Determining the Influence of Group II (G-U) Single-Nucleotide Bulges on the Thermodynamic Parameters of RNA Duplex Formation to Help Develop a Model to Predict Secondary Structure

E R I N C. B U C K L E Y

Allegheny College
Meadville, PA 16335

Determining the Influence of Group II (G-U) Single-Nucleotide Bulges on the Thermodynamic Parameters of RNA Duplex Formation to Help Develop a Model to Predict Secondary Structure

Erin C. Buckley

Senior Comprehensive Project
Allegheny College
Meadville, PA 16335

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

Faculty Committee:
Dr. Serra
Dr. Kleinschmidt
Dr. Deckert

April 7, 2009

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

_____

Erin C. Buckley

A C K N O W L E D G E M E N T S

‡ Dr. Serra- Thank you for being an amazing advisor and giving me my first research experience. Even though I asked a ton of questions you were always willing to help and give me guidance through this project.

‡ Dr. Kleinschmidt and Dr. Deckert- Thank you for being on my committee and guiding me throughout my project and giving me confidence in my work.

‡ Professor McGrath- Thank you for your guidance through Organic Chemistry, which is still my favorite class to this day. Even though you were did not have an office in the basement of Doane you still helped me through my struggles in figuring out the game of life. I appreciate all you have done for me.

‡ Dr. Serra's Compers- We are FINISHED!

- ‡ Alyssa- Thanks for staying up those late nights with me attempting to figure out what our data actually means. And keeping me focused (most of the time). Not only did we accomplish a successful comp but we got through PChem…somehow. You are a great friend and you will do wonderful things after Allegheny, no matter where that leads you.
- ‡ Jess, Ashley, Katie- Great times in our second home (aka Lab). You were all a big help when I needed to figure something out because god knows at least one of you were in the lab when I was there. Best of Luck!

‡ The Others in the Serra lab- Enjoy your time here it flies by…

- ‡ Geoff- Thanks for helping me make conclusions to my comp and guiding me around in lab…I needed it at times. Enjoy your senior year and best of luck on your comp.
- ‡ Michael- It was only a semester but I hope you enjoy RNA. Enjoy taking one of our desks in the "main lab" its sooo much more exciting in there. Thanks for helping me with math questions for PChem. Enjoy the rest of your years here at Allegheny. Best of Luck.

‡ Chrysanthemum- Where do I start? The amazing debriefing talks we had after long days of comping, your pep talks with kept me going when I was about to give up, your advice and guidance through my senior year, cooking lesson, Mum Chocolate Chip Cookies ☺, always giving me encouragement, and being enthusiastic about my comp topic, its simple I'm curing AIDS. You're a wonderful apartment-mate. You have grown so much in the time I have known you…you are going to do great things (just let me come visit where ever you are).

‡ Mike- It was fun living next to you and taking our puppies to the field. Thanks for taking Syd for walks when I couldn't get out of lab. It was great playing TA or should I say you teaching Alyssa and I PChem. Thanks for getting me through just about all of my classes starting with Calc 2. Good Luck at Dental School.

‡ Haley- You were always there for me and came to my rescue when I was about to rip my hair out. You are the ***best*** little & wonderful friend. I will miss you so much next year…but I always know I have a place to stay! Enjoy the remaining two years here it will fly by…don't work too hard its not worth it. Work hard but play harder! I love you.

‡ Casey- Thank you for being a great adopted little. You are a great sister & friend. Thanks for being there for me. Spring break, enough said. Keep dancing, you are a beautiful dancer and opportunities will come your way. Enjoy the remaining two years at Allegheny…go out as much as you can. Keep Josh out of the books. Go Pens! I love you and I will be coming to sleep on your couch next year.

‡ My AXΩ Companion- You were amazing with all the gifts. Every time I was having trouble seeing the end you brightened my day. Your letters were posted in my cubical to keep me motivated to meet you in the end! Thanks for all your hard work and keeping me sane. You're the best.

‡ Mom and Dad- Thanks for putting up with me throughout my four years. They started rough but with your help I found the place for me. Thanks for your guidance in figuring out my life and putting up with my grouchiness while I was comping. Thanks for the calls; they always seemed to come at the right time.

‡ Sean and Brian- Even though you still don't understand fully what a comp is thanks for trying to figure out what my title means…RNA "I know those letters". You two are great brothers. Sean I am so glad we are still so close I am sorry I did not get to come see you as much as I wanted…I was comping (no excuse I know). Brian enjoy college it goes so fast, just don't forget to work hard, I know you got the party part down ☺.

‡ Lisa- You're my best friend…and boy did I miss you, this year especially. You are one of the strongest people I know. I always looked forward to our phone conversations long and short. You helped me so much when you were at Allegheny I can't even explain it…even though I speak my mind a little more now. You listened to my struggles and joys. I wish you the best of luck finishing up your masters and wherever that leads you…MD or DO you will get there I have faith in you. I miss you so much and can't wait to see you again. I love you.

‡ Wicked Jim- It was wonderful having you up at school with me, you kept me sane.

‡ Sydney-You are the best companion. Thanks for hanging out in the science building with me, running around the halls to keep me up, and snuggling with me every night. Every time I was down you brought a smile to my face. We made it through senior year.

‡ Jeff- Thanks for always being there for me, especially when I had my mental breakdowns. Thanks for putting up with my complaining not only about comping but the four years of classes. The flowers and sour patch kids could not have come at a better time! Thanks for picking me up when I fall… ☺ and being my best friend for the last 4+ years ☺. I love you.

A B S T R A C T

Thermodynamic parameters ($\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and T_M) of oligonucleotide sequences containing single nucleotide bulge loops were determined through optical meeting in 1M NaCl. Seven sequences containing Group II bulge loops adjacent to a wobble base pair (G-U) were analyzed to see if the non-nearest neighbor model currently used for Group I single nucleotide bulges is an accurate model for Group II bulges. The identity of the bulge did not show an effect on the stability of the duplex containing the bulge loop; therefore, all Group II (G-U) bulges were considered one set. The bulge sequence was studied at varying positions along the duplex: 5', 3, and/or middle. There was a correlation between the least stable stem and the destabilization of the duplex. The free energy increments at 37°C for the introduction of a Group II (G-U) single nucleotide bulge loop ranged between 2.0 and 5.2 kcal/mol. The thermodynamic parameters, $\Delta G°_{37}$ and $\Delta H°$, were combined with previous data gathered for Group II (G-U) (Serra unpublished) and the free energies were compared to Group I (Blose *et al*. 2007). Based on this investigation the characteristics of Group II (G-U) bulges were found to show a relationship to Group I bulges; therefore, the non-nearest neighbor model would appear to be an accurate model for predicting the stability of duplexes containing a Group II (G-U) single nucleotide bulge.

T A B L E O F C O N T E N T S

Acknowledgements……. i

Abstract…… iii

Table of Contents…… iv

Introduction……. 1
RNA Function and Structure…… 1
Importance of Secondary Structural Motifs……. 5
Single Nucleotide Bulges……. 6
Thermodynamic Analysis……. 9
This Investigation……. 13

Material and Methods…… 14
RNA Synthesis and Purification……... 14
Optical Melting Experiments and Data Analysis……. 18
Statistical Analysis…… 19
Phylogenic Analysis of Group II (G-U) Nucleotide Bulge Loops……... 19

Results…….. 20
Thermodynamics…….. 20

Discussion…… 24
Analysis of Bulge Identity…… 25
Analysis of Less Destabilizing Effect…….. 26
Nearest Neighbor Effect……... 26
Less Stable Stem Effect…… 28

Future Research…….. 30

References…… 32

Appendicies:
A: Derivation of Equation 1-Meltwin Fitting Equation……. 35
B: Derivation of Equation 2-T_M Dependence Equation…… 38
C: Detailed Purification Procedure…… 40
D: Group II (G-U) – Raw Data……. 52
E: Sample Calculation – Gibb's Free Energy Associated with Bulge ($\Delta G^\circ_{37,\,bulge}$).. 61
F: Collective Group II (G-U) Thermodynamic parameters……... 62

I N T R O D U C T I O N

RNA Function and Structure

Ribonucleic acid (RNA) was a major focus for scientific research even before there was evidence of its role in gene regulation. RNA is now known to be involved in all the major steps of transmission of genetic information from the deoxyribonucleic acid (DNA) to the protein level. The roles of RNA include genetic information storage, binding of small molecules and proteins, enzymatic activities such as intron sequence splicing and polymerization, as well as peptide bond formation during protein biosynthesis (Yusupov *et al.* 2001, Sashital & Butcher, 2006, and Weaver).

The three main structural levels of RNA are primary, secondary, and tertiary. Initially, RNA has a primary structure of nucleotides linked covalently by phosphodiester bonds between the 5' phosphate group of one nucleotide and the 3' hydroxyl group of the next nucleotide. The complexity of RNA folding can be divided into two levels, secondary and tertiary structures. The more complex tertiary structure, or three-dimensional arrangement, determines the cellular functions. Preceding the native functional folding of the tertiary structure is the secondary conformation. The secondary structure of RNA is determined by either Watson-Crick base pairing, wobble base pairing, or unpaired nucleotides (Mathews *et al.* 1999). Formation of different tertiary structures of RNA could be understood if an accurate method to determine the secondary structure was developed by analyzing the secondary interactions (Blose *et al*. 2007; Mathews *et al.* 1999).

RNA single strands have a phosphodiester backbone which is formed from the 5' carbon of the sugar attaching to the 3' hydroxyl group of the following nucleotide

(Figure 1). The duplex is formed when two single strands hydrogen bond together in an anti-parallel manner. Traditionally, the primary RNA sequence is written and read 5' to 3'.

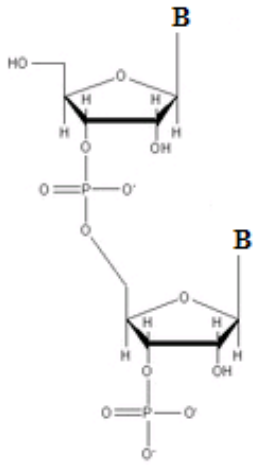


Figure 1. Phosphodiester backbone of RNA runs from the 5' carbon of the sugar to the 3' hydroxyl group of the following nucleotide. The B denotes a nitrogenous base. http://web.virginia.edu/Heidi/chapter11/Images/8883n11_17.jpg

RNA contains four different nitrogenous bases: uracil, cytosine, adenine, and guanine. The nucleotides are classified, based on their ring structure, as pyrimidines, a six member ring, or purines, a six- member ring fused to a five-member ring. Uracil (U) and cytosine(C) are pyrimidines and adenine (A) and guanine (G) are purines (Figure 2).

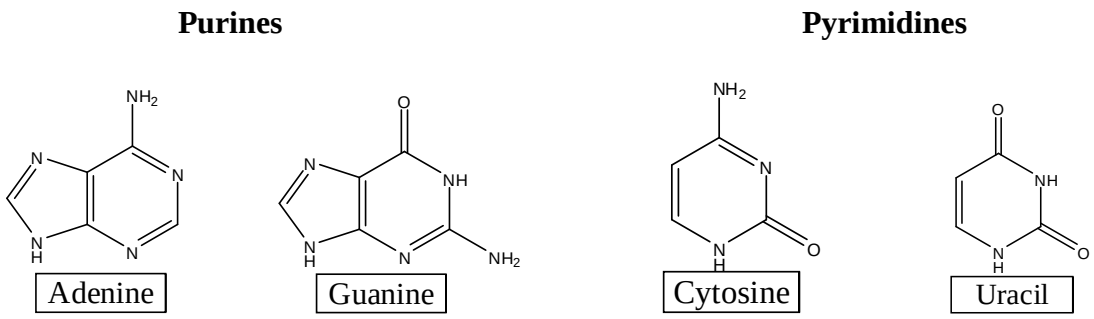


Figure 2. Nitrogenous bases of RNA. The bases are in two groups depending on their ring structure. Adenine (A) and guanine (G) contain a six-member ring connected to a five-member ring. Cytosine (C) and uracil (U) contain a six-member ring.

The secondary structure of RNA is formed from hydrogen bonding between base pairs either from a complementary strand of RNA or from a strand folding onto itself. The bases typically pair A-U and G-C, following the Watson-Crick base pairing model. Two and three hydrogen bonds form between A-U and G-C, respectfully. However, non-Watson-Crick base pairing does occur in nature; including mismatches such as A-G base pairing (Nagaswamy *et al.* 1999; Zacharias and Sklenar, 1999) or wobble base pairing G-U (Figure 3).

A. B.

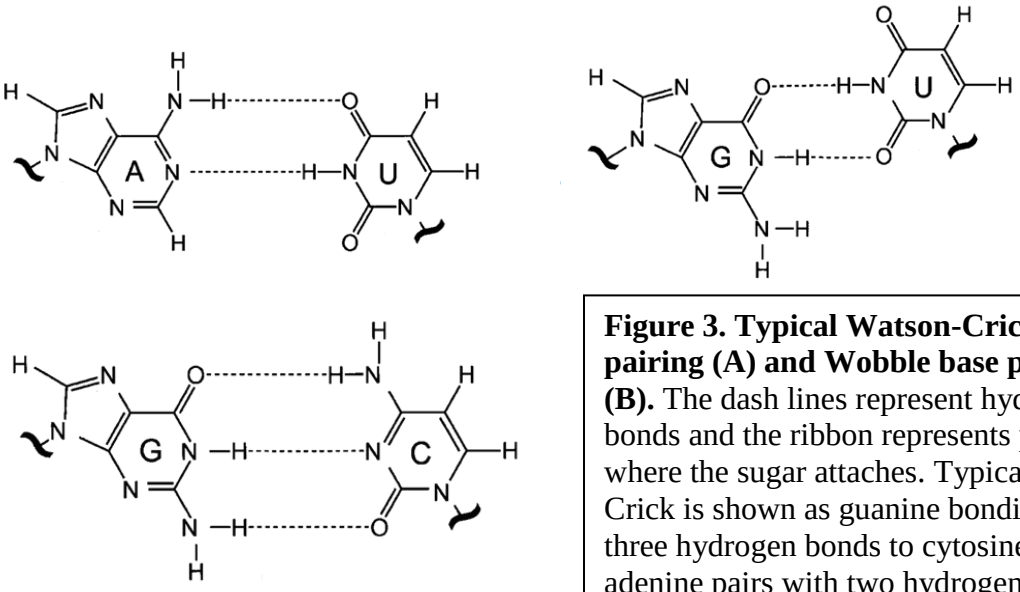


Figure 3. Typical Watson-Crick base pairing (A) and Wobble base pairing (B). The dash lines represent hydrogen bonds and the ribbon represents position where the sugar attaches. Typical Watson-Crick is shown as guanine bonding with three hydrogen bonds to cytosine while adenine pairs with two hydrogen bonds to uracil. The wobble base pair is guanine binding to uracil through two hydrogen bonds (Varani and McClain, 2000).

