}{RT}} \quad \text{(B9)}$$

Simplifying B9 gives:

$$\ln\left(\frac{C_{SS}{}^2}{C_{DS}}\right) = \frac{T\Delta S^0 - \Delta H^0}{RT} \quad \text{(B10)}$$

At T_M, the concentration of single stranded and double-stranded RNA are equal:

$$C_{SS1} + C_{SS2} = 2C_{SS} = C_{DS} \quad \text{(B11)}$$

Combining B11 and B10 at the melting temperature gives:

$$\ln\left(\frac{{C_{SS}}^2}{2C_{SS}}\right) = \frac{T_M\Delta S^0 - \Delta H^0}{RT_M} \quad \text{(B12)}$$

The transition equilibrium is now used, rather than the melting equilibrium $[SS_1] + [SS_2] \leftrightarrow [DS]$:

$$-\ln\left(\frac{C_{SS}}{2}\right) = \frac{(T_M\Delta S^0 - \Delta H^0)}{RT_M} \quad \text{(B13)}$$

$$-RT_M \ln\left(\frac{C_{SS}}{2}\right) = T_M\Delta S^0 - \Delta H^0$$

$$\Delta H^0 = T_M\Delta S^0 + RT_M\, ln\left(\frac{C_{SS}}{2}\right)$$

$$\frac{\Delta H^0}{T_M} = \Delta S^0 + RT_M\, ln\left(\frac{C_{SS}}{2}\right)$$

$${T_M}^{-1} = \left(\frac{\Delta S^0}{\Delta H^0}\right) + \left(\frac{R}{\Delta H^0}\right) * ln\left(\frac{C_{SS}}{2}\right) \quad \text{(B14)}$$

At high temperatures, C_T is calculated as $2C_{SS}$ for self complimentary, which gives Equation 3:

$${T_M}^{-1} = \left(\frac{R}{\Delta H^\circ}\right) * \ln(C_T) + \left(\frac{\Delta S^\circ}{\Delta H^\circ}\right)$$

At high temperatures, C_T is calculated as $\frac{C_{SS}}{2}$ for non-self complimentary:

$${T_M}^{-1} = \left(\frac{R}{\Delta H^\circ}\right) * \ln(C_T) + \left(\frac{\Delta S^\circ}{\Delta H^\circ}\right) \quad \text{(B15)}$$

Appendix C: Detailed Purification and Melting Process

Purification Procedure (1 μM scale)

Step 1

During the synthesis process, the oligomers are manufactured inside a small plastic vial which is sealed upon completion. The oligomers are synthesized while being chemically attached to the small glass beads. The standard scale is 30mg of bead per micromole of RNA. First, the sequence sample was recorded for later reference in the notebook. Using pliers, the small metal seal was cut and gently pulled away from one end of the sealed vial. The plastic cap was pried away from the vial and a small piece of filter paper was removed by using a needle. A 1.5 ml screw top vial and matched o-ring cap were acquired from the RNA purification drawer (in C104) and both the vial and lids were carefully labeled denoting each sequence in permanent ink. While working over a small sheet of 6" X 6" weighing paper, the beads were transferred into the 1.5 ml vial and 250 microliters of 100% ethanol was added with 750 microliters of cold, fresh ammonium hydroxide. The screw top vial was tightly sealed, wrapped in parafilm and placed in a 55 C water bath for 18 hours with frequent vortexing of the sample. The water bath was filled with distilled water to a point just above the test tube rack. During the incubation, the ammonium hydroxide cleaved the ester linkage between the bead and oligomer. The presence of the ethanol as a co-solvent allowed the non-polar protecting groups to dissolve.

Step 2

After the incubation period, the plastic vial was cooled in an ice bath. This prevented the hot ammonium hydroxide from overflowing onto the lab bench. After the vial had sufficiently cooled, the solution was transferred using a disposable Pasteur pipette into a disposable chromatography column (found in the RNA purification drawer) and the filtered solution collected in a test tube. The column separates the RNA, which was in solution, from the glass

beads which remained on the column. The screw top vial was then washed with 1.0 ml of distilled water and this resulting wash poured onto the column to wash the beads. After allowing the water to run through the column, the vial was washed with 50% acetonitrile: water and this was placed into the column. Finally, a 1.0 ml aliquot of pure acetonitrile was added to the vial, vortexed briefly, and added to the column. Because the ammonium hydroxide may not have removed all of the silyl protecting groups, the RNA may have differing polarities that the various washes of differing polarities will allow the RNA to pass through the column. All filtrates were collected in the same test tube and vortexed.

The test tube was placed into the Savant Speed-Vacuum concentrator centrifuge (abbr. speedvac) which was attached to a water aspirator in the sink of room C108 to remove the corrosive ammonia without ruining the pump. After one hour had passed, remove the sample form the speedvac and determine its volume. Divide the sample into 3 microcentrifuge tubes (not screw top) and dry with vacuum overnight.

Step 3

To each vial add 10 μl of DMF (dimethylformamide) and 90 μl TEAHF. Vortex for 30 s and place in the 55 ^{o}C water bath for 2 hours. Spin the tubes for 10 s and then add 1.2 ml of n-butanol, vortex for 30s and place in the freezer (-20 ^{o}C) for at least 1 hour. Centrifuge at 12K for 20 min. Check for precipitate, in no precipitate is visible, centrifuge an additional 15 min. Decant the liquid and leave the vial on a paper towel to absorb the excess liquid. Dissolve the precipitate (3 vials) into about 8 ml (total) of 2x Sep Pac buffer and Sep Pak.

Step 4

The 20X buffer is prepared by adding 3.953g (0.1 M) of solid ammonium bicarbonate to 500ml of water. In a 600ml beaker, the ammonium bicarbonate was dissolved in 450ml of deionized water using a magnetic stir bar to mix the solution. Using the electronic pH meter in the lab, the switch was turned from “standby” to “pH” and the pH of the solution was measured. Acetic acid was added to the buffer to adjust the pH to 4.5. As the acid was added, bubbling occurs as carbon dioxide was generated. The buffer was transferred into a 500ml bottle and

diluted to the 500ml mark. This is usually stored in the center shelf of room C108.

The Sep Pak chromatography cartridge removed the salts present in the solution. A Pyrex 6" by 9" glass pan was placed underneath a tall, three-tiered test tube rack. The long end of the Waters C18 Sep Pak chromatography cartridge was attached to a plastic syringe and placed in the test tube rack. The cartridge was flushed with 5 ml of anhydrous acetonitrile (stored in the dessicator in C108). A long neck (9 inch) disposable Pasteur pipet was used to tease the small air bubble out of the neck of the cartridge. A simple "trick" was used to eliminate this bubble every time liquid was placed inside the syringe. After the solution was added to the syringe, the Pasteur pipet was placed into the barrel of the syringe and liquid was drawn into the pipet. Next, the pipet was placed into the neck of the cartridge so the tip of the pipet was penetrating the air bubble. The bulb of the pipet was then compressed forcing liquid into the space occupied by the bubble and expelling the air up and out of the neck of the syringe. The acetonitrile (or any wash) then passed through the column easily. In case of a clogged cartridge, perhaps caused by small particle of silica, a syringe plunger was used to force the liquid through the column. After the acetonitrile had passed through the column, it was washed with 2 ml of 20X Sep Pak buffer, then with three separate 2 ml aliquots of 2X diluted buffer. The diluted buffer was prepared by adding 2.0 ml of 20X buffer to 20 ml of water. After the column had been rinsed with the aforementioned washes, 2 ml of 20X buffer was added to each dried sample and a labeled test tube (#1) was placed underneath the Sep Pak cartridge to collect the filtrate. The sample was then added to the column and allowed to pass through the cartridge. After the filtrate had passed through the cartridge into test tube #1, a second labeled test tube (#2) was placed into the rack behind tube #1 and the syringe was placed on top of the second test tube. The absorbance of the filtrate in tube #1 was measured at 280 nanometers using an UV-vis spectrophotometer blanked with water. A typical value for the first time a Sep Pak is done on a sample for tube #1 is around 0.4. If the value is higher, the filtrate was run through the column a second time and measured again. The RNA should bind to the column and only the inorganic salts should pass through. Two ml of 2X diluted Sep Pak buffer was added to wash the column

and this filtrate was collected in test tube #2. Again, this filtrate was measured to be sure all RNA had bound to the column and the absorbance for this tube should be near zero, perhaps 0.15. Finally, a third test tube, appropriately labeled #3, was placed into the rack and the cartridge placed above it. A solution of 30% acetonitrile: 70% water by volume was prepared and 3 ml of this was used to wash the RNA off of the column and into test tube #3. The measured absorbance of the liquid in this tube should be the highest as the acetonitrile: water wash flushed the RNA off of the column. A good purification will result in an absorbance of about 4.5 in the last tube (tube #3). Tube #3 was placed into the speedvac to dry overnight.

Step 5

A preparative TLC was used to separate the desired RNA from the sequence failures. After the RNA had dried in the speedvacuum, it was resuspended in 150 microliters of distilled water. A 500 micrometer thickness Whatman TLC plate (20cm X 20cm) was prepared by drawing a light pencil line using a ruler one inch from the edge of the plate. A disposable 25 microliter pipet was used to streak a thin line of liquid across the pencil line on the plate. An effective way to streak the plate is as follows: first place the pipet into the test tube and allow capillary action to draw the liquid into the pipet. Holding the pipet like a pencil in the left hand nearest the edge containing the liquid, the pipet should extend approximately 1 mm below the middle finger. The pipet was then lowered down to the right most edge of the plate completely perpendicular to the plate and the user's middle finger should be resting on the plate with the pipet remaining about 1 mm from the surface of the plate. The pipet was then lowered until it barely grazed the surface of the plate and the hand moves quickly right to left across the plate. The left middle finger glides along the plate to steady the pipet as a thin line of liquid is drawn onto the plate. This was repeated until all of the sample had been added to the plate. The test tube may be washed with a 20 microliter aliquot of water which is also added to the plate.

After allowing the RNA to dry, a UV lamp was used to visualize the band and ensure that a desirable band has been streaked across the surface of the plate. The dried plate was then placed into a clean developing chamber which contained approximately 100ml of freshly

prepared 55:35:10 (V:V:V) N-propanol, ammonium hydroxide, water. A piece of filter paper slightly smaller than the size of the plate was also placed inside the chamber to increase the gas-liquid equilibrium process. The tank was capped and a heavy object was placed on top of the cap to prevent evaporation. The mobile phase solvent was allowed to migrate to the top edge of the plate which took about five hours. After solvent migration, the plate was removed from the tank and placed within the fume hood to dry thoroughly. After the solvent had evaporated, the UV lamp was used to visualize the RNA bands that appeared on the plate. Careful analysis revealed several bands migrated to various heights above the origin. The higher bands on the plate, which were the smaller sequence failures, moved faster up the plate while the larger oligomers migrated at a slower rate up the plate. The desired RNA was the darkest in intensity while being the lowest band on the plate. This band of interest was then outlined in pencil. While working over a sheet of weighing paper, the band of silica containing RNA was gently scraped from the plate with a spatula onto the paper. The silica was poured from the paper into a disposable sterile centrifuge tube containing 2 ml of distilled water.

The resulting slurry was capped, shaken and vortexed for 3 minutes. Next, the tube was placed on the lab bench and left undisturbed for a period of ten minutes. This allotted amount of time ensured that the RNA had dissolved in the water. The tube was then placed into the centrifuge with a blank to balance the opposite position, and centrifuged for five minutes. After allowing the centrifuge to come to a complete stop, the supernatant solution containing the water and RNA was slowly decanted into another labeled plastic centrifuge tube. The first plastic tube containing the silica was washed with an additional 2mls of distilled water and this process was repeated one additional time for a total volume 6 ml of liquid in the second plastic tube.

The combined supernatant solutions were centrifuged and decanted a final time into a glass test tube to remove any particles of silica which will hinder the Sep Pak process.

Step 6

2 ml of 20X Sep Pak buffer was added to the sample and the procedure was the same as before with the RNA coming off the column into tube #3. The absorbance of each tube was

measured at 260 nm on every subsequent Sep pak following the first which is measured at 280nm. This tube was then placed into the speedvac and dried overnight. After drying, the sample was resuspended in 125 microliters of water and an analytical TLC plate was prepared.

Step 7

A small section of 250 micrometer TLC plate was cut (3 cm X10 cm) using a glass cutter and a light pencil line drawn 1cm from the bottom. The "trick" to cutting glass is to hold the glass cutter perpendicular to the sheet of glass when cutting. The suspended RNA was placed onto the plate as a single spot using a 5 microliter disposable pipette. After drying, the spot was visualized under UV light. A small jar with cap was used as the developing chamber. A small piece of filter paper was placed inside the chamber with the solvent 55:35:10 N-propanol, ammonium hydroxide, water and the solvent was allowed to reach gas-liquid equilibrium. The plate was placed inside the chamber and tightly capped. After approximately 20 minutes the solvent had migrated to within one cm of the top of the plate. The plate was removed and the solvent was allowed to evaporate under the fume hood. Visualization under the UV lamp allowed the purity to be qualitatively measured, and only one band should be present. If the analytical TLC shows impurities, the Preparative TLC was conducted again with repetition of the previous purification steps.

Once the purity had been determined, the desired RNA sequence was transferred into a small, sterile, plastic microcentrifuge tube and dried in the speedvac. The dry RNA was usually resuspended in 40 microliters of melt buffer.

Using the Lab Equipment:

Washing Glassware

Glassware was washed in distilled water with a small amount of soap, rinsed in distilled water, and placed upside-down in the drying rack. When dry, it was wrapped in aluminum foil and baked for three hours.

Changing the Vacuum Pump Oil

The waste pump oil container is stored under the sink in the Bio Organic prep. room. The exhaust hose, which is attached to the underside of the table, was unscrewed by first removing the metal bracket which holds the hose upright and then unscrewing the hose at the metal adapter (without disconnecting the hose from the metal adapter). The pump was placed on top of an inverted plastic bucket. The rear plug near the bottom of the pump was turned to open the valve that allowed the oil to drain through a small funnel into the waste container. Then the topmost plug was removed and the equalization of pressure allowed more oil to drain from the pump. The pump was tilted to drain as much of the old oil as possible. The lower plug was closed and fresh oil was poured into the upper hole while monitoring the oil level gauge on the back of the pump. The oil level should lie near the top portion of the gauge. Lastly, the upper plug was reinserted and the hoses were checked to ensure that they have been properly tightened and the metal bracket replaced to maintain the exhaust condensation column in a vertical position. In order to function properly, the dessicator attachment must be kept in an upright position.

Cleaning the Speed Vacuum Condensation Trap

The hoses were disconnected so the liquid inside the trap began to thaw. A pipet was used to withdraw and remove the melted liquid from the trap. The solvents in the trap were washed down the drain with copious amounts of water. Lastly, the hoses were reconnected.

Autoclaving Micropipet Tips

The autoclave is located in room B209. Both switches on the machine were turned on simultaneously (power and control). If the autoclave alarm were to go off, the “reset” button was selected and the tape was allowed to dispense. To the materials being autoclaved, small amounts of autoclave tape were added to random places. The materials were placed inside the autoclave allowing ample room for the steam to circulate between items. The door was closed and the wheel was turned clockwise until snug. The “gravity” option was selected with 30 minutes sterilize time and 10 minutes dry time, and the temperature increased as the cycle started. After the temperature drops, the door was opened slightly to allow small amounts of steam to escape

slowly. In the drawer with the autoclave tape (next to the autoclave) gloves are stored to handle the hot materials when the cycle has finished. If liquids are to be autoclaved, capped containers must be completely unscrewed with the lids resting loosely on their containers. The “liquids” option was selected instead of gravity.

The old pressure autoclave is kept in the room adjacent to the cold room and is used for small amounts of materials to be autoclaved. To use this autoclave the inner chamber was removed and the outer chamber filled partly with water. The inner chamber was reinserted and the lid well tightened. The needle on the lid was closed in the upright position. The timer was set for 50 minutes and the pressure allowed to increase to between 10 and 20 on the gauge. After autoclaving, the needle was knocked down with a ruler to cause the hot steam to escape.

Using the Speed Vacuum Concentrator

When turning on the Speedvac, the rotor is turned on first, then the pump is turned on and finally the pressure switch is closed. When turning off the speedvac, the order is reversed with the rotor being turned off last. When using heat to speed the drying process, the heater must be switched off overnight.

Using the Beckman Spectrophotometer to measure absorbance of samples

The quit icon was clicked on the diagnostic screen and “Fixed wavelength” was selected in the menu screen. The wavelength was selected by clicking on the current wavelength and typing in the desired one. The “blank” icon near the bottom of the screen was clicked to read the blank (cell one). A single right click on the mouse read the absorbance of the sample. Lastly, “quit without saving” was chosen to exit.

Using the Perkin Elmer Spectrophotometer to read absorbance

The power switch, located in the rear right of the machine, was toggled. Using the labeled Quartz cuvettes, one was filled with water to blank. This cuvette was placed into the first well position (closest rightmost position in cell block with the frosty side facing the operator). The “GO TO” (wavelength), button was pressed and “280” or “260” was typed into the machine at the prompt and the enter key was pressed. Finally, the “Back correct button” was pressed to

record the blank. The desired sample was placed into this well to read the absorbance.