Different structural motifs are found in the secondary structures including: hairpins, bulge loops, internal loops, and branched loops (Figure 4). Hydrogen bonding between the nitrogenous bases enable the duplex to be formed, however the structural motifs, mentioned above, are formed because a nitrogenous base does not have a

complement base to hydrogen bond. When these secondary structures bend, fold, and twist the tertiary structure is formed (Figure 5).

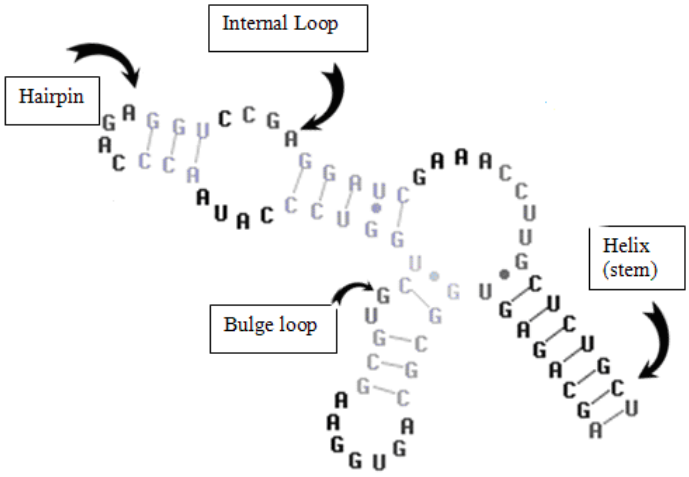


Figure 4. Common RNA Secondary Structure Motifs. The nucleotides are represented by their abbreviation G, C, A, U. The dashed lines represent hydrogen bonding between nucleotides following Watson-Crick base pairing (Goldgof, 2006).

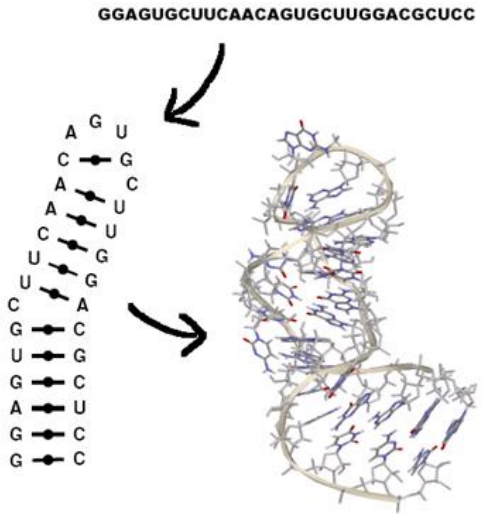


Figure 5. Primary, secondary, and tertiary structures displayed. The primary sequence (at top) is the nucleotides written 5' to 3'. The secondary structure shows the self-complementary folding of the primary structure (lower left). The tertiary structure shows the backbone and nucleotides in their functional state. http://www-lbit.iro.umontreal.ca/mcfold/logo.mcfold.png

Importance of Secondary Structural Motifs

Bulges were previously found to be the recognition feature of RNA in RNA-protein interactions. For instance many RNA/protein interactions including the coat protein of R17 coliphage, Aenopus protein TFIIIA, and E.coli ribosomal proteins were all found to interact with the secondary structure of the RNA. In one study the RNA's single nucleotide bulge was manipulated to see if the iron-responsive element binding protein (IRE-BP) still had the same affinity for the RNA. It was found that by changing or removing the cytosine in the RNA the affinity for the interaction decreased (Jaffrey *et al*. 1993).

Another protein/RNA interaction was found to be the human immunodeficiency virus type 1 (HIV-1) protein, Tat, binding to the trans-activation response region (TAR) allows for transcription to occur. Tat is considered a transactivator because it increases the rate of the viral gene expression. TAR was found on the 5' non-coding region of all HIV-1 mRNAs by mutating the RNA sequences to see how the affinity interaction changed. The RNA comprising the TAR region contains a stable stem-loop secondary motif structure. The region of nucleotides +18 through +44 is crucial for the trans-activation response. More specifically, the nucleotides which make up the loop, +23 to +25, contain a uracil in position +23 which was found to be most critical for the overall interaction between Tat and TAR; thus viral transcription (Figure 6) (Roy *et al.* 1990).



Figure 6. The viral TAR RNA with the bulge binding to the Tat protein for transcription (left). If the Tat protein is not bound to the TAR RNA (right) replication does not occur; no trans-activation. http://www.nature.com/nature/journal/v418/n6894/images/418252a-

Single-nucleotide Bulges

Double stranded regions of RNA are frequently interrupted by different secondary structural motifs. One commonly found motif is the bulge loop. The bulge is when one or more nucleotides are unpaired; therefore, protruding from the stem of the RNA helix. The influence of a single-nucleotide bulge motif has been studied and found to widen the major groove, revealing it for RNA interaction with either a protein or ligand. Bulges in many RNAs were found to be critical for RNA structure and function, including intron splicing and tertiary folding (Zacharias *et al.* 1999 and Znosko *et al.* 2002).

Investigations of bulges in the RNA secondary structure were conducted in order to find a trend in the bulge location and bulge stacking. Several methods have been utilized to study these secondary structures containing single nucleotide bulges including NMR spectroscopy, x-ray crystallization, and structural mapping using chemical and enzymatic probes (Znosko *et al.* 2002; Furtig *et al*. 2003; Jaeger *et al.* 1993; Ehresmann *et al.* 1987). The three dimensional structure of 23S rRNA of *Haloarcula marismortui*

and 16S rRNA of *Thermus thermophilus* were determined through X-ray crystallography allowing the single nucleotide bulges to be characterized (Znosko *et al.* 2002).

Bulge geometries were found to fall into three structural families: extrahelical, intercalated, and side-by-side, determined based on the geometry and direction of the bulges (Figure 7). Extrahelical bulges are most commonly found in nature, displaying the nucleotide bulge away from the helix (Znosko *et al.* 2002). This family can be distinguished based on the position of the phosphate group in relation to the minor groove. It can be lined up with the 5' or 3' nucleotide flanking the bulge. The intercalated bulges, which were found to be highly conserved in bacterial RNA, are typically comprised of an adenine bulge nucleotide adjacent to a guanine (Zacharias & Sklenar, 1999 and Znosko *et al.* 2002). The intercalated bulge creates a twist in the helix putting the bulge nucleotide between the adjacent nucleotides in the helix (Zacharias & Sklenar, 1999). Side-by-side bulges either have the bulge nucleotide side-by-side with the nucleotide on the 5' or 3' end. This type of bulge creates a more exaggerated twist of the helix and a break in the stacking interactions on the bulge side (Znosko *et al.* 2002).

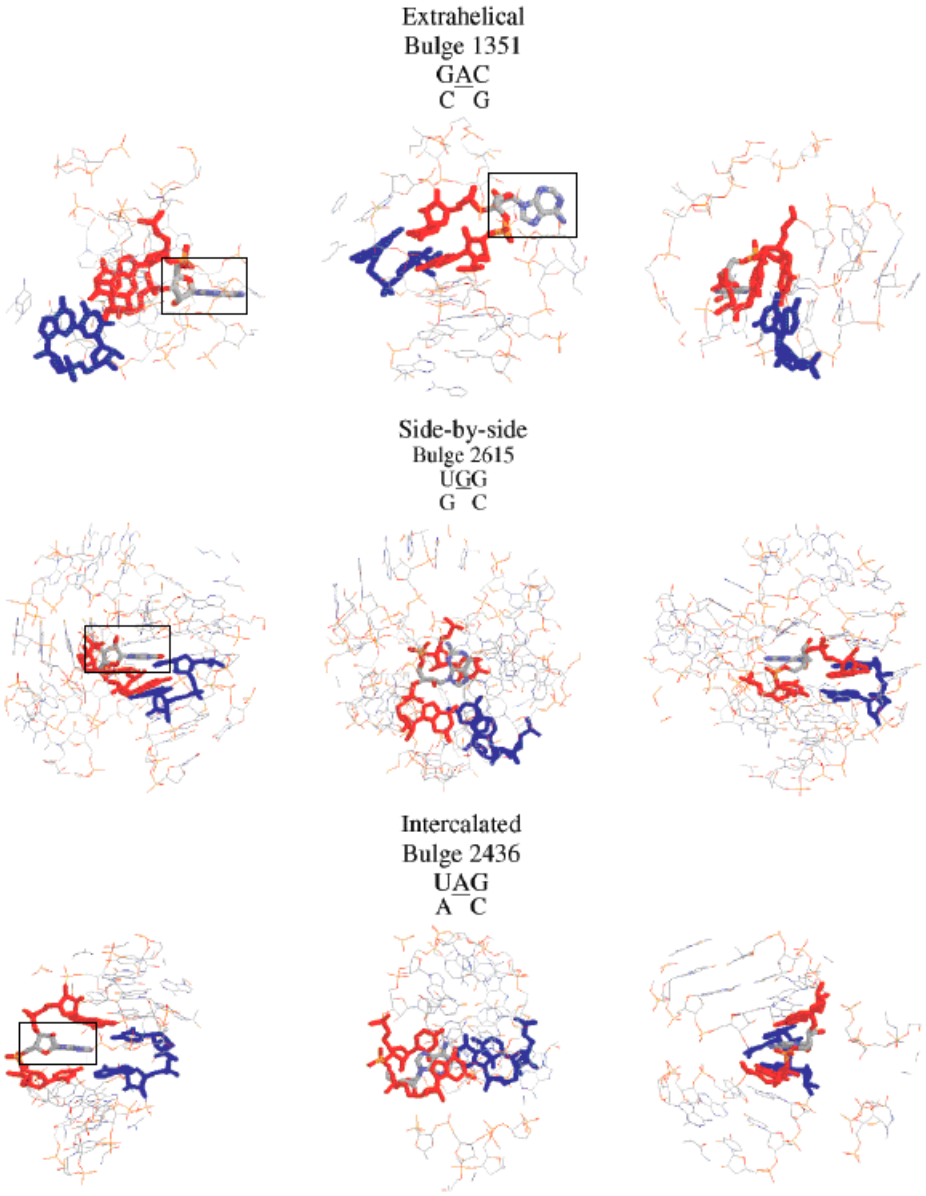


Figure 7. The three structural families of bulge nucleotide conformations. These examples of conformations are observed in the large ribosomal RNA subunit. The red strand contains the bulge nucleotide (gray and boxed) while the blue represents the complementary strand. The wire frame shows the surrounding molecules. (Znosko *et al*. 2002)

In addition to determining the geometry of different single nucleotide bulges by Znosko and colleagues (2002) classifications for different bulges were revised by Blose and colleagues (2007). Group I bulges are unique because they have one unpaired

nucleotide that is not identical to either of the flanking nucleotides. Group II bulges are ambiguous in location because the bulge nucleotide, or unpaired nucleotide, is identical to an adjacent nucleotide. The ambiguity of Group III arises from the potential of wobble base pairing with the adjacent nucleotide (Table 1) (Blose *et al.* 2007 and Znosko *et al.* 2002).

	Potential Bulge Positions	
Group I	5' CGC**A**GCC 3' 3' GCG CGG 5'	
Group II	5' CGC**C**GCC 3' 3' GC G CGG 5'	5' CG**C**CGCC 3' 3' GC G CGG 5'
Group III	5' CGC**C**UGG 3' 3' GCG GCC 5'	5' CGCC**U**GG 3' 3' GCGG CC 5'

Table 1. Group classifications for single nucleotide bulge loop sequences. The red letter represents the unpaired nucleotide. Group II and III show the ambiguity of bulge position.

Thermodynamic Analysis

The common method for studying the stability of RNA bulges is optical melting which provides useful thermodynamic information to determine the influence the bulge loop has on the stem stability. The optical melting procedure is conducted using a spectrophotometer, which uses a ultra-violet light to measure the absorbencies (Turner *et al.* 2000). The RNA duplexes are formed by single strands annealing together through hydrogen bonds. The disruption of the hydrogen bonds in the RNA duplex is caused as the duplex is heated to 85°C. As the duplex is denatured more light is absorbed, also known as hyperchromcity (Puglisi & Tinoco, 1989). Thus, the absorbance of the double strand is less than that of the single strands (Figure 8). The absorbance changes of the

oligomers were recorded as a function of temperature at a rate of 1°C/minute (Blose *et al.* 2007) (Figure 9).

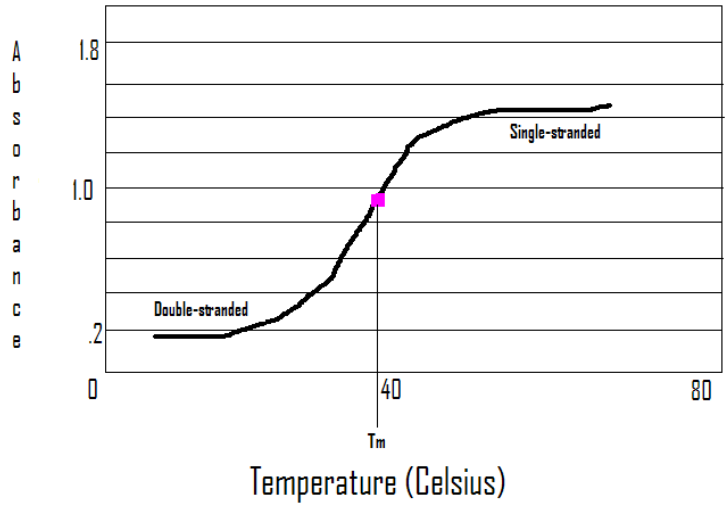


Figure 8. An example of a theoretical standard melt curve. The absorbance is the absorbance at 280nm (y-axis) verses the temperature measured in Celsius (x-axis). Using the Beckman DU 640 Spectrophotometer with High Performance Temperature Controller. The T_m measures the temperature where the molar ratios of the single and double strands of RNA in solution are equal.

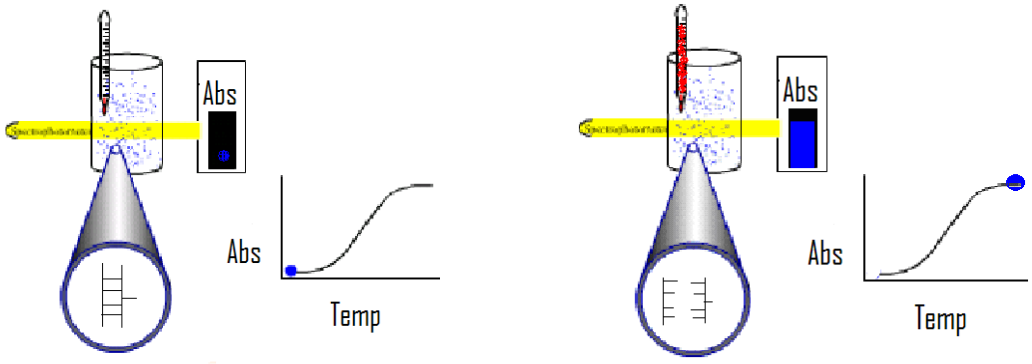


Figure 9. The diagram of the optical melting process looks at temperature and denaturation corresponding to the analysis graph. The image on the left shows the initial duplex at low temperatures. As the temperature increases the absorbance increases due to the denaturing of the RNA duplex into the single strands.
http://webpub.allegheny.edu/employee/m/mserra/research/demos/opticalMelt.htm

A plot of absorbance verses temperature (Celcius) provides a melt curve. The midpoint of the curve represents the melting temperature (T_m). The melting temperature corresponds to equal molar ratios of double and single-stranded RNA (Figure 8).