Using the Beckman Spectrophotometer for melts

The power switch in the rear right of the machine was turned on and one-half hour was allowed for the UV lamp to warm up. During this warm-up period, quartz cells were washed five times with water then washed five times with methanol. The cells were dried using a stream of air. The small cells required a needle to assist in blowing out the methanol. The cells were washed both before and after each use. The temperature housing has six spaces for cells. The first cell designation is the space nearest the electric line. The first cell was also designated as the blank which is filled with melt buffer. The samples were pipetted into the cells using a micropipet and buffer was added to raise the level of liquid to within 2mm from the top of the cell. The samples were measured for an absorbance between 2 and 0.2 at 280nm. If the measured absorbance was too high, melt buffer was added to dilute the concentration. Typically, the 0.1 cm cells were used with an appropriate RNA concentration to obtain an absorbance near 2. Usually, cells 2 and 5 contained the RNA sample while all other cells (1cm path) contained buffer. Once the samples were prepared for melting, and the absorbance was recorded in the notebook, a lid was screwed into the heating housing to prevent evaporation. The entire unit was then thumbscrewed into the spectrophotometer and the heating line was plugged in at the rear left-hand side the machine. The desired wavelength at which the melt will proceed was chosen under the “custom applications” menu. On the left-hand side of the spec, a separate heater unit has a toggle switch in the rear that must be switched on in order to heat the cells. Next, “temp control view” was clicked and the temperatures were adjusted as follows: “Start temp” is the temperature at which the machine begins to take readings and this was set to a value usually between 75 and 95 degrees. “End Temp” is the last temperature at which the machine takes a reading and this was set at a value between 25 and 5 degrees. “Final temp” is the temperature of the cells at the conclusion of the experiment and this was adjusted to “ambient.” A three minute delay was specified to allow the sample to come to the desired temperature and be uniformly heated while the “Ramp Rate” was set with 1 degree per minute with a “Read

Interval" of usually 0.2 min. This screen was exited and the blank was measured a final time. The right mouse button was clicked and the temperature increased. At the desired start temperature, the "start" icon was selected which began the melt (note that the specified delay time starts when this icon is selected). At the conclusion of the melt, the file was saved in a 3.5 floppy disk directory headed "TM_data." The "save clear" icon at the top of the screen was clicked and "B" drive on the spec was selected. The data was saved and the heater was turned off.

The cell holder was removed and the concentration of RNA in each cell was adjusted by removing 3 and 6 microliters into each 1 cm cell to the left and right of the 0.1 cm cell respectively. For a good melt, the RNA had to be well mixed within the cells. This was accomplished by slowly drawing up the liquid within a micropipet and quickly expelling it into the cell repeatedly. The heating unit was replaced and the entire process repeated. Upon the third repetition, the amounts removed from the 0.1 cm cells were 5 and 10 microliters, and the last repetition removed 10 and 20 microliters. These dilution guidelines should produce the proper concentration range in just four melts.

Following the final melt, the cells were cleaned. Before turning off the machine, "File utilities" icon was selected after quitting the melt screen. The spectrophotometer takes information in "lotus" format which must be converted into "ascii" for the Meltwin program. The data was converted by selecting "B_TM_data" and the file was highlighted. The "convert" option was selected at the top of the screen and the destination was specified to be the "B" drive. The format of the file was changed from lotus to "ascii" and the "OK" icon was clicked. This menu screen was exited and the power switch turned off.

Using the Meltwin Program

After opening the program, "File" was chosen and "import data" was selected. The files were chosen under "A:/ASCII" and under file of type: "Beckman files.asc" were chosen. The number of cells were entered and the wavelength at which the absorbance was measured was chosen after entering the cell information. The sequence type was specified and the melt data displayed.

At this point the file was saved in the temporary files of the computer in case the program shuts down during the fit process. The area of the melt can be customized by moving the left and right mouse buttons. The "Fit Wizard" icon was selected at the top of the screen and the selected cells were chosen. The blank was selected as the background and the desired secondary structure "hairpin," for example, was chosen which allowed the sequence to be entered. After completion of the fit, the "analyze data" icon was clicked and the results of the fit were displayed. Generally, the value of X^2 was desired to be in the 10^{-8} range. Finally the file was saved under melt files (.dat) a final time.

Bimolecular structure data was analyzed using the Meltwin program by importing the results of one fit and adding additional fits. By selecting the "analyze data" icon and selecting the necessary data points, the program will derive the thermodynamic parameters for a bimolecular structure.

Quantifying RNA

Using Beer's Law, the concentration of RNA can be determined. The RNA calc program was used to determine the molar absorptivity of the sequence. The RNA was first dried down completely and resuspended in a known quantity of water. An aliquot was removed and quantitatively diluted to give an absorbance reading between 0.2 and 2. This absorbance was corrected for the dilution factor, and using the molar absorptivity, the RNA concentration was determined.

Appendix D: Raw Optical Melt Data

CC 1-4 GC ACU**A**CG
BK 9 and BK 10 CG**A**UGA GC

Full Analysis

Date: 7 Apr 2013 Comment: Imported From Beckman DU Series 600

Melt Curve Results

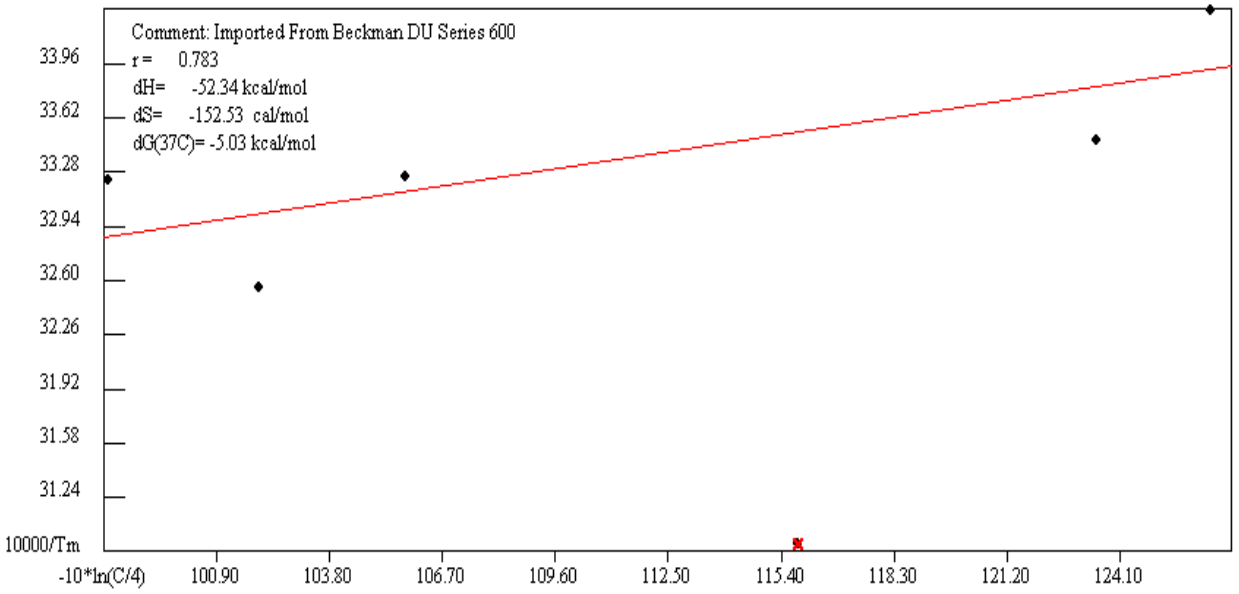
Concentration (M)	dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C)
2.190e-04	-33.59	-92.12	-5.02	27.75
1.488e-04	-64.73	-190.45	-5.66	34.03
1.022e-04	-44.63	-127.39	-5.12	27.57
1.294e-05	-36.35	-99.54	-5.47	18.41
~~3.729e-05~~	~~-17.33~~	~~-30.60~~	~~-7.84~~	~~50.01~~
1.731e-05	-72.61	-218.52	-4.83	25.56
Averages:	-50.38	-145.60	-5.22	26.67

Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-50.38 ± 17.407	-145.60 ± 56.220	-5.22 ± 0.340	29.1

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-52.34 ± 23.993	-152.53 ± 78.153	-5.03 ± 1.970	28.4



ΔH^0 (kcal/mol)	ΔG^0_{37} (kcal/mol)	Value
52.33	5.03	Measured
66.64	10.47	Parent
14.3	5.4	Effect of 2 Bulges

CC 5-8 GC**A**GGCC GC
BI 42 CG CCGG**A**CG

Full Analysis

Date: 7 Apr 2013 Comment: Imported From Beckman DU Series 600

Melt Curve Results

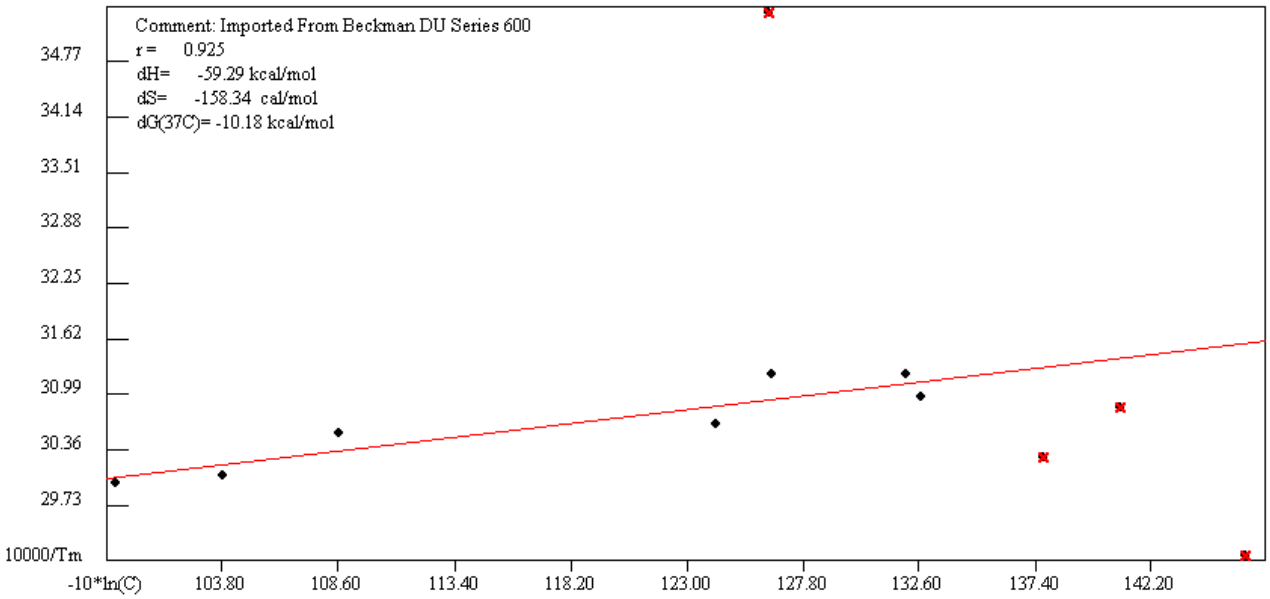
Concentration (M)	dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C)
4.822e-05	-54.37	-143.32	-9.92	60.27
~~1.042e-06~~	~~-46.48~~	~~-113.38~~	~~-11.31~~	~~57.06~~
~~4.523e-07~~	~~-10.62~~	~~-1.94~~	~~-10.02~~	~~69.80~~
3.094e-05	-53.41	-139.95	-10.01	59.47
1.837e-06	-42.50	-106.44	-9.49	47.15
~~7.581e-07~~	~~-35.42~~	~~-81.22~~	~~-10.22~~	~~51.09~~
1.920e-05	-49.15	-128.60	-9.26	54.11
3.194e-06	-45.62	-117.25	-9.26	47.24
1.724e-06	-54.05	-141.02	-10.32	49.77
~~3.234e-06~~	~~-21.90~~	~~-52.22~~	~~-5.70~~	~~-9.98~~
4.020e-06	-53.87	-140.46	-10.31	53.06
Averages:	-50.43	-131.01	-9.79	53.01

Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-50.43 ± 4.780	-131.01 ± 14.258	-9.79 ± 0.459	64.6

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-59.29 ± 10.889	-158.34 ± 33.135	-10.18 ± 0.807	62.5



ΔH^0 (kcal/mol)	ΔG^0_{37} (kcal/mol)	Value
59.30	10.18	Measured
83.00	16.20	Parent
23.70	6.02	Effect of 2 Bulges

CC 9-12 A GUC**A**UCA CUG
BL7 and BL8 CAG AGU**A**GAC

Full Analysis

Date: 7 Apr 2013 Comment: Imported From Beckman DU Series 600

Melt Curve Results

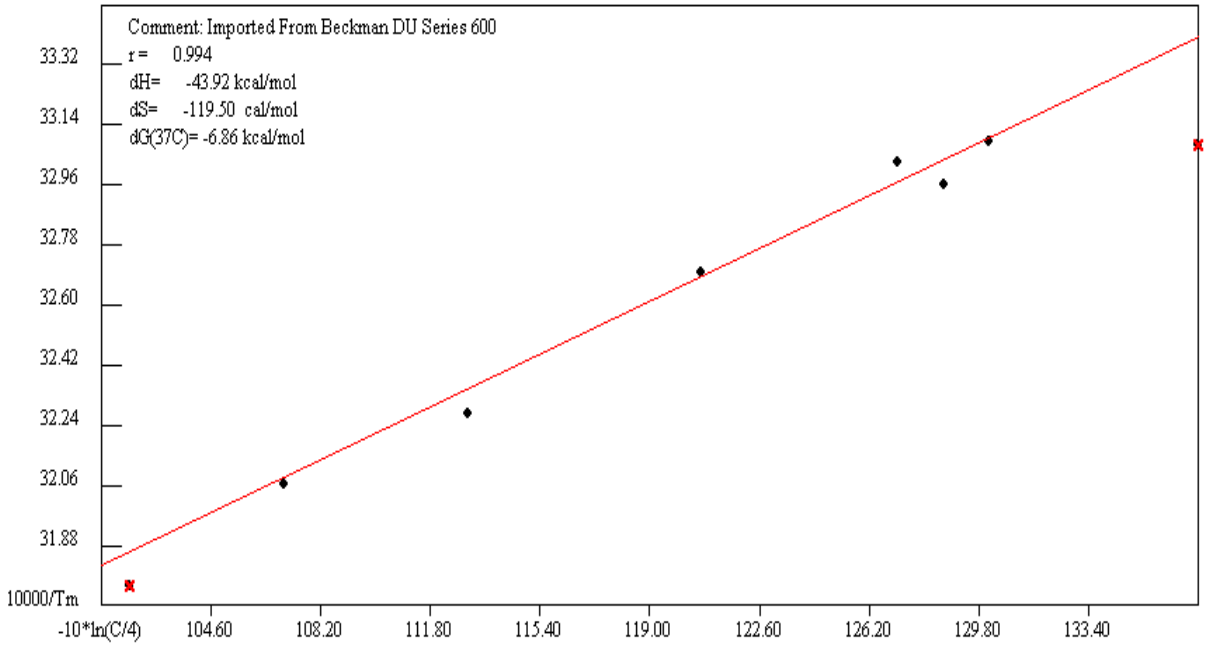
Concentration (M)	dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C)
~~1.496e-04~~	~~-86.51~~	~~-254.50~~	~~-7.58~~	~~41.71~~
9.030e-05	-57.40	-162.80	-6.91	38.72
8.921e-06	-62.02	-179.37	-6.39	29.05
4.937e-05	-58.08	-165.02	-6.90	36.67
1.037e-05	-60.31	-173.23	-6.58	30.22
~~4.500e-06~~	~~-74.73~~	~~-219.96~~	~~-6.51~~	~~29.17~~
2.294e-05	-43.61	-118.59	-6.83	32.70
1.203e-05	-61.08	-176.49	-6.35	29.62
Averages:	-57.08	-162.58	-6.66	32.83

Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-57.08 ± 6.833	-162.58 ± 22.491	-6.66 ± 0.256	37.7

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-43.92 ± 2.508	-119.50 ± 8.195	-6.86 ± 0.059	39.3



ΔH^0 (kcal/mol)	ΔG^0_{37} (kcal/mol)	Value
43.91	6.86	Measured
85.91	13.47	Parent
42.00	6.6	Effect of 2 Bulges