The data from the optical melt curve was gathered for a specific sequence to determine the thermodynamic parameters: enthalpy (ΔH°), entropy (ΔS°), Gibb's free energy (ΔG°), and melting temperature (T_m) (Puglisi & Tinoco, 1989). The change in enthalpy is the measure of the amount of heat lost or gained in a reaction and the change in entropy is the measure of disorder in a system. To analyze this data Meltwin 3.0, a non-linear least-square program, uses a standard two state model (double /single strand) (Hang *et al.* 1998 and Meltwin 3.0). The data from the individual melts are combined and plotted for normalized absorbance vs. temperature (Celcius) and then the curve of the line is fitted using a derivative of the Beer's Law. Meltwin 3.0 program uses the following equation in order to fit the thermodynamic data (derivation in Appendix A) (Xia *et al.* 1998).

$$A_t = (\{[(e^{(T\Delta S^\circ - \Delta H^\circ)/RT})^2 + 2(e^{(T\Delta S^\circ - \Delta H^\circ)/RT})C_T]^{1/2} - e^{(T\Delta S^\circ - \Delta H^\circ)/RT}\}/2)(2M_{SS}T + 2b_{SS} - M_{DS}T - b_{DS}) + (C_T/2)(M_{DS}T + b_{DS})$$

(Equation 1)

After fitting the thermodynamic data a graph is compiled. The graph displays the relationship between the inverse of the melting temperature (T_m) and the natural logarithm of the total concentration of the single stranded RNA ($\ln(C_T/4)$). The line of best fit varies with relation to the concentration of single stranded RNA (equation 2; derivative in Appendix B). Equation 2 follows the typical y=mx+b line equation; therefore, m is the slope, x is the x-intercept, b is the y-intercept. Using this information gathered Gibb's free energy (ΔG°_{37}) is calculated by Meltwin 3.0 using the equation $\Delta G^\circ_{37} = \Delta H^\circ - T\Delta S^\circ$ (T=310.15K). (Znosko *et al.* 2002).

$$T_M^{-1} = (R/\Delta H^\circ)\ln(C_T/4) + (\Delta S^\circ/\Delta H^\circ)$$
$$y = m x + b$$

(Equation 2)

The nearest neighbor model is the most commonly used method for predicting the free energy of an RNA duplex formation (Xia *et al.* 1998). This model uses: $\Delta G^\circ_{37\ (bulge)} = \Delta G^\circ_{37\ (measured)} - \Delta G^\circ_{37\ (duplex)}$, to calculate the free energy associated with the bulge. The other thermodynamic parameters, enthalpy and entropy, use the same model only using the respective values (ie. $\Delta H^\circ_{\ (bulge)} = \Delta H^\circ_{\ (measured)} - \Delta H^\circ_{\ (duplex)}$) (Blose *et al.* 2007). Here the measured value is the experimental data from the optical melt and the duplex data is calculated from the nearest-neighbor model as if the bulge was not present. After a full set of free energy values were obtained for Group I bulge loops a new model, non-nearest-neighbor model, was established as a revised nearest neighbor model proposed by Znosko et al. (2002).

Blose and colleagues (2007) found that by adding a single pyrimidine bulge into the duplex formation of the duplex is destabilized by 3.9 ± 0.8 kcal/mol while the purine added 4.2 ± 0.8 kcal/mol; therefore, statistically the identity of the bulge nucleotide does not need to be considered in calculating duplex stability. The stability of the stem containing Watson-Crick base pairs next to the bulge was compared to a stem containing wobble base pairing adjacent to the bulge. This was done to see if there was a relationship between the stability of the duplex, with the thermodynamic contribution of the bulge on the stem formation. In comparing the linear fit of data for bulges flanked by Watson-Crick and wobble base pairing there was no significant variation; therefore, the data was combined (Table 2). The data showed that the position of the wobble base pair

on either the 5’ or 3’ end of the bulge nucleotide and the orientation of the wobble base pair had an influence on the stability (Blose *et al.* 2007). Therefore, bonuses and penalties were accounted for in the model depending on the orientation and position of the wobble base pair relative to the bulge (Table 2) (Blose *et al.* 2007). The stability of the helix stem influences the bulge through the tight packing and electrostatic strain present in the RNA duplex (Blose *et al.* 2007). The wobble pairs decrease the stacking interactions in the duplex making the insertion of a bulge nucleotide less destabilizing, and resulting in a positive free energy for the duplex formation. The destabilization provided by the wobble base pair is often necessary for the folding of the secondary structure (Blose *et al.* 2007).

Free Energy increment Equation for adding Group I bulge loop Into RNA duplex	$\Delta G^\circ_{37,\text{ bulge loop}} = -\ 0.62\Delta G^\circ_{37,\text{ less stable stem}} + 0.25(\text{kcal/mol})$
Nearest-Neighbor Bonus: – 0.6 kcal/mol	$\left(\frac{5'UBX3'}{3'G-Y5'}\right)\left(\frac{5'GBG3'}{3'U-U5'}\right)\left(\frac{5'UBU3'}{3'G-G5'}\right)$
Nearest-Neighbor Penalty: +0.4 kcal/mol	$\left(\frac{5'GBX3'}{3'U-Y5'}\right)\left(\frac{5'XBU3'}{3'Y-G5'}\right)$
Enthalpic increment for Inserting Group I Single-Nucleotide bulge Loop into RNA duplex	$H^\circ_{\text{ bulge loop}} = 16.5$ kcal/mol (independent of stem or nearest-neighbor effects)

Table 2. Group I single-nucleotide bulge loop thermodynamic parameters for non-nearest-neighbor model. X and Y represent a Watson-Crick base pair in the sequence. The B is a nucleotide bulge, not identical to an adjacent nucleotide. The (-) represents no hydrogen bonding; therefore, a bulge nucleotide (Blose *et al*. 2007).

This Investigation

The previous studies on single nucleotide bulge loops were completed with Group I bulges, which now have a comprehensive set of data (Blose *et al.* 2007). Trends in the thermodynamic parameters of Group I single nucleotide bulges heavily depends on the stability of the helix stem and the orientation and position of the flanking wobble base pair in relation to the bulge nucleotide. Group I bulges were compared to the limited data on Group II single-nucleotide bulges and it was concluded, thus far, that Group II bulges are more stable than Group I bulges (Znosko *et al.* 2002).

For this investigation, Group II sequences with at least one G-U wobble base pair adjacent to bulge will be analyzed. Initially, the bulge sequence $\frac{5'U\mathbf{A}\mathbf{A}G3'}{3'GU-C5'}$ (bold letters represent the bulge ambiguity) was analyzed due to it consisting of nearly 50% of the Group II (G-U) bulges found in RNA that are in a database of known secondary sequences. The next bulged sequences analyzed will focus on the bulge nucleotide being adjacent to two wobble base pairs $\frac{5'U\mathbf{A}\mathbf{A}U3'}{3'GU-G5'}$, $\frac{5'G\mathbf{U}\mathbf{U}G3'}{3'UA-U5'}$, and $\frac{5'G\mathbf{U}\mathbf{U}G3'}{3'CA-U5'}$ for comparison. The sequences used will contain the bulge sequence in a duplex of seven to ten nucleotides in varying locations along the duplex to determine if the adjacent stem stability effects the destabilization of inserting a bulge nucleotide as seen with Group I bulges. The thermodynamic parameters obtained from this study will be combined with data from previous studies in order to develop a more defined model for predicting secondary structures with a single nucleotide Group II (G-U) bulges.

MATERIAL AND METHODS

RNA Synthesis and Purification

The oligonucleotides were synthesized at University of Rochester or W.M. Keck Oligonucleotide Synthesis at Yale University. They were made on a 1 μmole scale on a controlled pore glass (CPG) solid supports (Applied Biosystems 392 DNA/RNA synthesizer) utilizing phosphoramidites which were acquired from Proligo©. The 2'-hydroxyl group of the ribose sugar was protected with a *tert*-butyldimethylsilyl ether and a benzoyl group protects the amine of each base (Figure 10; Blose *et al*. 2007; Vecenie *et al.* 2006).

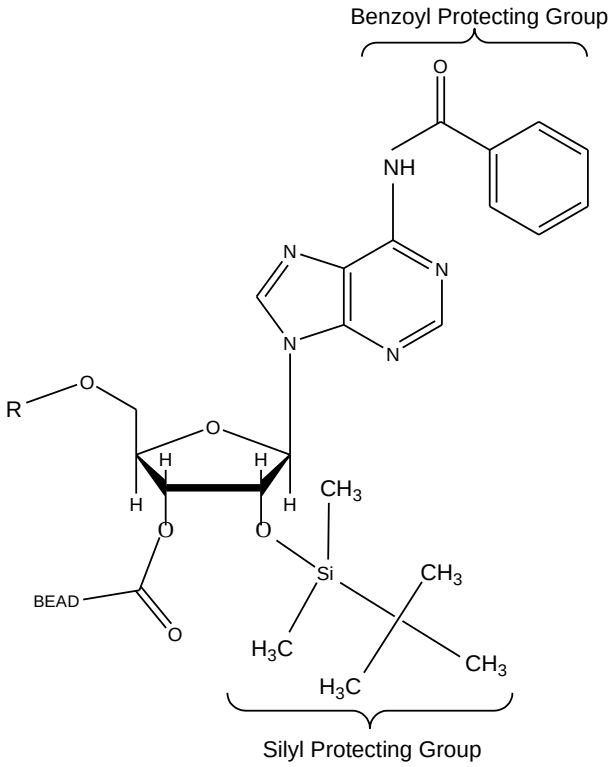


Figure 10. The nucleotide pictured above is an example of how the RNA is protected from reactivity of the oligomer when synthesized. Here an adenine nucleotide has a silyl protecting group on the 2' position of the ribose sugar and a benzoyl protecting group on the amine of the nucleotide in order to prevent reactivity of the oligomer. The nucleotide is attached to a controlled pore glass solid support (CPG), by an ester linkage.

The RNA oligomer was removed from the CPG or beads via cleavage of the ester linkage. The ester bond and the benzoyl protecting groups were cleaved using a 3:1 solution cold, fresh ammonium hydroxide: 100% ethanol: in a 55°C water bath for 18 hours. The protecting groups are non-polar; therefore, they are able to dissolve in the ethanol which acts as a co-solvent (Figure 11).

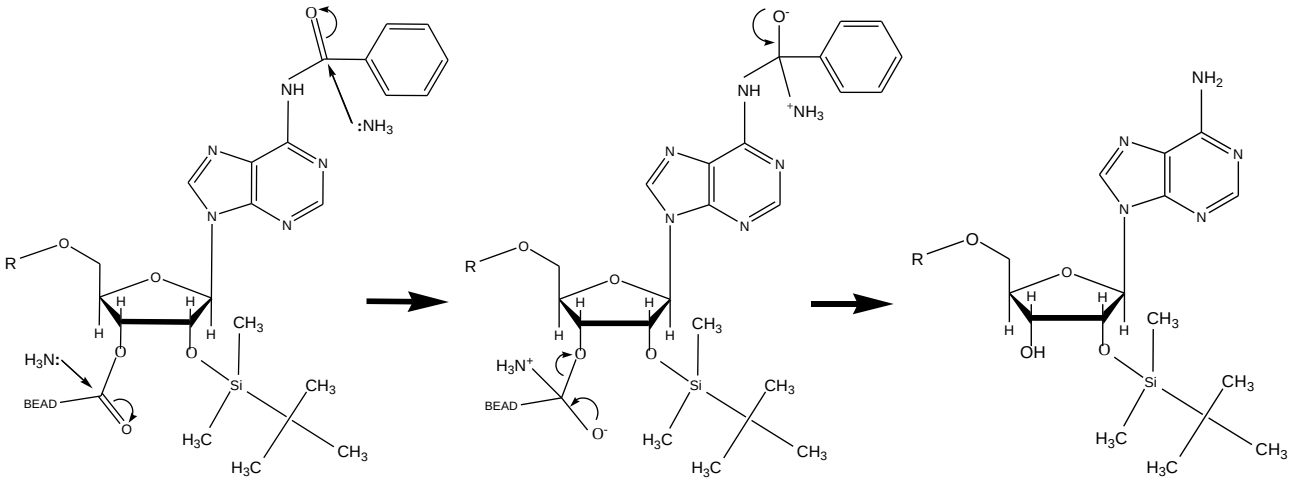


Figure 11. Removal of benzoyl protecting group and the bead. Ammonium hydroxide cleaved the ester bond that attaches the adenosine to the bead and the benzoyl protecting group on the amide of the nucleotide.

The RNA solution was washed from the beads using a disposable chromatography column using various polarities of solvents: distilled water, 50% acetonitrile, and 100% acetonitrile. The corrosive ammonia was removed using a Savant Speed-Vacuum concentrator centrifuge attached to a water aspirator. After one hour, the solvent was then dried overnight using the Speed-Vacuum concentrator centrifuge attached to a vacuum pump.

The silyl protecting groups on the 2' hydroxyl position of the RNA was removed using a 9:1 solution of triethylamine hydrogen fluoride (TEAHF) and dimethylformamide

(DMF) for two hours at 55°C. The TEAHF removed the protecting silyl group by cleaving the ester bond (Figure 12**)** while DMF acted as a solvent.

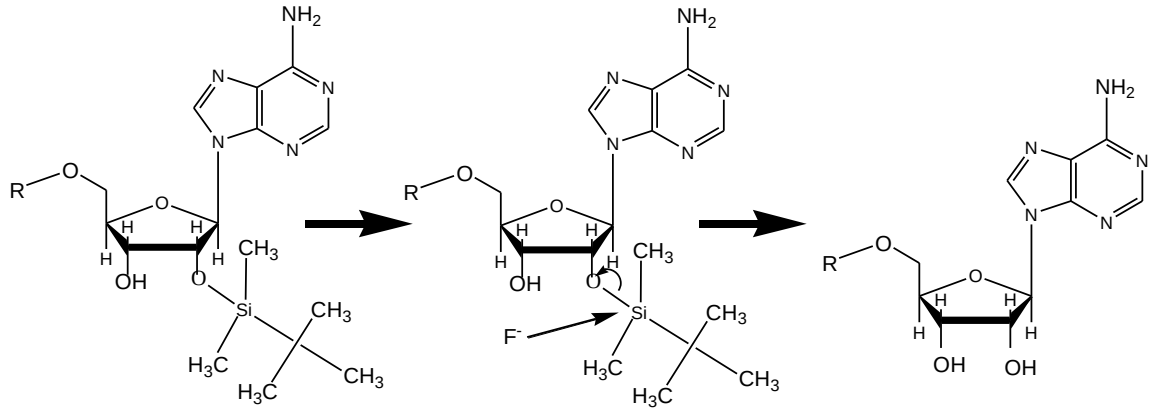


Figure 12. Removal of the silyl protecting group. The silyl group on the 2' hydroxyl position of the ribose sugar was removed during the incubation of the RNA sample with triethylamine hydrogen fluoride (TEAHF) and dimethylformamide (DMF). DMF acted as the solvent while the TEAHF removed the silyl group by cleaving the ester bond.

The RNA was then resuspended in a total of 8mL of 20X Sep Pak buffer (0.1M ammonium bicarbonate, pH 4.5). A Water Sep Pak C-18 chromatography column was used to remove the inorganic salts from the oligomers. Once the RNA was separated, which was determined by UV spectroscopy at wavelength 280nm, it was dried overnight and resuspended in distilled water. Thin Layer Chromatography (TLC) was used to separate the desired RNA from the sequence failures. The RNA sample was spotted on the preparative TLC plate (20X20 cm, 500μm thickness silica with fluorescence indicator) using a 5μL pipet. The TLC plate was placed in a chromatography chamber with a mobile phase of a 55:35:10 solution of n-propanol: ammonium hydroxide: water. Ultraviolet light was used to visualize the bands on the plate. The lowest darkest band, which contained the desired RNA sequence, was scrapped off the plate into a sterile test tube. Distilled water was used to remove the RNA from the silica. The RNA was

combined with 2mL of 20X Sep Pak buffer and the remaining salts in the sample were removed using another Water Sep Pak C-18 chromatography column. The sample was then dried in the Speedvac and resuspended in a known volume of distilled water.

The purity of the RNA sample was then determined using reverse-phase High Performance Liquid Chromatography (HPLC) using a concentration gradient of 0% Buffer B (50:50 solution of 10mM sodium phosphate 12-Hydrate pH 7.2: HPLC grade methanol) to 40% Buffer B. A pure sample would result in a single peak on the HPLC output. If more than one peak was present the purification process was started again from the analytical TLC plate. If impurities continued to show on the HPLC output the TLC was run with two mobile phases, first a 9:1, chloroform: methanol followed by the original mobile phase. The added mobile phase was to help remove any salts still present in the RNA sample.

The absorbance of a 1:100 solution of the pure RNA sample was measured at 260 nm and 280 nm at 85°C for 120 seconds using a Beckman DU 640 Spectrophotometer with High Performance Temperature Controller. The Beer's Law formula was used to calculate the concentration of each RNA oligomer. Equimolar amounts of each RNA strand were mixed together and dried in the Speedvac and resuspended in melt buffer (1M NaCl, 10mM cacodylic acid, 0.5mM EDTA, pH 7) so that the RNA single strands would anneal to include the Group II bulge.
(See Appendix C for a detailed version of the purification process).

Optical Melting Experiments and Data Analysis

All samples were melted in the melt buffer (1M NaCl, 10mM cacodylic acid, 0.5mM EDTA, pH 7) and the absorbance of the RNA solution containing the bulge was

measured at 280 nm starting at 85°C as the temperature changed at a cooling rate of 1°C/min using a Beckman DU 640 Spectrophotometer with High Performance Temperature Controller. If the melting curve had a good upper base line the temperature was truncated in order to keep from over heating the RNA. As a result of the sequences being rich in guanine and cytosine the absorbance was taken at 280 nm due to the greater hyperchromicity, of G-C base pairs, thus resulting in better optical melting data. The absorbencies of RNA solution were measured at 11 different concentrations to achieve a sample concentration range of at least 100 fold. Meltwin 3.0, a least square regression program, was used to fit the optical melting curve data assuming a two state model with no intermediates. Meltwin 3.0 was used to determine the thermodynamic parameters of the non-self complementary RNA duplex by fitting the optical melting data for RNA strands using Beer's Law (Equation 1; Appendix A). Then by graphing the inverse of the melting temperature (T_M) verses the natural logarithm of the total concentration of the single stranded RNA ($\ln(C_T/4)$) to determine how concentration changes with temperature (Equation 2; Appendix B). The thermodynamic parameters were calculated by averaging the values obtained from the Meltwin 3.0 least squares regression fit and the plot of T_M^{-1} verses $\ln(C_T/4)$. The two parameters from Meltwin 3.0 regression line and the plot of the T_M^{-1} must agree within 15% to adhere to the two state model (Freier *et al*. 1984).

Statistical Analysis

The thermodynamics of Group II bulge loops closed by one or two wobble (G-U) base pairs were compared to sequences containing Group I bulge loops adjacent to a wobble (G-U) base pairs and illustrated graphically using GraphPad Prism graphing

program. GraphPad InStat statistics program was then used to determine whether the wobble base pair(s) affects the stability of the RNA duplex.

Phylogenic Analysis of Group II (G-U) Nucleotide Bulge Loops

In order to determine which Group II bulges closed by one or two wobble base pair(s) should be studied, a Hairpin/Bulge Calculator (Visyak unpublished) program was used. This program calculates the frequency of Group II bulge found in naturally occurring sequences. The sequences searched were from 16S rRNA, 23S rRNA, and catalytic intron RNAs (Group II introns). The structural files searched by this program were from the University of Texas Comparative RNA Web Site, www.rna.icmb.utexas.edu (Cannone *et al.* 2002).

R E S U L T S

Thermodynamics

The Group II bulges closed by one or two guanine-uracil wobble base pair(s) (Figure 13) were investigated out of the 43 combinations (Appendix F). The thermodynamic parameters were found for seven different bulge sequences containing the desired Group II (G-U) specifications. Some bulge sequences were used more than once by changing the bulge positioning in the duplex sequence from 5', 3', and/or middle of the sequence. The RNA oligomers studied typically were between seven and ten nucleotides in length.

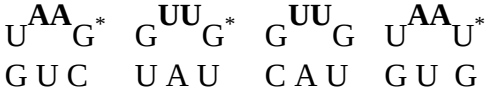


Figure 13. The Group II (G-U) bulges analyzed in this study. The bold nucleotides are the ambiguous bulge and the adjacent nucleotides either follow Watson-Crick base pairing or Wobble base pairing. * Sequences are ones that were used in multiple sequences varying the position within the RNA duplex (ie. 5', 3', or middle).

The thermodynamic parameters: enthalpy ($\Delta H°$), entropy ($\Delta S°$), free energy ($\Delta G_{37}°$), and melting temperature (T_M) were obtained using the optical melting procedure described previously. These parameters were calculated from the melting curve analysis and the T_M dependence plot using the Meltwin 3.0 least squares regression program (Appendix D). All successful sequences showed hyperchromicity and two state function. The melts ran from 85°C to 5°C (Figure 14). The data acquired from the melt curve analysis and T_M dependence plot from Meltwin 3.0 program, according to Freier and colleagues (1984), must agree within 15% to be consistent with the two state model (Appendix D). If the sequences met the agreement, the thermodynamic parameters were calculated to determine the influence the bulge had on the duplex stability (Table 3). The free energy of the sequence containing the bulge, which was experimentally determined, was compared to the free energy of the perfect duplex excluding the bulge nucleotide. This comparison allowed for the free energy associated with the bulge to be calculated (Table 4; Appendix E). The data was combined with previous results for Group II (G-U) bulges (Appendix F).

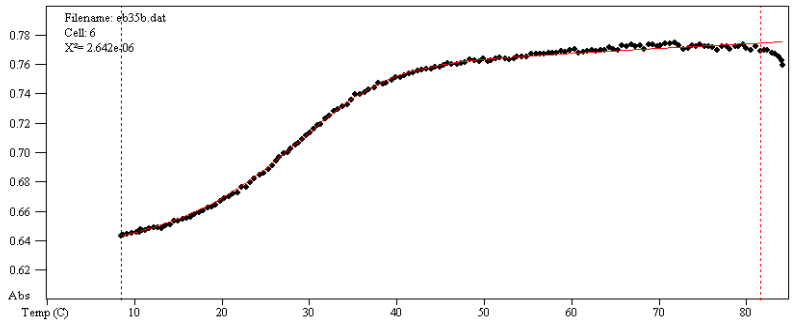


Figure 14. Example optical melt curve fitted by Meltwin 3.0 program. All eight sequences melted successfully resulted in a melting curve similar to this one. It is evident that the duplex demonstrates two state model due to the transition from double strand (lower base line) to single strand (upper base line). The blue and red dashed lines truncate trailing off points.

Table 3. Experimental thermodynamic parameters for duplex formation of Group II bulges closed by one or two Wobble base pairing(s)[a].

	Averages of Melt-Curve fits				T_M^{-1} vs. log C_T plots			
Sequences [b]	ΔH° (kcal/mol)	ΔS° (cal/K*mol)	$\Delta G_{37}°$ (kcal/mol)	T_M[c] (°C)	ΔH° (kcal/mol)	ΔS° (cal/K*mol)	$\Delta G_{37}°$ (kcal/mol)	T_M[c] (°C)
1. 5'–GUAAGUGCC– 3' 3'–CGU CACGG – 5'	-56.59	-150.41	-9.94	56.9	-58.77	-157.36	-9.96	56.2
2. 5'–GACGUAAG – 3' 3'–CUGCGU C – 5'	-49.36	-137.32	-6.77	38.5	-55.73	-157.86	-6.77	38.3
3. 5'–GAGUAAGAGC– 3' 3'–CUCGU CUCG – 5'	-47.27	-121.26	-9.67	59.0	-51.87	-135.71	-9.78	57.7
4. 5'–GUCGUUGCC– 3' 3'–CAGUA UGG – 5'	-50.44	-146.79	-4.91	27.4	-47.27	-136.48	-4.94	26.9
5. 5'–GUCGUUGCC– 3' 3'–CAGCA UGG – 5'	-55.32	-153.88	-7.59	43.1	-63.32	-179.75	-7.57	42.2
7. 5'–GGUACUAAGC– 3' 3'–CCAUGGU CG – 5'	-61.42	-168.51	-9.16	50.9	-63.29	-174.53	-9.16	50.4
8. 5'–ACUGUAAUG– 3' 3'–UGACGU GC – 5'	-50.54	-143.30	-6.10	34.4	-44.75	-124.43	-6.16	34.5

[a]The thermodynamic parameters for each duplex formation of Group II bulges closed by one or two Wobble base pairings were determined through optical melting analysis. [b]The nucleotide not paired, along with the identical nucleotide adjacent to it, are the possible Group II bulge (numbered according to raw data Appendix D). [c]The thermodynamic values ΔH°, ΔS°, $\Delta G_{37}°$, and T_M, were calculated using Meltwin 3.0 least square regression program.

Table 4. Gibb's Free Energy Associated with Group II (G-U) Single Nucleotide Bulges.

Oligomer[a]	Bulge	$\Delta°G_{37,\ bulge}$[b] (kcal/mol)
GUAAGUGCC CGU CACGC	A	2.58
GACGUAAG CUGCGU C	A	1.68
GAGUAAGAGC CUCGU CUCG	A	4.05
GUCGUUGCC CAGUA UGG	U	5.11
GUCGUUGCC CAGCA UGG	U	4.29
GGUACUAAGC CCAUGGU CG	A	4.43
ACUGUAAUG UGACGU GC	A	2.04

[a]The nucleotides in red are the bulge residues. The wobble base pairs are highlighted in green. [b]The free energy associated with the bulge was calculated using the experimental thermodynamic parameters and the nearest model by Xia et al. 1998 and Mathews et al. 1999 (Appendix E). [c] The different bulge position options for the Group II sequences observed in this investigation. [d]The average free energy was calculated because the exact position of the bulge is unknown for Group II bulges.

The data gathered focused on Gibb's free energy in order to determine the influence the bulge had on the duplex. There is too much variation in enthalpy and entropy to draw any conclusions. The influence of the bulge (Table 4) was calculated using the nearest neighbor model $\Delta G°_{37,bulge} = \Delta G°_{37,measured} - \Delta G°_{37,duplex}$. The positive free energies were expected due to the destabilizing effect an insertion of a bulge has on a duplex. The range of free energy values were also expected due to the positioning of the bulge sequence in the helix. The insertion of the bulge closer to the terminus is less destabilizing than in the middle of the duplex, as expected.

For a more complete understanding of these sequences 5'-CUGCGUC-3' and 5'-GGUAUGAC-3' single strands were melted to see how the absorbance of a single strand compares to the duplex when heated. The single strands should show little to no

transition (Figure 15). These melts will verify that no other interactions are happening with the duplex, such as hairpin formations.

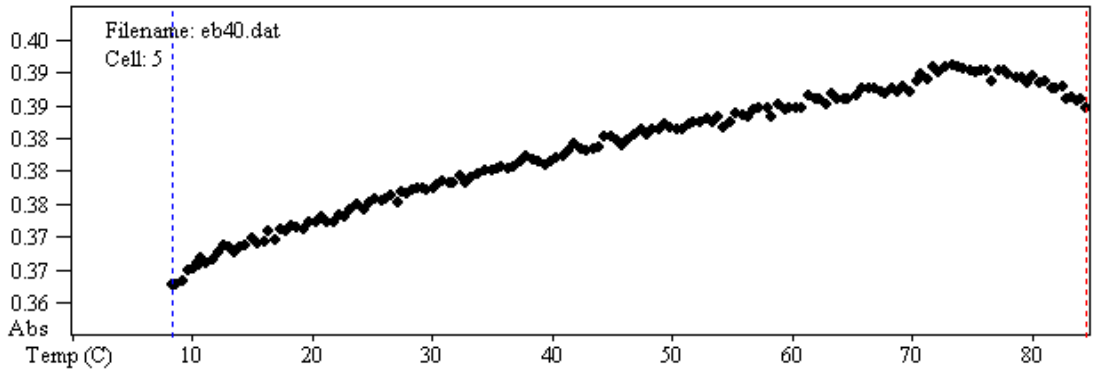


Figure 15. Single strand melt curve data with no transition. The data is from Meltwin 3.0 program and the points on the graph represent no transitional shift in the curve. Thermodynamic parameters can not be calculated from this graph.