CC 13-20A and 29-32A CG**A**UGA GC
BN1 and B010; BN1 and BL1 GC ACU**A**CG

Full Analysis

Date: 7 Apr 2013 Comment: Imported From Beckman DU Series 600

Melt Curve Results

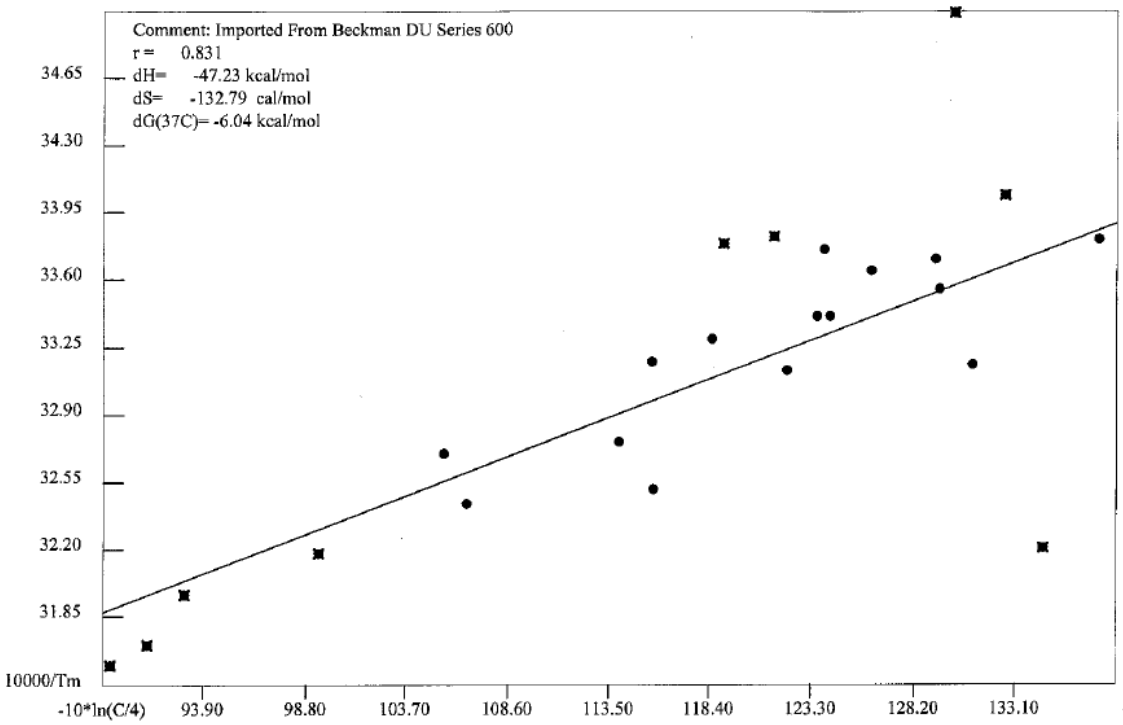
Concentration (M)	dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C)
~~8.893e-06~~	~~-26.94~~	~~-68.39~~	~~-5.73~~	~~12.66~~
~~6.951e-06~~	~~-51.71~~	~~-149.64~~	~~-5.30~~	~~20.65~~
~~2.117e-05~~	~~-43.23~~	~~-122.07~~	~~-5.37~~	~~22.50~~
4.392e-06	-41.16	-111.91	-6.46	22.63
~~2.704e-05~~	~~-29.14~~	~~-74.80~~	~~-5.94~~	~~22.81~~
1.671e-05	-51.94	-150.73	-5.19	23.07
9.702e-06	-55.80	-162.39	-5.43	23.52
1.323e-05	-44.39	-124.30	-5.84	24.02
9.522e-06	-57.37	-166.74	-5.66	24.92
1.626e-05	-50.04	-142.55	-5.83	26.14
1.725e-05	-46.74	-131.61	-5.92	26.18
2.861e-05	-63.66	-188.38	-5.23	27.22
3.806e-05	-47.47	-134.48	-5.76	28.33
8.101e-06	-52.16	-146.89	-6.60	28.47
1.992e-05	-41.86	-114.40	-6.38	28.72
4.459e-05	-62.92	-183.44	-6.02	32.12
1.048e-04	-57.77	-167.93	-5.69	32.71
3.791e-05	-70.90	-207.53	-6.53	34.41
9.364e-05	-55.01	-157.28	-6.24	35.12
~~5.764e-06~~	~~-14.14~~	~~-18.81~~	~~-8.30~~	~~37.34~~
~~1.914e-04~~	~~-88.40~~	~~-264.70~~	~~-6.30~~	~~37.60~~
~~3.657e-04~~	~~-52.55~~	~~-149.47~~	~~-6.19~~	~~39.72~~
~~4.394e-04~~	~~-98.04~~	~~-283.22~~	~~-7.20~~	~~42.23~~
~~5.246e-04~~	~~-81.63~~	~~-240.16~~	~~-7.14~~	~~43.33~~
Averages:	-53.28	-152.70	-5.92	27.84

Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-53.28 ± 8.466	-152.70 ± 27.547	-5.92 ± 0.452	33.5

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-47.23 ± 8.781	-132.79 ± 28.539	-6.04 ± 0.778	33.8



ΔH^0 (kcal/mol)	ΔG^0_{37} (kcal/mol)	Value
47.23	6.04	Measured
66.67	10.47	Parent
19.44	4.43	Effect of 2 Bulges

CC13-16B GC**A**UGCGCA GC
BL 23 CG ACGCGU**A**CG

Full Analysis

Date: 7 Apr 2013 Comment: Imported From Beckman DU Series 600

Melt Curve Results

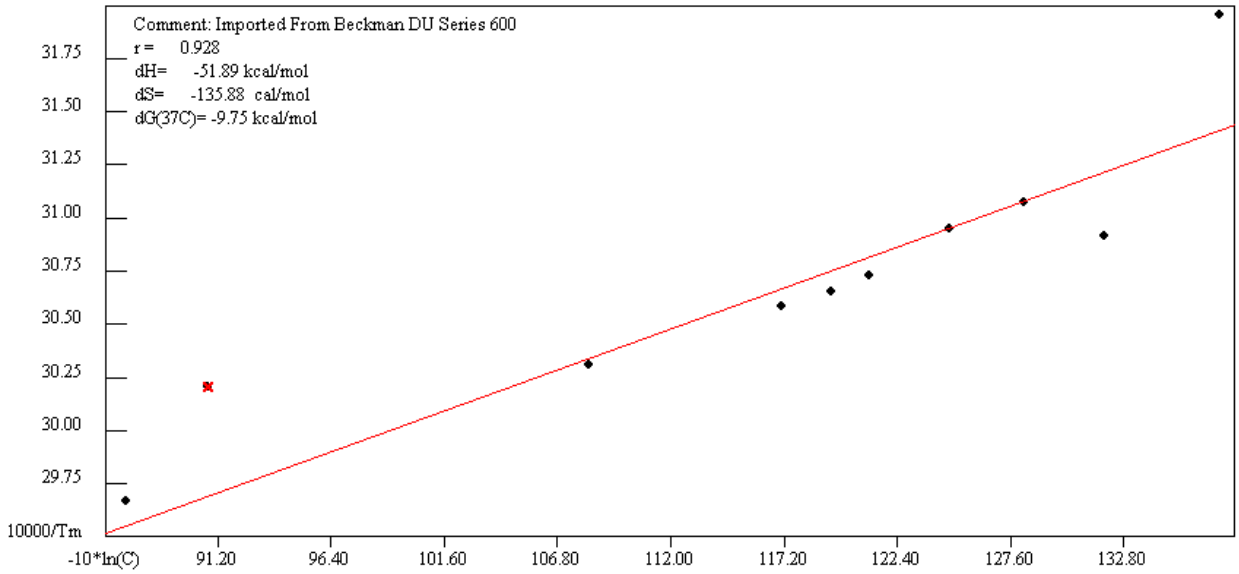
Concentration (M)	dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C)
~~1.146e-04~~	~~-44.72~~	~~-117.06~~	~~-8.42~~	~~57.90~~
1.095e-06	-56.96	-154.75	-8.97	39.80
1.996e-05	-47.32	-121.92	-9.51	56.77
3.796e-06	-51.41	-134.31	-9.75	49.95
2.696e-06	-51.10	-133.32	-9.75	48.65
6.532e-06	-50.60	-131.40	-9.85	53.05
8.213e-06	-54.34	-142.92	-10.01	53.83
5.487e-06	-56.21	-148.63	-10.11	52.31
1.674e-04	-61.73	-165.89	-10.28	63.87
1.863e-06	-52.72	-136.75	-10.30	50.32
Averages:	-53.60	-141.10	-9.84	52.06

Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-53.60 ± 4.258	-141.10 ± 13.455	-9.84 ± 0.418	63.1

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-51.89 ± 7.895	-135.88 ± 24.086	-9.75 ± 0.607	63.4



ΔH^0 (kcal/mol)	ΔG^0_{37} (kcal/mol)	Value
51.89	9.75	Measured
108.39	19.90	Parent
56.5	10.15	Effect of 2 Bulges

CC17-20B GC**A**UGCGU**A**CG
BL3 and BL 17 CG ACGCA GC

Full Analysis

Date: 7 Apr 2013 Comment: Imported From Beckman DU Series 600

Melt Curve Results

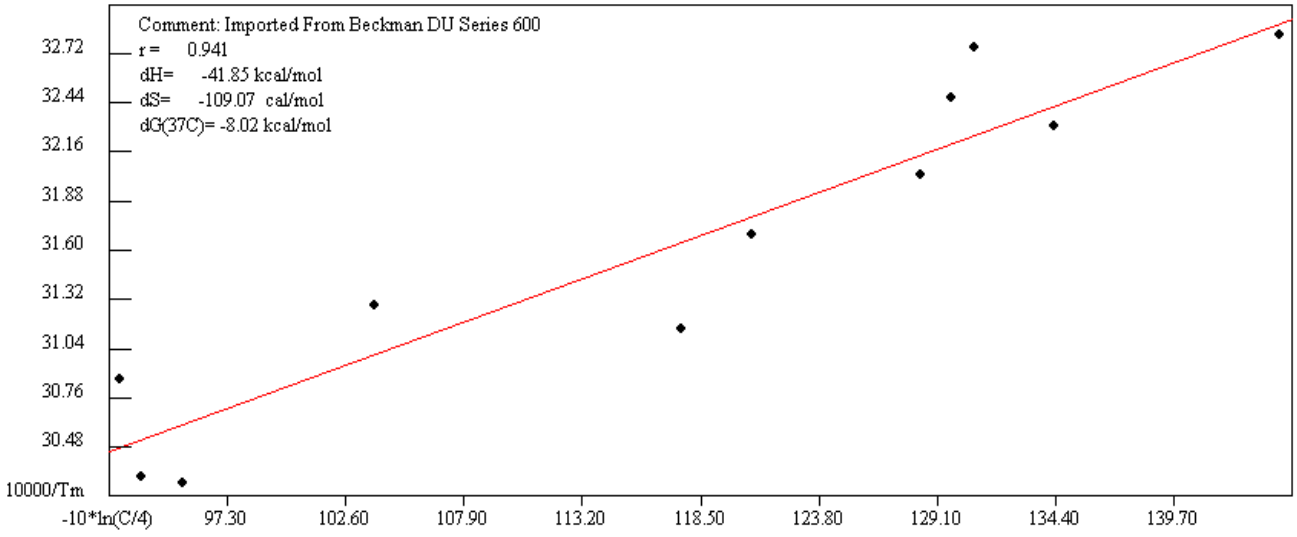
Concentration (M)	dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C)
3.493e-04	-41.37	-106.84	-8.23	56.72
5.871e-06	-27.20	-61.17	-8.22	36.39
2.132e-06	-27.76	-62.43	-8.40	31.47
2.892e-04	-59.53	-161.29	-9.51	57.15
1.067e-05	-29.46	-68.85	-8.10	39.06
9.281e-06	-35.82	-90.51	-7.74	34.83
3.833e-04	-33.22	-84.17	-7.12	50.79
3.116e-05	-39.08	-98.39	-8.57	47.82
8.368e-06	-23.40	-50.63	-7.69	32.19
1.229e-04	-27.78	-66.26	-7.23	46.50
2.275e-05	-28.28	-65.62	-7.93	42.39
Averages:	-33.90	-83.29	-8.07	43.21

Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-33.90 ± 10.127	-83.29 ± 31.178	-8.07 ± 0.660	51.7

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-41.85 ± 5.012	-109.07 ± 15.747	-8.02 ± 0.326	48.5



ΔH^0 (kcal/mol)	ΔG^0_{37} (kcal/mol)	Value
41.85	8.02	Measured
92.19	16.25	Parent
46.34	8.23	Effect of 2 Bulges

CC21-24A and CC33-36A CAG G**A**CUG
BM1 and BM2U GUC**A**C GAC

Full Analysis

Date: 7 Apr 2013 Comment: Imported From Beckman DU Series 600

Melt Curve Results

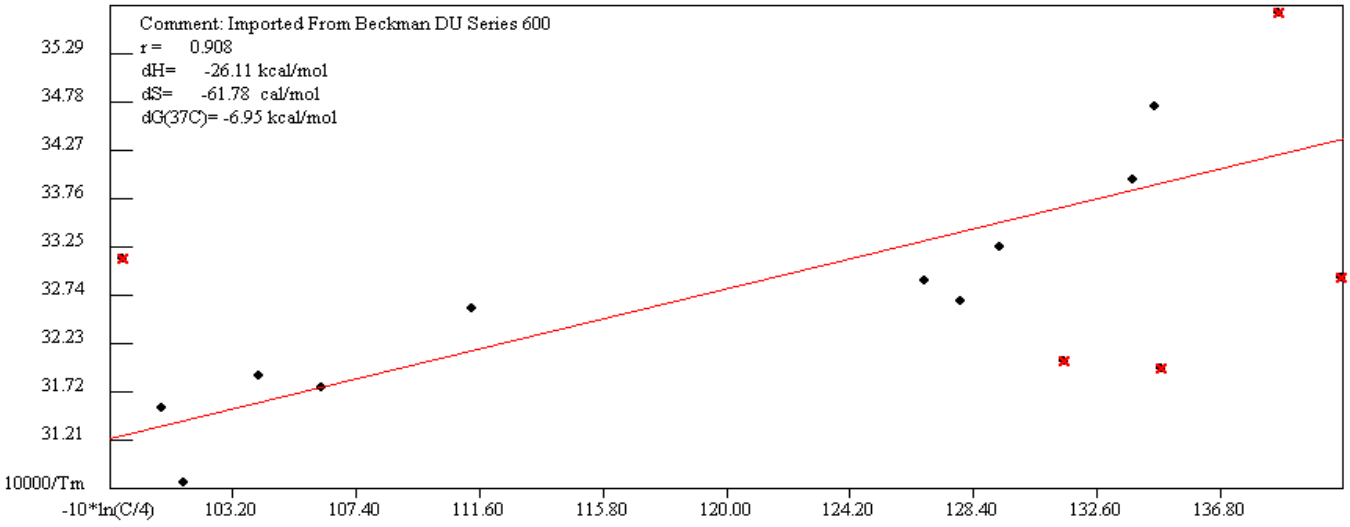
Concentration (M)	dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C)
~~1.923e-04~~	~~-39.68~~	~~-111.66~~	~~-5.05~~	~~28.78~~
5.850e-05	-32.64	-84.29	-6.50	33.57
5.720e-06	-22.60	-51.74	-6.56	14.86
1.679e-04	-19.31	-40.93	-6.62	43.69
1.207e-04	-22.08	-49.74	-6.65	40.34
9.771e-05	-28.97	-70.94	-6.97	41.57
9.678e-06	-26.37	-61.98	-7.15	27.60
1.554e-04	-23.73	-52.85	-7.34	51.81
6.179e-06	-17.03	-31.22	-7.34	21.35
~~3.748e-06~~	~~-10.80~~	~~-10.99~~	~~-7.39~~	~~-6.88~~
1.256e-05	-13.42	-18.96	-7.54	30.79
1.106e-05	-13.84	-19.81	-7.70	32.86
~~3.038e-06~~	~~-22.62~~	~~-46.50~~	~~-8.20~~	~~30.55~~
~~7.778e-06~~	~~-35.60~~	~~-87.98~~	~~-8.32~~	~~38.86~~
~~5.611e-06~~	~~-42.85~~	~~-110.26~~	~~-8.66~~	~~39.57~~
Averages:	-22.00	-48.25	-7.04	33.85

Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-22.00 ± 6.292	-48.25 ± 21.208	-7.04 ± 0.440	44.3

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-26.11 ± 4.263	-61.78 ± 13.667	-6.95 ± 0.502	42.1



ΔH^0 (kcal/mol)	ΔG^0_{37} (kcal/mol)	Value
26.11	6.95	Measured
66.50	10.97	Parent
40.39	4.02	Effect of 2 Bulges

CC21-24B CG GAUGCA GC
BO11 and BN3 CG**A**CUACGU**A**CG

Full Analysis

Date: 7 Apr 2013 Comment: Imported From Beckman DU Series 600

Melt Curve Results

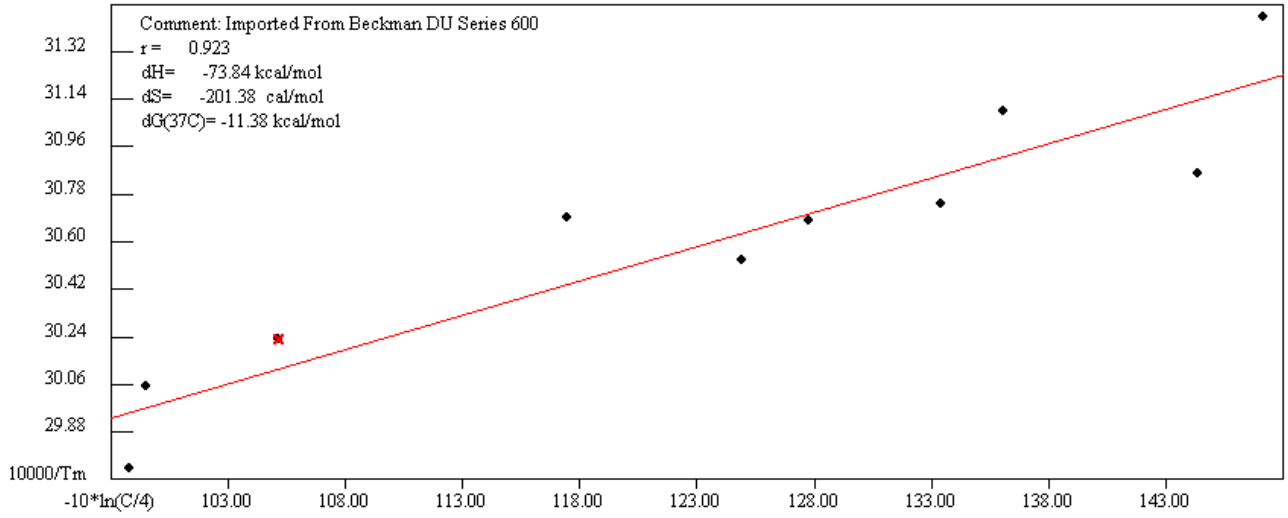
Concentration (M)	dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C)
2.044e-04	-52.35	-136.06	-10.15	63.07
2.150e-06	-44.56	-108.80	-10.81	50.93
1.635e-06	-43.35	-107.11	-10.13	44.82
1.908e-04	-50.21	-131.15	-9.54	59.56
4.930e-06	-49.19	-125.94	-10.13	48.42
~~1.087e-04~~	~~-79.64~~	~~-219.87~~	~~-11.44~~	~~57.62~~
1.132e-05	-46.50	-117.27	-10.13	52.79
6.448e-06	-52.83	-135.90	-10.68	52.14
3.173e-05	-50.34	-131.17	-9.65	52.64
1.503e-05	-43.35	-107.51	-10.00	54.42
Averages:	-48.07	-122.32	-10.14	53.20

Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-48.07 ± 3.730	-122.32 ± 12.256	-10.14 ± 0.412	62.1

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-73.84 ± 11.602	-201.38 ± 35.322	-11.38 ± 0.808	58.8