D I S C U S S I O N

Previously, thermodynamic models have been developed in order to predict the stability of sequences containing a single nucleotide bulge (Longfellow *et al*. 1990; Znosko *et al*. 2002; Blose *et al.* 2007). Znosko and colleagues (2002) proposed an expanded nearest neighbor model which predicted the free energy of Group I and Group II single nucleotide bulges, by looking at the nearest neighbors and the identity of the bulge. A more recent study to revise the penalty model for Group I, II, and III single nucleotide bulges was completed because the addition of Group III classification of bulge loops, which reassigned some of the Group I and II bulges as Group III. As a consequence, the proposed model by Znosko *et al.* (2002) did not take into account the effect the wobble base pair had on stability nor was the orientation of the bulge analyzed (Manni unpublished 2005). Blose and colleagues (2007) were able to use a comprehensive set of Group I bulge free energies to develop the non-nearest neighbor

model, which improved the accuracy of predicting stability of duplexes containing Group I single nucleotide bulges.

The non-nearest neighbor model has not yet been determined as an accurate model for predicting the stability of duplexes containing Group II bulge loops. Group II bulges adjacent to Watson-Crick base pairings have been characterized. Therefore, Group II bulges adjacent to a wobble base pair (Group II (G-U)) have to be characterized in order to have a more complete data set. This investigation focused on different Group II bulges that are closed by either one or two wobble base pairing(s). The bulge sequences studied varied in orientation and positioning within the duplex. The results from this investigation, combined with other Group II (G-U) results, (Serra unpublished) were analyzed to identify the characteristics of Group II (G-U) bulges that influence the stability of the duplex. The non-nearest neighbor model for Group I bulges found that the identity of the bulge and the nearest neighbors had no determining factor on the stability of the duplex. However, a correlation between the least stable stem and the destabilization effect of the bulge was found. Therefore, bulge identity, nearest neighbors, and least stable stem were analyzed in this study to determine if the non-nearest neighbor model could be used to predict stability of Group II bulges adjacent to a wobble (G-U) base pair.

Analysis of Bulge Identity

The data gathered from this investigation was combined with other Group II (G-U) sequences previously studied (Table 5; Serra unpublished). GraphPad Instat was used to analyze if the identity of the bulge, purine or pyrimidine, affected the stability of the duplex (Table 6). The values were found to be statistically the same; therefore, the

analysis process could proceed combining Group II (G-U) bulges as one data set, similar to Group I data in the Blose *et al.* (2007) study.

Table 6. Average Gibb's Free Energy of all Group II (G-U) bulges.

Group II (G-U)	**$\Delta G°_{average}$**
Purine	1.4 ± 2.0 (kcal/mol)
Pyrimidine	2.0 ± 2.1 (kcal/mol)
Combined	1.7 ± 2.1 (kcal/mol)

Results calculated using GraphPad Instat. The free energies were statistically the same; therefore, an average Group II (G-U) free energy was calculated.

Analysis of Less Destabilizing Effect

The average free energy associated with Group II (G-U) is 1.7 ± 2.1 kcal/mol (Table 6) compared to the average free energy associated with Group II Watson-Crick (3.5 ± 1.3 kcal/mol; Manni unpublished 2005). This shows that the insertion of a bulge next to a wobble base pair is less destabilizing. The wobble base pair disrupts the typical stacking interactions of the duplex (Blose *et al*. 2007) making the stem less stable. Thus, the insertion of an unpaired nucleotide has less destabilizing effect. When comparing all Group II (G-U) bulges to the complete Group I data set they were found to be significantly different upon completing a t-test in GraphPad Instat (P=0.0001) ($\Delta G°_{Group\ I}$ = 4.0 ± 0.8 kcal/mol, $\Delta G°_{Group\ II(G-U)}$ = 1.7 ± 2.1 kcal/mol). According to these statistics Group II bulges adjacent to a wobble base pair is less destabilizing when compared to Group I. The conformational freedom of Group II bulges is due to the bulge nucleotide being identical to the adjacent nucleotide; which was earlier found by Znosko *et al.* (2002).

Nearest Neighbor Effect

Due to the ambiguity of the bulge in Group II bulges the free energy associated with the nearest neighbors for each bulge was calculated using the nearest neighbor model (Table 7; Xia *et al.* 1998; Mathews *et al*. 1999). The two free energies calculated were averaged because without further structural studies on each of the sequences, it is impossible to know which nucleotide is the bulge. This was done in order to see if the nearest neighbors influenced the stability of the bulge. GraphPad Prism was used to graph the free energy of the nearest neighbors (NN) versus free energy of the bulge (Figure 16). After completing a t-test the nearest neighbors were found to be influential on the stability of duplexes containing a Group II (G-U) bulge because the y-intercepts were found to be statistically different (P=0.0122). Due to the positioning of the bulge sequence closer to the terminus (5' or 3') in this study, could cause the appearance that the nearest neighbors have an influence on stability. However, the least stable stem needs to be analyzed in order to make a conclusion on the nearest neighbors effect.

Table 7. Nearest neighbor $\Delta°G_{37}$ for the two possible bulge positions.

Oligomer[a] Bulge right[c]	**Nearest Neighbors (NN) [b] $\Delta°G_{37}$(kcal/mol)**	**Oligomer[a] Bulge left[c]**	**Nearest Neighbors $\Delta°G_{37}$(kcal/mol)**	**Average $\Delta°G_{37}$(kcal/mol) of NN [d].**
GUAAGUGCC CGU CACGC	AG UC -2.08	GUAAGUGCC CG UCACGC	UA GU -1.00	-1.54

[a]The oligomer is an example oligomer examined in this investigation. The row is the same oligomer with the bulge nucleotide moved. [b]The free energy was calculated using the nearest neighbor model (Xia *et al.*1998; Mathews *et al* 1999). [c]The different bulge position options for the Group II sequences observed in this investigation. [d]The average free energy was calculated because the exact position of the bulge is unknown for Group II bulges.

Based on what we know about Group I bulged loops, the least stable stem influences the insertion of a bulge as being more or less destabilizing to the duplex (Blose

et al. 2007). The bulge was positioned in the middle of the duplex for all Group I bulge sequences. In the sequences investigated containing Group II (G-U) bulges, the bulge sequences were positioned on the 5', 3', or middle of the duplex. Therefore, the influence of the nearest neighbors on the stability of the duplex could be a result of having the bulge positioned closer to the terminus of the double stranded RNA, creating a more drastic difference in the stability of the least stable stem.

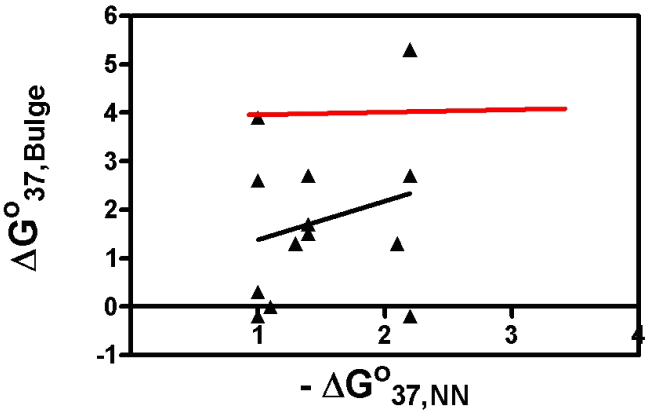


Figure 16. Least-Square Regression Analysis of Group II (G-U) Bulges. The Gibb's free energy of the nearest neighbors (NN) and free energy of the bulge ($\Delta G^\circ_{37,\ bulge}$) was plotted using GraphPad Prism to compare how the nearest neighbors affects Group II (G-U) in relation to Group I. The black line is Group II (G-U) data and the red line is Group I. The ▲ represents Group II data points.

Less Stable Stem Effect

When looking at the non-nearest neighbor model, the stability of the less stable stem adjacent to a Group I bulge was found to have a direct effect on the duplex destabilization upon bulge insertion. The free energy of the stems on either side of Group II bulges were calculated using the nearest neighbor model utilizing the Hairpin/Bulge Calculator (Visyak unpublished). In the case of Group II bulges, four stems were calculated due to the ambiguity of the bulge position (Table 8). The least negative ΔG°_{37} was considered the least stable stem.

Table 8. Example calculation of the least stable stems for Group II bulges.

$\Delta G°_{37,stem}$ [c] (kcal/mol)	Stem sequence[b]	Duplex sequence[a]	Stem sequence[b]	$\Delta G°_{37,stem}$ [c] (kcal/mol)
-7.49	GAGUA GUGCU	5'-GAGUA**A**GAGC-3' 3'-GUGCU CUCG-5'	GAGC CUCG	-7.85
-6.49	GAGU CUCG	5'-GAGU**A**AGAGC-3' 3'-GUGC UCUCG-5'	AGAGC UCUCG	-9.48

[a] Duplex sequence used here is an example of a sequence containing a Group II bulge. [b] The stems were determined by taking the duplex to the left or right of the bulge nucleotide. [c] Gibb's free energy was calculated using the Hairpin/Bulge Calculator (Visyak unpublished). The least negative value is the least stable stem. The highlighted stem sequence had the least stable free energy value.

The least stable stems were used in comparison to the $\Delta G°_{37}$ of the bulge in order to see if a similar relationship was found as the non-nearest neighbor model found for Group I bulges. The least stable stem, was plotted against the $\Delta G°_{37,bulge}$ using GraphPad Prism (Figure 17). In the analysis, the slopes of Group I and Group II are not significantly different (P=0.2937) and the intercepts were equivalent, meaning the y-intercept, 0.619 kcal/mol, could be calculated to combine Group I with Group II data (P=0.05459).

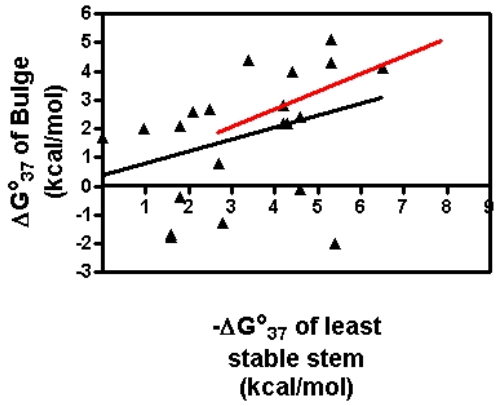


Figure 17. Least-square Regression of the least stable stem. The least stable stem was plotted with the $\Delta G°_{37,bulge}$ using GraphPad Prism to determine if the least stable stem influenced the stability of the duplex. The black line represents the best fit line for the Group II (G-U) data that is represented by ▲. The red line represents the best fit line for Group I data.

Conclusion

In this investigation the aspects of the non-nearest neighbor model were analyzed, including the effects of the identity of the bulge nucleotide and the least stable stem for Group II (G-U) bulges. The identity of the bulge was found to not have a significant affect on the stability of the duplex. Additionally, a strong correlation was found between the least stable stem and the destabilization of the duplex, similar to Group I. Therefore, the aspects of sequence stability were looked at in comparison to Group I data. Based on these results, the non-nearest neighbor model would appear to be an accurate model for determining the stability of Group II sequences adjacent to a wobble base pair.

FUTURE RESEARCH

In order for a more concrete conclusion to be made concerning the non-nearest neighbor model's accuracy for Group II (G-U) sequences, the remaining bulge sequences need to be analyzed. Sequences that need to be looked at include those that have the wobble base pair as the bulge nucleotide as well as different orientations of the adjacent wobble base pair (Appendix F). The research would include determining if there is a relationship between destabilization ($\Delta G°_{37,bulge}$) and the orientation of the wobble base pair adjacent to the bulge nucleotide; for example $\begin{matrix} 5'GAAG \\ 3'UU-U \end{matrix}$ compared to $\begin{matrix} 5'UAAU \\ 3'GU-G \end{matrix}$.

Due to the ambiguity of the bulge nucleotide in the Group II classification of single nucleotide bulge loops, final research for Group II data will be the structural analysis of the bulge in the duplex through In-line Probing. This would allow a more concrete understanding to why the least stable stem statistically relates to Group I data as well as confirm if the non-nearest neighbor model is the most accurate method to predict the stability of duplexes containing Group II bulges.

R E F E R E N C E S

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

Ehresmann C., Baudin F., Mougel M., Romby P., Ebel J., and Ehresmann B. (1987) Probing the structure of RNAs in solution. *Nucleic Acids Research*. 15, 9109-1928.

Freier S., Petersheim M., Hickey D., Turner D.H. (1984) Thermodynamic Studies of RNA Stability. *J. Biomol. Struct. Dyn*. 1, 1229-1242.

Furtig B., Richter C., Wohnert J., and Schwalbe H. (2003) NMR Spectroscopy of RNA. *ChemBioChem*. 4, 936-962.

Goldgof, Greg. 2006. RNA secondary structure prediction and runtime optimization. (Powerpoint). CS374 Presentation; Standford University. <ai.stanford.edu/~serafim/CS374_2006/presentations/lecture4.ppt>.

Hang B., Sagi J., and Singer B. (1998) Correlation between Sequence-dependent Glycosylase Repair and the Thermal Stability of Oligonucleotide Duplexes Containing 1,N^6-Ethenoadenine. *Journal of Biological Chemistry*. 273, 33406-33413.

Hetzer M., Schweyen R., Mueller M. (1997) DNA polymerization catalysed by a group II intron RNA *in vitro*. *Nucleic Acid Research*. 25, 1825-1829.

Jaeger J., SantaLucia J., and Tinoco I. (1993) Determination of RNA structure and thermodynamics. *Annual Review of Biochemistry*. 62, 255-285.

Jaffrey S., Haile D., Klausner R., and Harford J. (1993) The interaction between the iron-responsive element binding protein and its cognate RNA is highly dependent upon both RNA sequence and structure. *Nucleic Acid Research*. 21, 4627-4631.

Klapec K. (2003) *A Thermodynamic Study of Single Nucleotide Bulge Loops in Ribosomal RNA*. Allegheny College Senior Comprehensive Project. Unpublished.

Mathews D., Sabina J., Zuker M., and Turner D. (1999). Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure. *J. Mol. Biol*. 288, 911-940.

Manni M. (2005) *A Thermodynamic Study of Group I Single Nucleotide Bulge Loops with Wobble Pair Nearest Neighbors and Group II Bulges in RNA Duplexes*. Allegheny College Senior Comprehensive Project. Unpublished.

Meltwin Program. <www.meltwin.com>. 20 November 2008.

Nagaswamy U., Voss N., Zhang Z., and Fox G. (2000) Database of non-canonical base pairs found in known RNA structures. *Nucleic Acids Research*. 28, 375-376.

Puglisi J., and Tinoco I. (1989) Absorbance Melting Curves of RNA. *Methods of Enzymology*. 180, 304-325.

Roy S., Delling U., Chen C., Rosen C., and Sonenberg N. (1990) A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated *trans*-activation. *Genes & Development*. 4, 1365-1373.

Sashital D., and Butcher S. (2006) Flipping off the riboswitch: RNA structures that control gene expression. *ACS Chem. Biol*. 1, 341-345.

Serra M. J. (2009) Personal communication.

Serra M. J. (2009) Unpublished results.

Serra M.J., and Turner D.H. (1995) Predicting Thermodynamic Properties of RNA. *Methods of Enzymology*. 259, 242-260.

Turner D. (2000) Bulges in Nucleic Acids. Conformational Changes. In: Bloomfield V., Crothers D., Tinoco I., *Nucleic Acids: Structure, Properties, and Function;* University Science Books: Sausalito, CA. pp 259-334.

Varani G, and McClain W. (2000) The G-U wobble base pair. *EMBO reports*. 1, 18-23.

Vecenie C., Morrow C., Zyra A., and Serra M. (2006) Sequence Dependence of the Stability of RNA Hairpin Molecules with Six Nucleotide Loops. *Biochemistry*. 45, 1400-1407.

Weaver R.F, *Molecular Biology*; McGraw-Hill Companies, Inc: New York, NY, 2008; 4^{th} ed.

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

Yusupov M., Yusupova G, Baucom A., Lieberman K., Earnest T., Cate J., and Noller H. (2001) Crystal structure of the ribosome at 5.5 Å resolution. *Science*. 292, 883-896.

Zacharias M., and Sklenar H. (1999) Conformational Analysis of Single-base Bulges in A-form DNA and RNA using a Hierarchical Approach and Energetic Evaluation with a Continuum Solvent Model. *J. Mol. Biol*. 289, 261-275.

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

Appendix A: Derivation of Equation 1-Meltwin Fitting Equation

Sources referenced for derivation: (Serra and Turner 1995); (Serra personal communication)

Beer's Law:

$$A = \varepsilon\beta C \tag{A1}$$

Where: A is absorbance
ε is the molar absorptivity
ß is the path-length
C is the concentration

From the equilibrium expression for a two state transition: [DS]↔[SS$_1$] +[SS$_2$], where SS are non-complementary single-stranded RNA and DS is the double-stranded form of RNA:

$$A_T = A_{SS1} + A_{SS2} + A_{DS} \tag{A2}$$

Where: A_T is the total absorbance of the sample
A_{SS1} & A_{SS2} are the absorbencies of the non-complementary single-stranded RNA
A_{DS} is the absorbance of the double-stranded RNA

Combining expressions A1 and A2:

$$A_T = \varepsilon_{SS1}\beta C_{SS1} + \varepsilon_{SS2}\beta C_{SS2} + \varepsilon_{DS}\beta C_{DS} \tag{A3}$$

Where: A_T is the total absorbance of the sample
ß is the path-length of the cell
ε_{SS1} & ε_{SS2} are the extinction coefficients for single-stranded forms of RNA
ε_{DS} is the extinction coefficient for double-stranded RNA
C_{SS1} & C_{SS2} are the concentrations of single-stranded forms of RNA
C_{DS} is the concentration of the double-stranded RNA

Since there are equal molar amounts of each single-stranded RNA it can be assumed:

$$C_{SS1} = C_{SS2} = C_{DS} \tag{A4}$$

Substituting A4 into A3:

$$A = 2C_{SS}\varepsilon_{SS}\beta + \varepsilon_{DS}\beta C_{DS} \tag{A5}$$

Where: ε_{SS} is the average molar absorptivity of SS_1 and SS_2

Since the total concentration of single-stranded RNA is equal to the sum, [SS] + [DS]:

$$C_T = 2C_{SS} + 2C_{DS} \quad \text{(A6)}$$

Where: C_T is the total concentration of single-stranded RNA

Solving for C_{DS} :

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

Substituting A7 into A5 and assuming a path-length of 1 cm:

$$A = 2C_{SS}\varepsilon_{SS} + \varepsilon_{DS}(C_T / 2 - C_{SS}) \quad \text{(A8)}$$

The molar absorptivity can be expreGssed as a linear function in terms of temperature:

$$\varepsilon_{SS} = M_{SS}T + b_{SS} \quad \text{(A9)}$$

Where: M_{SS} is the slope of the linear sloping baseline for the single-stranded RNA
T is the temperature of the cell in Kelvin

$$\varepsilon_{DS} = M_{DS}T + b_{DS} \quad \text{(A10)}$$

Where: M_{DS} is the slope of the linear sloping baseline for the double-stranded RNA
T is the temperature of the cell in Kelvin
b_{DS} is the y-intercept of the linear sloping baseline for the double-stranded RNA

Substituting A9-A10 into A8:

$$A = 2C_{SS}(M_{SS}T + b_{SS}) + (M_{DS}T + b_{DS})(C_T / 2 - C_{SS}) \quad \text{(A11)}$$

Rearranging A11:

$$A = (C_{SS})(2M_{SS}T + 2b_{SS} - M_{DS}T - b_{DS}) + (C_T / 2)(M_{DS}T + b_{DS}) \quad \text{(A12)}$$

Using the expressions for Gibb's Free Energy:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \quad \text{(A13)}$$

$$\Delta G^\circ = -RT \ln K \quad \text{(A14)}$$

Where: R=1.986 cal/(mol*K)

Because the expressions as A13 and A14 are equal:

$$\Delta H^\circ - T\Delta S^\circ = -RT \ln K \tag{A15}$$

Rearranging and simplifying A15:

$$K = e^{(T\Delta S^\circ - \Delta H^\circ)/RT} \tag{A16}$$

From the equilibrium expression [DS] ↔ [SS$_1$] + [SS$_2$]:

$$K = (C_{SS1} C_{SS2})/C_{DS} \tag{A17}$$

Substituting A4 into A17:

$$C_{SS}{}^2 = KC_{DS} \tag{A18}$$

Substituting A7 into A18:

$$C_{SS}{}^2 = K(C_T/2 - C_{SS}) \tag{A19}$$

$$2C_{SS}{}^2 = KC_T - 2C_{SS} \tag{A20}$$

$$0 = 2C_{SS}{}^2 - KC_T + 2C_{SS} \tag{A21}$$

Solving A21 using the quadratic equation:

$$C_{SS} = [(K^2 + 2KC_T)^{1/2} - K]/2 \tag{A22}$$

Substituting A16 into A22:

$$C_{SS} = \{[(e^{(T\Delta S^\circ - \Delta H^\circ)/RT})^2 + 2(e^{(T\Delta S^\circ - \Delta H^\circ)/RT})C_T]^{1/2} - e^{(T\Delta S^\circ - \Delta H^\circ)/RT}\}/2 \tag{A23}$$

Substituting A23 into A12 gives the absorbance at any temperature in terms of the total concentration of the single-stranded RNA:

$$A_t = (\{[(e^{(T\Delta S^\circ - \Delta H^\circ)/RT})^2 + 2(e^{(T\Delta S^\circ - \Delta H^\circ)/RT})C_T]^{1/2} - e^{(T\Delta S^\circ - \Delta H^\circ)/RT}\}/2)(2M_{SS}T + 2b_{SS} - M_{DS}T - b_{DS}) + (C_T/2)(M_{DS}T + b_{DS})$$

(Equation 1 *)

* Used by Meltwin 3.0 to fit thermodynamic data with ΔH°, ΔS°, M_{DS}, M_{SS}, b_{SS}, and b_{DS} as adjustable parameters.

Appendix B: Derivation of Equation 2 – T_M Dependence Equation

Sources referenced for derivation: (Serra and Turner 1995); (Serra personal communication)

From the equilibrium expression $[DS] \leftrightarrow [SS_1] + [SS_2]$:

$$K = (C_{SS1}C_{SS2})/C_{DS} \tag{B1}$$

The initial concentrations of the single-stranded RNA are equal:

$$C_{SS1} = C_{SS2} \tag{B2}$$

Substituting B2 into B1:

$$K = C_{SS}{}^2/C_{DS} \tag{B3}$$

Using the expressions for Gibb's Free Energy:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \tag{B4}$$

$$\Delta G^\circ = -RT\ln K \tag{B5}$$

Because the expressions of B4 and B5 are equal:

$$\Delta H^\circ - T\Delta S^\circ = -RT\ln K \tag{B6}$$

Rearranging and simplifying B6:

$$K = e^{(T\Delta S^\circ - \Delta H^\circ)/RT} \tag{B7}$$

Substituting B4 into B7:

$$K = e^{(-\Delta G^\circ)/RT} \tag{B8}$$

Setting B8 and B3 equal to one another:

$$e^{(-\Delta G^\circ)/RT} = C_{SS}{}^2/C_{DS} \tag{B9}$$

$$\ln(C_{SS}{}^2/C_{DS}) = (-\Delta G^\circ)/RT \tag{B10}$$

Substituting B4 into B9:

$$\ln(C_{SS}{}^2/C_{DS}) = (T\Delta S^\circ - \Delta H^\circ)/RT \tag{B11}$$

At T_M the concentrations of the single and double stranded RNA are equal, $T = T_M$:

$$C_{SS1} + C_{SS2} = 2C_{SS} = C_{DS} \quad \text{(B12)}$$

Substituting B12 into B11:

$$\ln(C_{SS}{}^2 / 2C_{SS}) = (T_M \Delta S^\circ - \Delta H^\circ) / RT_M \quad \text{(B13)}$$

$$-\ln(C_{SS} / 2) = (T_M \Delta S^\circ - \Delta H^\circ) RT_M \quad \text{(B14)}$$

$$-RT_M \ln(C_{SS} / 2) = (T_M \Delta S^\circ - \Delta H^\circ) \quad \text{(B15)}$$

$$\Delta H^\circ = T_M \Delta S^\circ + RT_M \ln(C_{SS} / 2) \quad \text{(B16)}$$

$$\Delta H^\circ / T_M = \Delta S^\circ + R \ln(C_{SS} / 2) \quad \text{(B17)}$$

$$T_M{}^{-1} = (\Delta S^\circ / \Delta H^\circ) + (R / \Delta H^\circ) * \ln(C_{SS} / 2) \quad \text{(B18)}$$

At high temperatures, C_T is calculated as 2 C_{SS}:

$$T_M{}^{-1} = (\Delta S^\circ / \Delta H^\circ) + (R / \Delta H^\circ) * \ln(C_T / 4) \quad \text{(Equation 2)}$$

Appendix C: Detailed Purification Procedure

Purification Procedure (1 μM scale)

Step 1 (20 minutes)

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 (15 minutes)

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 (3:30)

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 ? (1 Hour)

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. A plunger was used to blow out the small amount of liquid which remained in the column into the last tube. 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 8 (2 Hours)

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 9 (1 Hour)

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 10 (1 Hour)

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: (Dr. Serra has greatly improved the method)

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: Group II (G-U) – Raw Data

Sequence #1

5’-GUAAGUGCC-3’
3’-CGU CACGG-5’

Melt Curve Data

Full Analysis

Date: 22 Nov 2008 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)
~~4.896e-04~~	~~-71.75~~	~~-196.49~~	~~-10.81~~	~~61.53~~
9.835e-07	-50.95	-131.66	-10.11	41.54
4.269e-06	-55.23	-146.73	-9.72	44.15
3.851e-04	-61.05	-163.75	-10.26	62.03
8.961e-06	-50.15	-130.10	-9.80	48.42
9.502e-06	-54.74	-145.61	-9.58	46.30
2.565e-04	-64.40	-174.93	-10.15	58.63
1.851e-05	-53.88	-142.48	-9.69	49.67
1.615e-05	-55.13	-146.64	-9.65	48.64
9.007e-05	-62.51	-167.64	-10.52	57.78
3.180e-05	-57.87	-154.50	-9.95	52.28
Averages:	-56.59	-150.41	-9.94	50.95

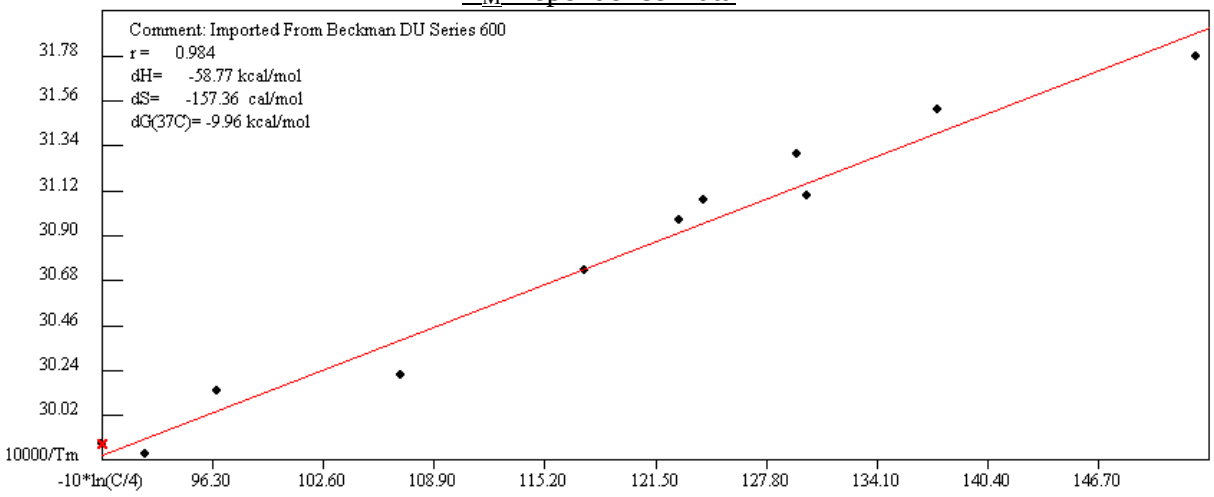
Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-56.59 ± 4.781	-150.41 ± 14.795	-9.94 ± 0.310	56.9

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-58.77 ± 3.794	-157.36 ± 11.704	-9.96 ± 0.187	56.2

T_M Dependence Data



Analysis of Melt Curve and T_M Dependence Data

	ΔH (kcal/mol)	ΔS (cal/mol*K)	ΔG (kcal/mol)	T_M (°C)
Melting Curve Fits/errors	-56.59±4.781	-150.41±14.795	-9.94±0.310	56.9
T_M Dependence/errors	-58.77±3.794	-157.36±11.704	-9.96±0.187	56.2
% Agreement	3.71%	4.42%	0.20%	

Sequence #2

5’-GACGUAAG-3’
3’-CUGCGU C-5’

Melt Curve Data

Full Analysis

Date: 22 Nov 2008 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)
7.771e-04	-52.88	-148.34	-6.87	46.71
5.173e-06	-50.35	-139.46	-7.10	29.45
6.620e-04	-44.89	-123.52	-6.58	45.60
1.703e-05	-61.38	-175.10	-7.07	34.25
1.316e-05	-50.68	-141.44	-6.81	31.17
4.224e-04	-52.08	-146.61	-6.61	42.85
3.164e-05	-48.62	-135.18	-6.70	33.58
2.320e-05	-46.59	-128.53	-6.73	32.37
1.570e-04	-48.21	-133.94	-6.67	39.71
5.016e-05	-37.96	-101.13	-6.60	34.09
Averages:	-49.36	-137.32	-6.77	36.98