ΔH^0 (kcal/mol)	ΔG^0_{37} (kcal/mol)	Value
73.84	11.38	Measured
103.36	18.12	Parent
29.52	6.74	Effect of 2 Bulges

CC25-28A GUC**A**UCA CUG
BL7 and BN5 CAG AGU**A**GAC

Full Analysis

Date: 7 Apr 2013 Comment: Imported From Beckman DU Series 600

Melt Curve Results

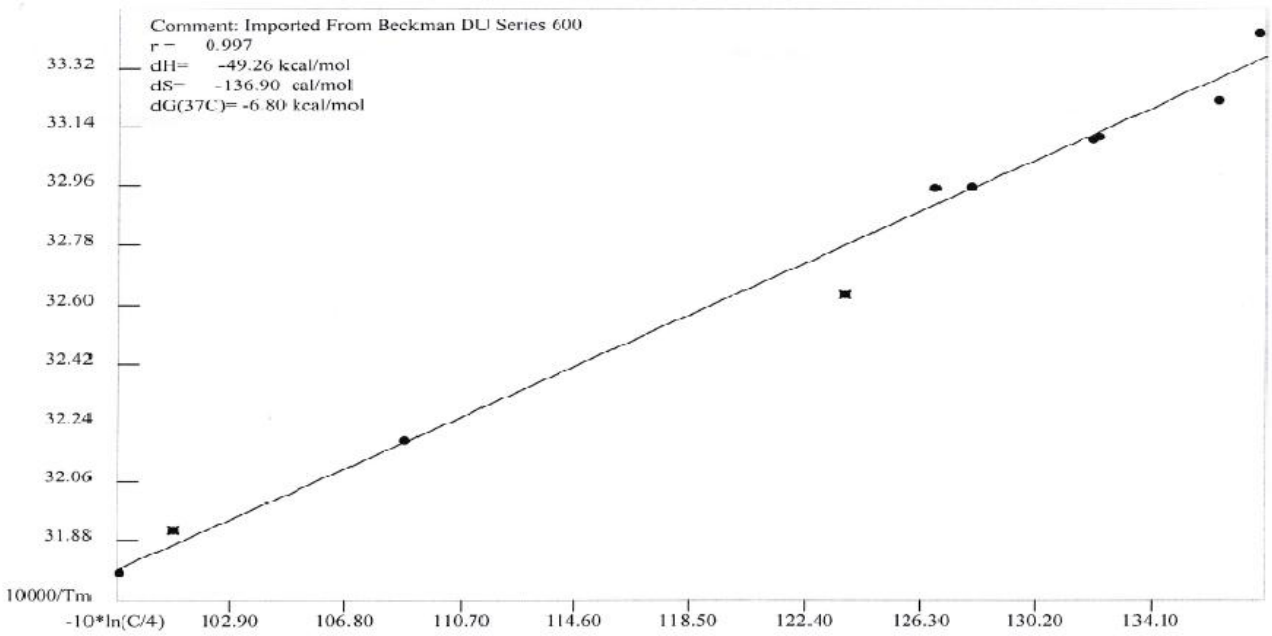
Concentration (M)	dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C)
1.980e-04	-51.61	-144.34	-6.85	41.51
4.139e-06	-47.68	-131.96	-6.75	26.05
~~1.649e-04~~	~~-93.25~~	~~-277.54~~	~~-7.17~~	~~40.20~~
7.188e-06	-62.81	-181.66	-6.47	28.91
4.784e-06	-60.22	-172.93	-6.59	27.92
7.520e-05	-54.97	-155.31	-6.81	37.56
1.106e-05	-59.04	-169.13	-6.59	30.31
7.332e-06	-59.00	-169.01	-6.58	28.98
~~1.687e-05~~	~~-87.56~~	~~-261.10~~	~~-6.58~~	~~33.33~~
1.251e-05	-53.97	-152.63	-6.63	30.34
Averages:	-56.16	-159.62	-6.66	31.45

Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-56.16 ± 5.014	-159.62 ± 16.508	-6.66 ± 0.130	37.7

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-49.26 ± 1.633	-136.91 ± 5.364	-6.80 ± 0.037	38.7



ΔH^0 (kcal/mol)	ΔG^0_{37} (kcal/mol)	Value
49.26	6.80	Measured
85.91	13.47	Parent
36.65	6.67	Effect of 2 Bulges

CC29-32B GUC ACU**A**CUG
BN2 and BL5 CAG**A**UGA GAC

Full Analysis

Date: 7 Apr 2013 Comment: Imported From Beckman DU Series 600

Melt Curve Results

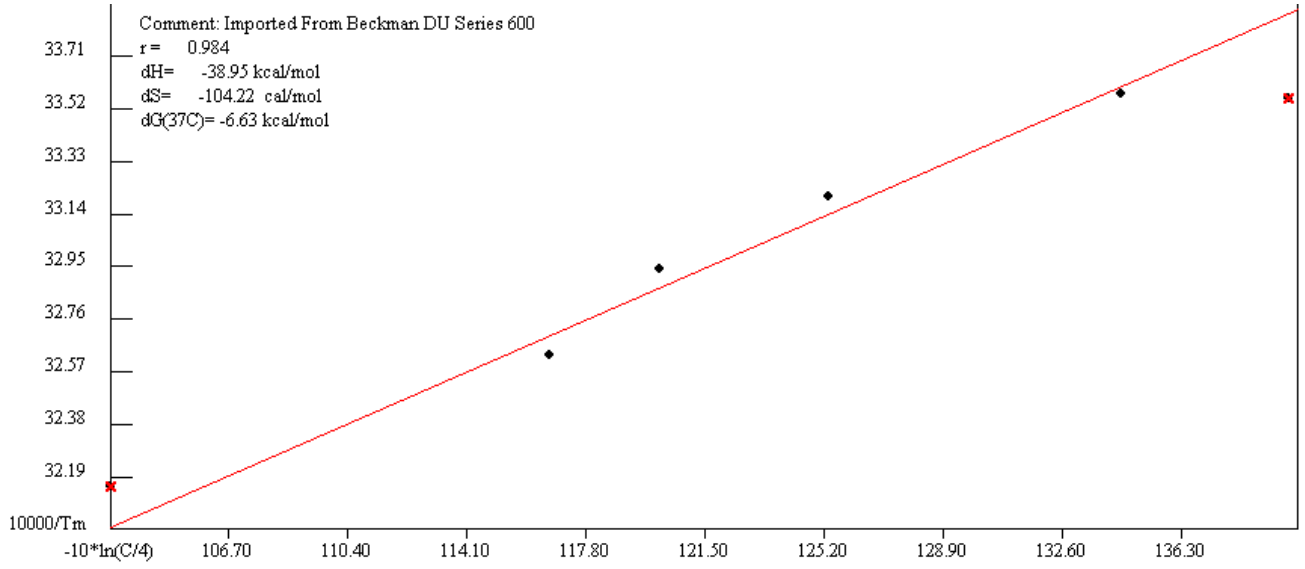
Concentration (M)	dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C)
5.804e-06	-44.48	-122.61	-6.45	24.70
~~3.439e-06~~	~~-22.13~~	~~-46.50~~	~~-7.71~~	~~24.85~~
1.443e-05	-34.28	-88.92	-6.70	28.02
2.440e-05	-34.09	-88.43	-6.66	30.40
~~1.343e-04~~	~~-129.20~~	~~-394.94~~	~~-6.71~~	~~37.86~~
3.425e-05	-51.36	-144.41	-6.58	33.34
Averages:	-41.05	-111.09	-6.60	29.11

Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-41.05 ± 8.416	-111.09 ± 27.371	-6.60 ± 0.111	37.5

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-38.95 ± 4.933	-104.22 ± 16.286	-6.63 ± 0.206	37.8



ΔH^0 (kcal/mol)	ΔG^0_{37} (kcal/mol)	Value
38.95	6.63	Measured
85.91	13.47	Parent
46.96	6.84	Effect of 2 Bulges

Appendix E: Sample Free Energy Calculations

A. Sample Calculation to Predict the Thermodynamic Stability of the Parent Strand and of the Bulge Using the Average Method

Sequence:

$$^{5'}GC\ \ ACUACG^{3'}$$
$$^{3'}CGAUGA\ \ GC^{5'}$$

1. Using the INN-HB model (Xia et al., 1998) and the duplex calculator the thermodynamic stability of the parent duplex without any bulges can be determined.