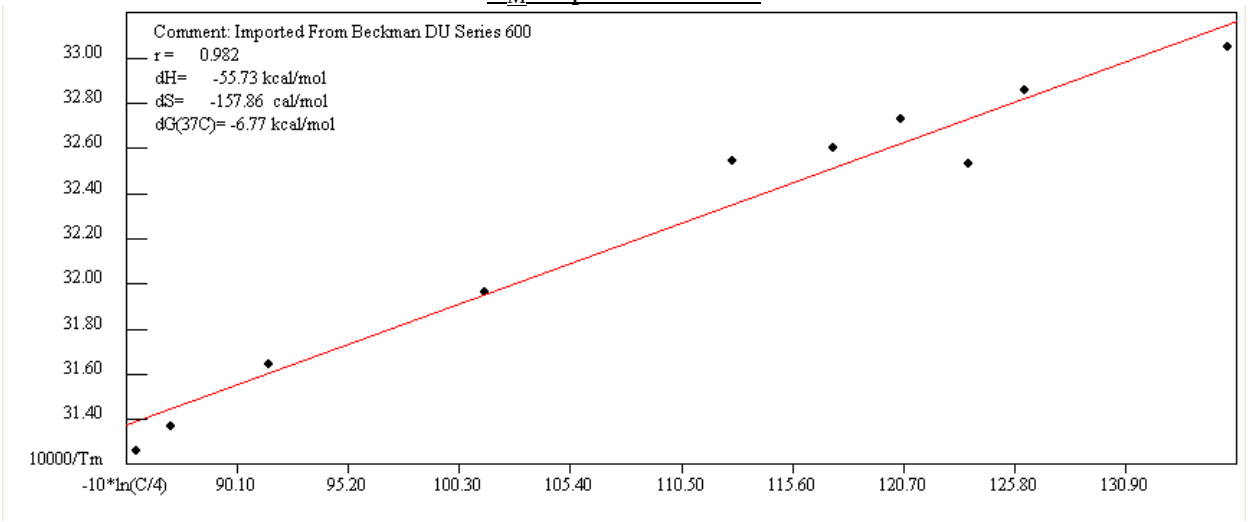
Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-49.36 ± 6.022	-137.32 ± 19.003	-6.77 ± 0.189	38.5

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-55.73 ± 3.781	-157.86 ± 12.190	-6.77 ± 0.088	38.3

T_M Dependence Data



Analysis of Melt Curve and T_M Dependence Data

	ΔH (kcal/mol)	ΔS (cal/mol*K)	ΔG (kcal/mol)	T_M (°C)
Melting Curve Fits/errors	-49.36±6.022	-137.32±19.003	-6.77±0.189	38.5
T_M Dependence/errors	-55.73±3.781	-157.86±12.190	-6.77±0.088	38.3
Agreement within	11.43%	13.01%	0.0%	

Sequence #3

5’-GAGUAAGAGC-3’
3’-CUCGU CUCG-5’

Melt Curve Data

Full Analysis

Date: 13 Mar 2009 Comment: Imported From ASCII

Melt Curve Results

Concentration (M)	dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C)
3.818e-06	-55.25	-146.18	-9.91	44.89
~~3.602e-06~~	~~-64.41~~	~~-171.97~~	~~-11.08~~	~~49.52~~
6.954e-06	-55.54	-147.00	-9.95	47.25
2.243e-06	-49.91	-129.32	-9.81	42.92
4.132e-05	-48.76	-126.19	-9.62	54.08
4.329e-06	-52.30	-136.78	-9.87	45.57
9.217e-07	-31.43	-70.75	-9.49	37.69
~~7.570e-05~~	~~-52.05~~	~~-137.96~~	~~-9.26~~	~~53.02~~
1.881e-06	-46.37	-117.98	-9.78	42.46
9.726e-05	-38.63	-95.90	-8.89	57.00
~~1.169e-06~~	~~-21.95~~	~~-39.46~~	~~-9.71~~	~~43.32~~
Averages:	-47.27	-121.26	-9.67	46.48

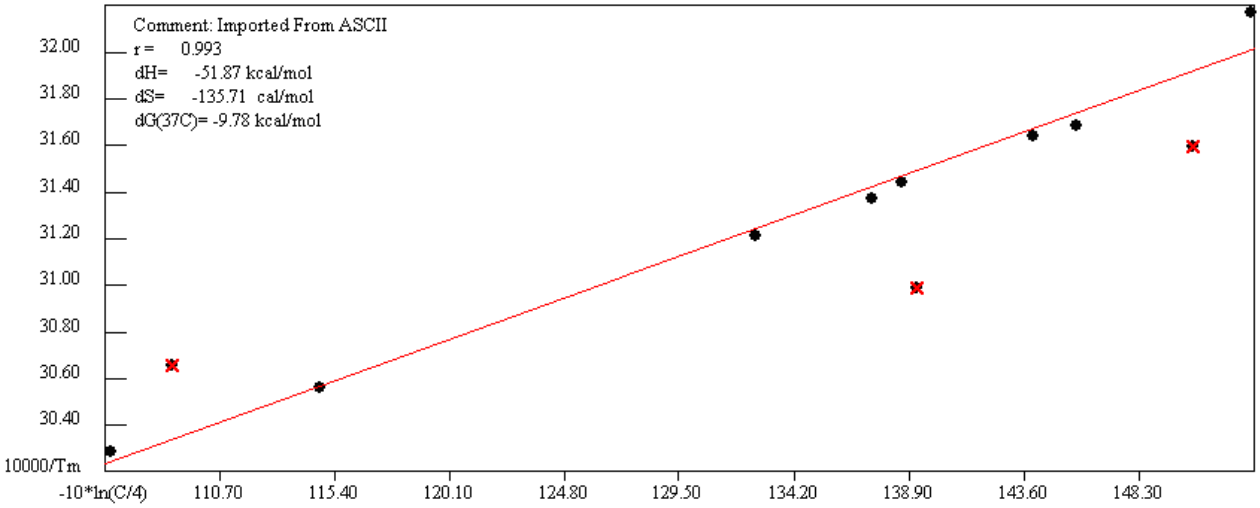
Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-47.27 ± 8.391	-121.26 ± 26.238	-9.67 ± 0.350	59.0

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-51.87 ± 2.500	-135.71 ± 7.825	-9.78 ± 0.087	57.7

T_M Dependence Data



Analysis of Melt Curve and T_M Dependence Data

	ΔH (kcal/mol)	ΔS (cal/mol*K)	ΔG (kcal/mol)	T_M (°C)
Melting Curve Fits/errors	-47.27±8.391	-121.26±26.238	-9.67±0.350	59.0
T_M Dependence/errors	-51.87±2.500	-135.71±7.825	-9.78±0.087	57.7
Agreement within	8.87%	10.65%	1.12%	

Sequence #4

5’-GGUACUAAGC-3’
3’-CCAUGGU CG-5’

Melt Curve Data

Full Analysis

Date: 27 Feb 2009 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)
1.935e-05	-43.68	-113.55	-8.46	43.66
9.970e-06	-90.40	-259.13	-10.03	44.30
3.826e-06	-49.04	-129.31	-8.94	39.52
2.021e-05	-74.91	-209.80	-9.84	46.94
9.776e-06	-58.41	-159.09	-9.06	42.95
3.269e-06	-53.74	-144.06	-9.06	39.43
2.193e-04	-55.81	-151.89	-8.70	52.46
8.805e-06	-61.83	-169.85	-9.15	42.71
3.391e-04	-64.62	-178.50	-9.26	54.66
4.080e-06	-61.78	-169.92	-9.08	39.90
Averages:	-61.42	-168.51	-9.16	44.65

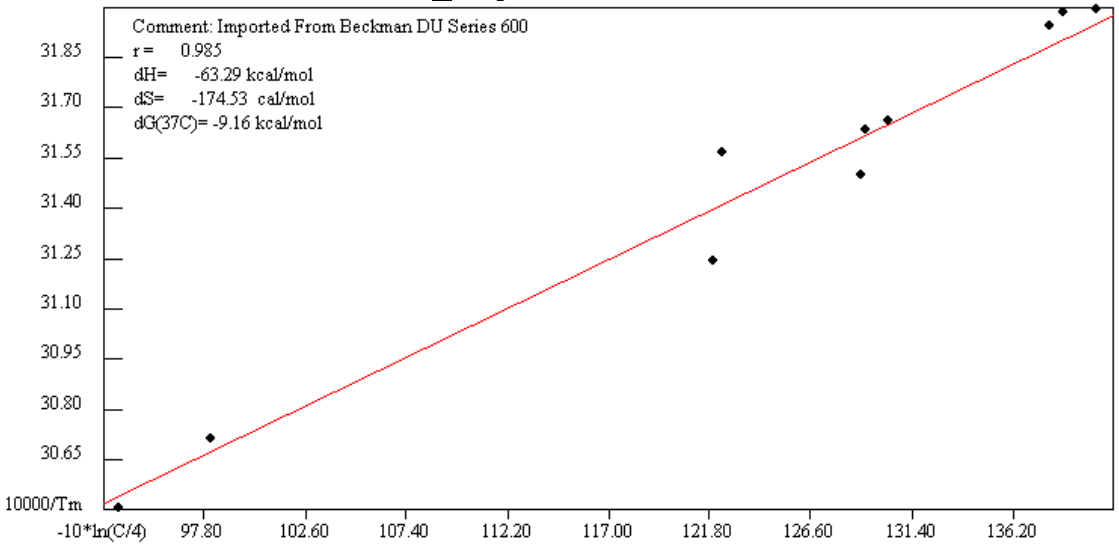
Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-61.42 ± 13.334	-168.51 ± 41.555	-9.16 ± 0.473	50.9

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-63.29 ± 3.940	-174.53 ± 12.391	-9.16 ± 0.128	50.4

T_M Dependence Data



Analysis of Melt Curve and T_M Dependence Data

	ΔH (kcal/mol)	ΔS (cal/mol*K)	ΔG (kcal/mol)	T_M (°C)
Melting Curve Fits/errors	-61.42±13.334	-168.51±41.555	-9.16±0.473	50.9
T_M Dependence/errors	-63.29±3.940	-174.53±12.391	-9.16±0.128	50.4
Agreement within	2.95%	3.45%	0.0%	

Sequence #5

5’-GUCGUUGCC-3’
3’-CAGUA UGG-5’

Melt Curve Data

Full Analysis

Date: 12 Feb 2009 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)
4.381e-04	-47.34	-136.76	-4.92	32.50
2.927e-06	-55.15	-163.49	-4.44	14.73
2.708e-04	-70.64	-213.49	-4.42	30.58
9.627e-06	-38.48	-106.36	-5.49	18.23
~~7.065e-06~~	~~-58.57~~	~~-170.28~~	~~-5.76~~	~~24.77~~
1.519e-04	-61.75	-184.06	-4.66	29.10
1.556e-05	-41.55	-117.36	-5.15	19.23
~~1.008e-05~~	~~-57.35~~	~~-165.47~~	~~-6.03~~	~~26.99~~
6.832e-05	-49.20	-142.73	-4.93	25.87
2.246e-05	-39.40	-110.02	-5.27	20.76
Averages:	-50.44	-146.79	-4.91	23.87

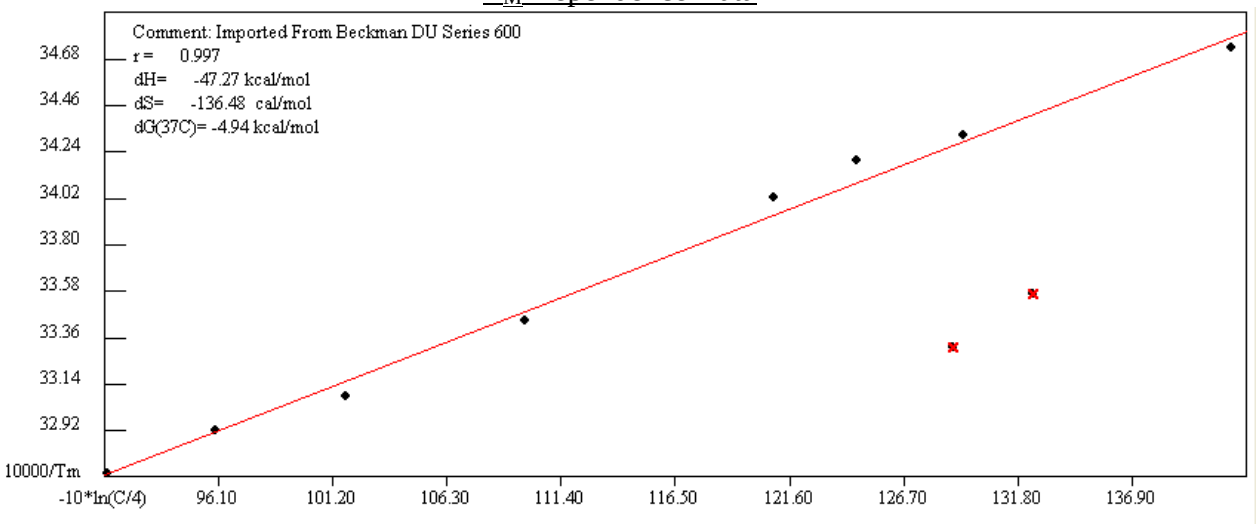
Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-50.44 ± 11.421	-146.79 ± 37.962	-4.91 ± 0.388	27.4

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-47.27 ± 1.593	-136.48 ± 5.370	-4.94 ± 0.075	26.9

T_M Dependence Data



Analysis of Melt Curve and T_M Dependence Data

	ΔH (kcal/mol)	ΔS (cal/mol*K)	ΔG (kcal/mol)	T_M (°C)
Melting Curve Fits/errors	-50.44±11.421	-146.79±37.962	-4.91±0.388	27.4
T_M Dependence/errors	-47.27±1.593	-136.48±5.370	-4.94±0.075	26.9
Agreement within	6.28%	7.02%	0.61%	

Sequence #6

5’-GUCGUUGCC-3’
3’-CAGCA UGG-5’

Melt Curve Data

Full Analysis

Date: 12 Feb 2009 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)
1.309e-05	-52.87	-145.88	-7.63	36.09
9.818e-05	-58.60	-164.49	-7.58	42.60
2.814e-05	-51.80	-143.03	-7.44	37.75
3.748e-04	-63.47	-180.26	-7.57	46.31
3.589e-06	-55.45	-153.42	-7.87	33.06
3.189e-04	-64.60	-183.61	-7.66	46.10
7.075e-06	-41.46	-108.61	-7.77	34.09
2.321e-04	-63.99	-181.94	-7.56	44.68
1.505e-05	-48.34	-131.96	-7.41	35.17
1.309e-05	-52.87	-145.88	-7.63	36.09
9.818e-05	-58.60	-164.49	-7.58	42.60
2.814e-05	-51.80	-143.03	-7.44	37.75
Averages:	-55.32	-153.88	-7.59	39.36

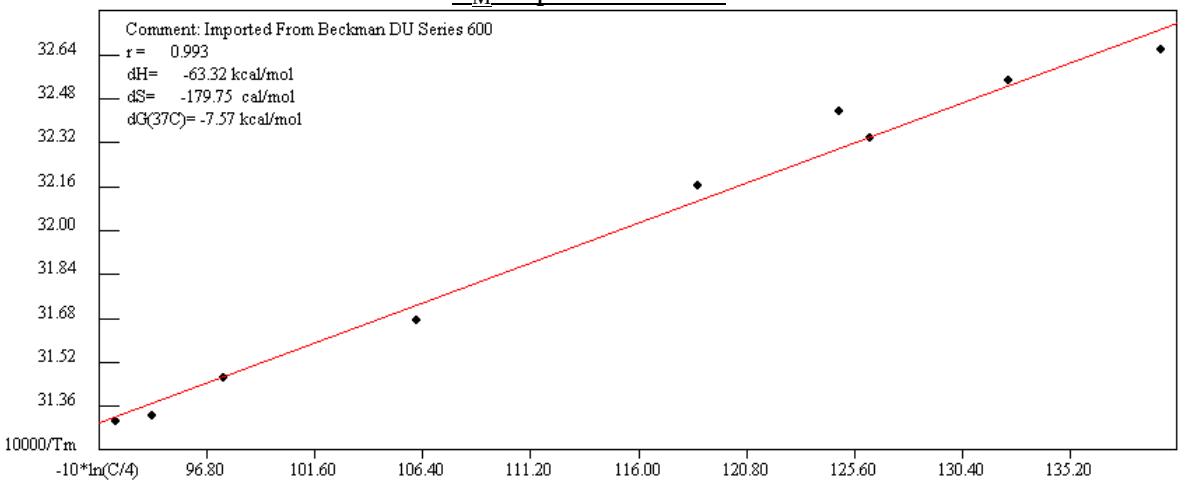
Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-55.32 ± 6.921	-153.88 ± 22.330	-7.59 ± 0.134	43.1

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K-mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-63.32 ± 2.401	-179.75 ± 7.687	-7.57 ± 0.037	42.2

T_M Dependence Data



Analysis of Melt Curve and T_M Dependence Data

	ΔH (kcal/mol)	ΔS (cal/mol*K)	ΔG (kcal/mol)	T_M (°C)
Melting Curve Fits/errors	-55.32±6.921	-153.88±22.330	-7.59±0.134	43.1
T_M Dependence/errors	-63.32±2.401	-179.75±7.687	-7.57±0.037	42.2
Agreement within	12.63%	14.39%	0.26%	

Sequence #7

5’-GCCGUUGC-3’
3’-CGGUA UG-5’

Data collected from optical melt did not have consistent hyperchromicity nor did it adhere to the two state model.

Sequence #8

5’-ACUGUAAUG-3’
3’-UGACGU GC-5’

Melt Curve Data

Full Analysis

Date: 26 Mar 2009 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)
5.321e-04	-64.57	-187.37	-6.45	41.65
2.190e-06	-47.39	-132.94	-6.16	20.15
~~3.376e-06~~	~~-68.47~~	~~-205.54~~	~~-4.72~~	~~20.31~~
3.864e-04	-53.25	-151.21	-6.35	40.86
1.139e-05	-51.53	-147.50	-5.79	24.96
8.182e-06	-45.39	-127.18	-5.95	23.12
2.437e-04	-49.24	-138.56	-6.26	38.76
2.135e-05	-45.73	-127.75	-6.10	27.93
1.476e-05	-50.83	-145.12	-5.82	25.88
9.253e-05	-44.76	-124.47	-6.16	34.13
3.414e-05	-52.72	-150.91	-5.91	29.64
Averages:	-50.54	-143.30	-6.10	30.71

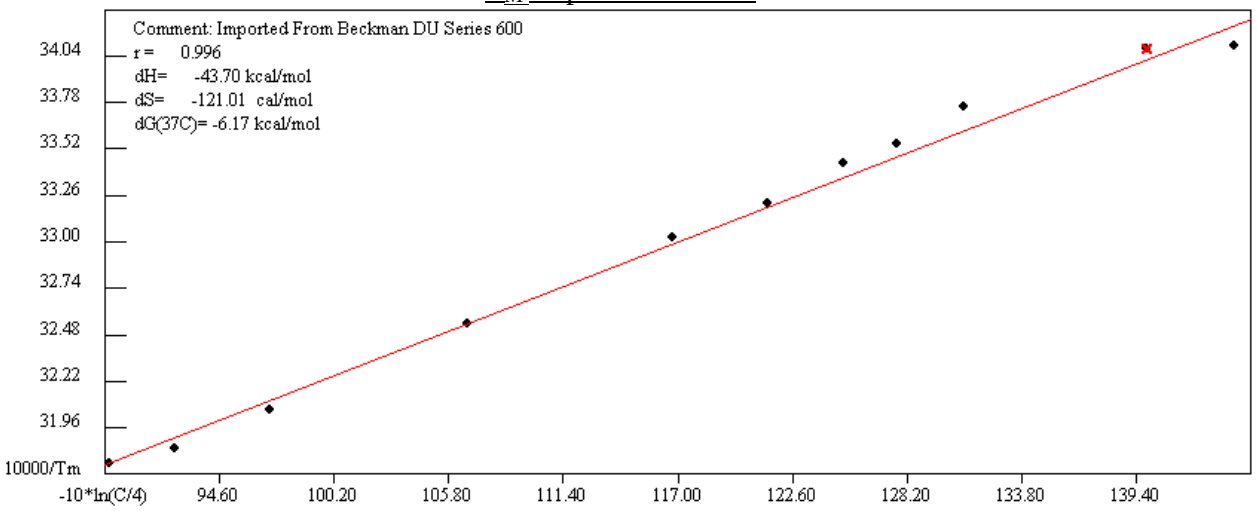
Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-50.54 ± 5.817	-143.30 ± 18.468	-6.10 ± 0.225	34.4

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-43.70 ± 1.330	-121.01 ± 4.386	-6.17 ± 0.035	34.4

T_M Dependence Data



Analysis of Melt Curve and T_M Dependence Data

	ΔH (kcal/mol)	ΔS (cal/mol*K)	ΔG (kcal/mol)	T_M (°C)
Melting Curve Fits/errors	-50.54±5.817	-143.30±18.468	-6.10±0.225	34.4
T_M Dependence/errors	-44.75±1.191	-124.43±3.929	-6.16±0.032	34.5
Agreement within	11.46%	13.17%	0.97%	

Sequence #9

5’-GACUAAUGUC-3’
3’-CUGGU GCAG-5’

Melt Curve Data

Melt Curve Results

Concentration (M)	dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C)
4.920e-04	-63.52	-183.22	-6.69	42.69
~~3.312e-06~~	~~-32.22~~	~~-74.16~~	~~-9.22~~	~~42.79~~
3.712e-04	-47.09	-131.31	-6.36	41.28
8.698e-06	-20.67	-41.63	-7.76	32.86
7.880e-06	-20.93	-43.28	-7.51	28.54
2.176e-04	-41.53	-113.46	-6.34	39.14
1.866e-05	-26.26	-61.37	-7.23	33.09
1.391e-05	-35.08	-89.37	-7.36	33.61
5.715e-05	-38.90	-103.70	-6.73	35.88
3.112e-05	-31.27	-78.79	-6.84	32.96
4.938e-04	-41.80	-114.11	-6.41	43.58
~~2.939e-06~~	~~-16.75~~	~~-27.08~~	~~-8.35~~	~~30.64~~
3.973e-06	-26.60	-60.10	-7.96	30.65
3.946e-04	-46.80	-130.70	-6.26	40.89
1.025e-05	-26.15	-60.31	-7.45	31.33
~~8.838e-06~~	~~-24.40~~	~~-56.61~~	~~-6.84~~	~~22.66~~
2.378e-04	-59.72	-172.09	-6.35	38.85
1.897e-05	-34.93	-89.41	-7.20	33.86
1.504e-05	-30.25	-73.01	-7.61	36.07
9.332e-05	-47.53	-131.27	-6.82	38.60
3.065e-05	-37.22	-96.05	-7.43	38.45
Averages:	-37.57	-98.51	-7.02	36.24

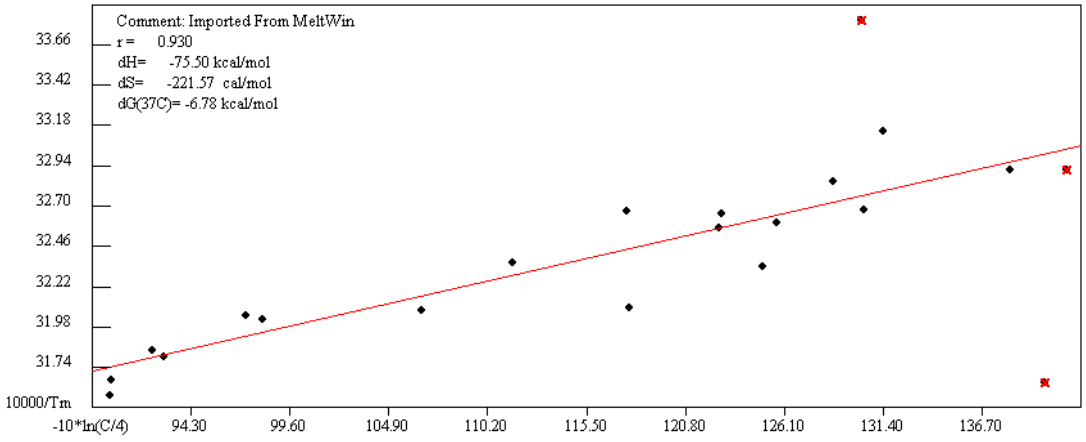
Analysis of Melt Curve Fits/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-37.57 ± 12.180	-98.51 ± 40.642	-7.02 ± 0.548	41.1

Analysis of Tm Dependence/Errors

dH (kcal/mol)	dS (cal/K·mol)	dG (kcal/mol, 37°C)	Tm (°C, 1.0e-4M)
-75.50 ± 7.476	-221.57 ± 24.046	-6.78 ± 0.258	38.0

T_M Dependence Data



Analysis of Melt Curve and T_M Dependence Data

	ΔH (kcal/mol)	ΔS (cal/mol*K)	ΔG (kcal/mol)	T_M (°C)
Melting Curve Fits/errors	-37.57±12.180	-98.51±40.642	-7.02±0.548	41.1
T_M Dependence/errors	-75.50±7.476	-221.57±24.046	-6.78±0.258	38.0
Agreement within	50.24%	55.54%	3.42%	

Appendix E: Sample Calculation – Gibb's Free Energy Associated with Bulge ($\Delta G°_{37,\,bulge}$)

Sample Calculation to Predict Thermodynamic Stability of Group II (G-U) bulges

Sequence: 5'-GUA**A**GUGCC -3'
3'-CGU CACGG-5'

1. The stability of the perfect duplex, without the bulge (red letter A), is calculated using the INN-HB nearest neighbor model (Xia *et al.* 1998; Mathews *et al.* 1999) as follows:

Nearest Neighbors	$\Delta G°_{37}$ (kcal/mol)
GU CG	-2.51
UA GU	-1.00
AG UC	-2.08
GU CA	-2.24
UG AC	-2.11
GC CG	-3.42
CC GG	-3.26
Initiation	+4.09
$\Delta G°_{37}$ =	-12.53 kcal/mol

2. In order to determine the influence of the bulge the $\Delta G°_{37}$ is averaged from the melt curve data and T_M dependence data. The experimental data was found to be:

Analysis of Melt Curve: $\Delta G°_{37}$ = -9.94 kcal/mol
Analysis of T_M Dependence Data: $\Delta G°_{37}$ = -9.96 kcal/mol
Average $\Delta G°_{37}$: (-9.94 + -9.96)/2 = -9.95 kcal/mol

3. The $\Delta G°_{37}$ for this Group II (G-U) bulge was found by subtracting the $\Delta G°_{37}$ of the duplex without the bulge (calculated from INN-HB model) from the $\Delta G°_{37}$ for the duplex with the bulge (experimentally determined).

$\Delta G°_{37,bulge}$ = -9.95 – (-12.53) = 2.58 kcal/mol

Appendix F. Collective Group II (G-U) Thermodynamic parameters

bulge sequence	$\Delta G^{o}_{37bulge}$ [a] (kcal/mol)	ΔH^{o}_{bulge} [a] (kcal/mol)	Percentage of naturally occurring group II sequences	bulge sequence	$\Delta G^{o}_{37bulge}$ [a] (kcal/mol)	ΔH^{o}_{bulge} [a] (kcal/mol)	Percentage of naturally occurring group II sequences
C**AA**U GU G			0.9	A**CC**G UG U	1.5	24.7	0.0
G**AA**U CU G			0.0	G**CC**G CG U	-2.0	2.3	0.5
U**AA**U AU G			0.0	U**CC**G AG U			0.0
U**AA**U GU G	2.0	11.3	9.2	G**CC**G UG U			0.0
U**AA**C GU G			11.5	G**CC**A UG U	2.2	29.6	0.5
U**AA**G GU C	2.6, 1.7 4.1,4.4 2.7, 1.7	18.9, 8.4 37.6, 22.4 18.9, 8.4	50.2	G**CC**G UG C			1.8
U**AA**U GU A			0.9	G**CC**U UG A			0.0
C**GG**U GC G			0.5	A**UU**G UA U			0.0
U**GG**U AC G			0.0	G**UU**G CA U	4.3	15.9	0.0
U**GG**U GC G			5.5	G**UU**G UA U	5.1	18.8	1.4
C**GG**U GU G			0.0	A**UU**G UG U			0.0
U**GG**U AU G	-1.3,-0.4 -1.7, -1.8	-3.8, 6.1 18.0, 2.9	0.0	G**UU**G CG U			0.0
U**GG**U GU G			0.0	G**UU**G UG U			0.9
U**GG**C GU G			0.5	G**UU**A UG U			1.8
U**GG**U GU A	0.8	-13.1	0.0	G**UU**G UG C	2.4 2.2	23.4 6.0	1.4
U**GG**C GC G			0.9	G**UU**A UA U			1.8
U**GG**U GC A			1.4	G**UU**G UA C			0.0
C**GG**C GU G	4.0 0.8	1.3 29.2	0.9	A**UU**A UG U	-0.1	15.4	4.6
C**GG**U GU A			0.0	A**UU**G UG C			1.4
U**GG**C AU G	2.1	11.2	0.0	G**UU**A CG U	2.8	16.9	0.0
U**GG**U AU A			0.0	G**UU**G CG C			1.4
U**GG**U AU G			0.0				

[a]Values calculated from experimental data and perfect duplex data utilizing the nearest neighbor model. Values in blue represent the sequences that were completed in this study. Other results were from Serra (unpublished).