Parameters	ΔG^0_{37} (kcal/mol)
CU GA	-2.08
CA GU	-2.11
GU CA	-2.24
GA CU	-2.35
CG GC	-2.36
GC CG	-3.42
Initiation	+4.09
ΔG^0_{37} = -10.47 kcal/mol	(-2.08)+(-2.11)+(-2.24)+(-2.35)+(-2.36)+(-3.42)+4.09

2. The average ΔG^0_{37} of the bulge can be determined using the data that was collected from the Melt Curve fit and the T_m dependence linear regression. Using the values for the T_m dependence graph the calculations can be performed:

Analysis of Melt Curve: ΔG^0_{37} = -5.22 kcal/mol
Analysis of T_m Dependence graph: ΔG^0_{37} = -5.03 kcal.mol

3. The ΔG^0_{37} of the two bulges combined is found by subtracting the free energy of the parent duplex from the experimentally determined ΔG^0_{37} found experimentally. This is done using Equation 4.

$$\Delta G^0_{37,\text{ bulges}} = [(-5.03)-(-10.47)] = 5.44 \text{ kcal/mol}$$
$$\Delta G^0_{37,\text{ bulge average}} = [(-5.03)-(-10.47)]\ /2\ = 2.72 \text{ kcal/mol}$$

B. Sample Calculation to Predict the Thermodynamic Stability of the 2nd bulge

Sequence:

$$^{5'}GC\ \ ACU\mathbf{A}CG^{3'}$$
$$^{3'}CG\mathbf{A}UGA\ \ GC^{5'}$$

1. The contribution of one bulge is found using the revised Blose model (McCann et al., 2011), focusing on the weaker stem that is adjacent to the bulge. The bulge that is used for the calculations is the second bulge closest to the 3' on the top strand.

ΔG^0_{37} for $\frac{GC}{CG}$ is -3.42 kcal/mol

ΔG^0_{37} for $\frac{ACU}{UGA}$ is -4.32 kcal/mol

ΔG^0_{37} for $\frac{CG}{GC}$ is -2.36 kcal/mol (less stable stem)

$$\Delta G^0_{37\ (\text{bulge loop})} = (-0.53^* \Delta G^0_{37\ \text{less stable stem}}) + 0.7 \text{ kcal/mol}$$
$$\mathbf{\Delta G^0_{37\ (bulge\ loop)} = (-0.53^* -2.36) + 0.7\ kcal/mol = 1.95\ kcal/mol}$$

2. To determine the free energy of the duplex with one bulge the contribution of the bulge is added to the free energy of the parent duplex:

$$\Delta G^0_{37\ (\text{duplex + bulge})} = 1.95 + (-10.47) = -8.52 \text{ kcal/mol}$$

3. The $\Delta G^0_{37\ (\text{duplex + bulge})}$ can then be subtracted from what was found experimentally for the duplex with both bulges:

$$\Delta G^0_{37\ (\text{2nd bulge})} = (-5.03) - (-8.52) = 3.49 \text{ kcal/mol}$$

C. Sample Calculation for the Free Energy of the Duplex with both bulges:
Sequence:

$$^{5'}GC\ \ ACU\boldsymbol{A}CG^{3'}$$
$$^{3'}CG\boldsymbol{A}UGA\ \ GC^{5'}$$

1. By using the revised Blose model (McCann et al., 2011) the contribution of one bulge is calculated. The stems have different stabilities based on the order of the base pairs.

ΔG^0_{37} for $\frac{GC}{CG}$ is -3.42 kcal/mol
ΔG^0_{37} for $\frac{ACU}{UGA}$ is -4.32 kcal/mol
ΔG^0_{37} for $\frac{CG}{GC}$ is -2.36 kcal/mol (least stable stem)

$\Delta G^0_{37\text{ (bulge loop)}}$ = (-0.53* $\Delta G^0_{37\text{ less stable stem}}$) + 0.7 kcal/mol
$\Delta G^0_{37\text{ (bulge)}}$ = (-0.53*-2.36) +0.7 = 1.95 kcal/mol

2. The predicted ΔG^0_{37} of the duplex with both bulges is determined by adding the contribution of each bulge to the parent duplex's free energy

$\Delta G^0_{37\text{ (sequence)}}$ = (1.95 *2) + -10.47 kcal/mol = **-6.57 kcal/mol**

Appendix F: Statistical Analysis via *GraphPad Prism*

$\Delta G^{\circ}_{37\text{ least stable stem}}$ versus $\Delta G^{\circ}_{37,\text{ 2nd Bulge}}$ and $\Delta G^{\circ}_{37,\text{ average of bulges}}$

		A	B
		Average of bulges	2nd Bulge
		Y	Y
1	Best-fit values		
2	Slope	-0.4259 ± 0.1165	-0.3156 ± 0.2220
3	Y-intercept when X=0.0	2.340 ± 0.4155	3.933 ± 0.7921
4	X-intercept when Y=0.0	5.494	12.46
5	1/slope	-2.348	-3.169
6	95% Confidence Intervals		
7	Slope	-0.6696 to -0.1822	-0.7803 to 0.1491
8	Y-intercept when X=0.0	1.470 to 3.210	2.275 to 5.591
9	X-intercept when Y=0.0	2.278 to 16.98	
10	Goodness of Fit		
11	r^2	0.4131	0.09613
12	Sy.x	1.034	1.972
13	Is slope significantly non-zero?		
14	F	13.38	2.021
15	DFn, DFd	1.000, 19.00	1.000, 19.00
16	P value	0.0017	0.1714
17	Deviation from zero?	Significant	Not Significant
18	Data		
19	Number of X values	21	21
20	Maximum number of Y replicates	1	1
21	Total number of values	21	21
22	Number of missing values	0	0

$\Delta G°_{37\ parent}$ versus $\Delta G°_{37,\ 2nd\ Bulge}$ and $\Delta G°_{37,\ average\ of\ bulges}$

		A	B
		Average of bulges	2nd Bulge
		Y	Y
1	Best-fit values		
2	Slope	-0.2721 ± 0.03950	-0.3464 ± 0.08103
3	Y-intercept when X=0.0	-0.4963 ± 0.6174	-0.3576 ± 1.266
4	X-intercept when Y=0.0	-1.824	-1.032
5	1/slope	-3.675	-2.887
6	95% Confidence Intervals		
7	Slope	-0.3548 to -0.1894	-0.5160 to -0.1768
8	Y-intercept when X=0.0	-1.789 to 0.7959	-3.008 to 2.293
9	X-intercept when Y=0.0	-5.099 to 4.154	-5.932 to 12.74
10	Goodness of Fit		
11	r^2	0.7141	0.4903
12	Sy.x	0.7220	1.481
13	Is slope significantly non-zero?		
14	F	47.45	18.28
15	DFn, DFd	1.000, 19.00	1.000, 19.00
16	P value	< 0.0001	0.0004
17	Deviation from zero?	Significant	Significant
18	Data		
19	Number of X values	21	21
20	Maximum number of Y replicates	1	1
21	Total number of values	21	21
22	Number of missing values	0	0

$\Delta G^{\circ}_{37\text{ least stable stem}}$ versus $\Delta G^{\circ}_{37,\text{ predicted}}$ and $\Delta G^{\circ}_{37,\text{ 2nd bulge}}$

		A	B
		Predicted Bulge	2nd Bulge
		Y	Y
1	Best-fit values		
2	Slope	-0.5301 ± 0.00007099	-0.3154 ± 0.2221
3	Y-intercept when X=0.0	0.6996 ± 0.0002533	3.933 ± 0.7924
4	X-intercept when Y=0.0	1.320	12.47
5	1/slope	-1.887	-3.170
6	95% Confidence Intervals		
7	Slope	-0.5302 to -0.5299	-0.7802 to 0.1494
8	Y-intercept when X=0.0	0.6991 to 0.7002	2.274 to 5.591
9	X-intercept when Y=0.0	1.319 to 1.321	
10	Goodness of Fit		
11	r^2	1.000	0.09598
12	Sy.x	0.0006305	1.972
13	Is slope significantly non-zero?		
14	F	55750000	2.017
15	DFn, DFd	1.000, 19.00	1.000, 19.00
16	P value	< 0.0001	0.1717
17	Deviation from zero?	Significant	Not Significant
18	Data		
19	Number of X values	21	21
20	Maximum number of Y replicates	1	1
21	Total number of values	21	21
22	Number of missing values	0	0

$\Delta G°_{37 \text{ least stable stem}}$ versus $\Delta G°_{37, \text{predicted}}$ and $\Delta G°_{37, \text{Average}}$

		A	B
		Predicted Bulge	Average Bulge
		Y	Y
1	Best-fit values		
2	Slope	-0.5301 ± 0.00007099	-0.4259 ± 0.1165
3	Y-intercept when X=0.0	0.6996 ± 0.0002533	2.340 ± 0.4155
4	X-intercept when Y=0.0	1.320	5.494
5	1/slope	-1.887	-2.348
6	95% Confidence Intervals		
7	Slope	-0.5302 to -0.5299	-0.6696 to -0.1822
8	Y-intercept when X=0.0	0.6991 to 0.7002	1.470 to 3.210
9	X-intercept when Y=0.0	1.319 to 1.321	2.278 to 16.98
10	Goodness of Fit		
11	r^2	1.000	0.4131
12	Sy.x	0.0006305	1.034
13	Is slope significantly non-zero?		
14	F	55750000	13.38
15	DFn, DFd	1.000, 19.00	1.000, 19.00
16	P value	< 0.0001	0.0017
17	Deviation from zero?	Significant	Significant
18	Data		
19	Number of X values	21	21
20	Maximum number of Y replicates	1	1
21	Total number of values	21	21
22	Number of missing values	0	0