

Stability of single-nucleotide bulge loops embedded in a GAAA RNA hairpin stem

GEOFFREY F.S. LIM, GREGORY E. MERZ, MICHAEL D. MCCANN, JOCELYN M. GRUSKIEWICZ, and MARTIN J. SERRA¹

Department of Chemistry, Allegheny College, Meadville, Pennsylvania 16335, USA

ABSTRACT

Forty-six RNA hairpins containing combinations of 3' or 5' bulge loops and a 3' or 5' fluorescein label were optically melted in 1 M NaCl, and the thermodynamic parameters ΔH° , ΔS° , ΔG_{37}° , and T_M for each hairpin were determined. The bulge loops were of the group I variety, in which the identity of the bulge is known, and the group II variety, in which the bulged nucleotide is identical to one of its nearest neighbors, leading to ambiguity as to the exact position of the bulge. The fluorescein label at either the 3' end or 5' end of the hairpin did not significantly influence the stability of the hairpin. As observed with bulge loops inserted into a duplex motif, the insertion of a bulge loop into the stem of a hairpin loop was destabilizing. The model developed to predict the influence of bulge loops on the stability of duplex formation was extended to predict the influence of bulge loops on hairpin stability. Specifically, the influence of the bulge is related to the stability of the hairpin stem distal from the hairpin loop.

Keywords: secondary structure; thermodynamics

INTRODUCTION

The stability of fully base-paired duplex regions of RNA has been extensively investigated (Turner et al. 1988; Xia et al. 1998). To develop the nearest neighbor model to predict the secondary structure of RNA from sequence, the thermodynamics of various single-stranded RNA motifs (helices, bulges, internal loops, hairpins, dangling ends, multibranch loops, etc.) were studied in short RNA oligonucleotides. Based partially on these thermodynamic parameters, an algorithm was developed to predict secondary structure from sequence, and it currently predicts 73% of known base pairs (Mathews et al. 1999). The accuracy of the prediction is based on the measured thermodynamics of the various motifs. However, to fully use the rapid accumulation of nucleic acid sequence information, further refinements in structure prediction are necessary (Mathews et al. 1999, 2004; Andronescu et al. 2007).

One of the more common secondary structural motifs, a bulge loop, occurs when a duplex is interrupted by one or more unpaired nucleotides in one of the strands. Bulge loops are important in many biological processes, including protein binding (Peattie et al. 1981; Gott et al. 1991; Harper

and Logsdon 1991; Lilley 1995), intron splicing (Schmelzer and Schweyen 1986; Parker et al. 1987), feedback regulation (Climie and Friesen 1987), tertiary folding (Woese and Gutell 1989), and miRNA:mRNA interactions (Cevec et al. 2010).

Recently the thermodynamic stability of the complete set of duplexes with the group I (Blose et al. 2007) and group II bulge loops (McCann et al. 2011) has been determined. Group I bulge loops are those in which the bulged nucleotide is not identical to either of its nearest neighbor nucleotides (Fig. 1). Group II bulge loops are identical to at least one of their nearest neighbors, leading to ambiguity as to the exact position of the bulge (Fig. 1). The free-energy increment for the insertion of a bulge into an RNA duplex was primarily influenced by non-nearest-neighbor interactions. In particular, the stability of the stems adjacent to the bulged nucleotide influences the impact of the bulged nucleotide on the stability of the duplex. From these results, a simple model could be derived to predict the thermodynamic contribution (in kilocalories per mole) for inserting a bulge loop into a duplex:

$$\Delta G_{37}^\circ \text{ bulge loop} = -0.53 \Delta G_{\text{stem}}^\circ + 0.7 \quad (1)$$

where $\Delta G_{\text{stem}}^\circ$ is the less stable stem for group I bulge loops and the second least stable stem for group II bulge loops. In the investigation of the group II bulge loops, in-

¹Corresponding author.

E-mail mserra@allegheny.edu.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.028308.111>.

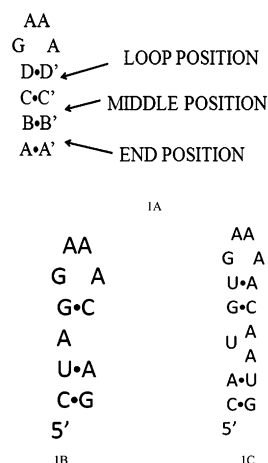


FIGURE 1. Secondary structure of hairpins. (A) N•N' represents base pairs, and End, Middle, and Loop Positions indicate the positions of bulge loop insertion. (B) Example hairpin with group I bulge loop on the 5' side, middle position of the stem. (C) Example hairpin with group II bulge loop on the 3' side, middle position of the stem.

line structure mapping was used to investigate the structural ambiguity of the bulge. The bulged nucleotide was found to be the one positioned further from the hairpin loop. To carry out the in-line probing, the bulge loop needed to be inserted into a hairpin stem (McCann et al. 2011). It is, therefore, important to determine how the insertion of a bulge loop into the stem of a hairpin influences the stability of hairpin formation. This is especially true because nearly 70% of rRNA is found in small hairpin structures (Gutell 1993; Gutell et al. 1993). In this investigation, we examined the influence of inserting a bulge loop into a hairpin stem and found that the model developed for the insertion of bulge loops into duplexes can be extended to predict the influence of bulge loops into hairpin stems.

RESULTS

The thermodynamics for hairpin formation for 46 RNA hairpins containing combinations of 3' or 5' group I or II bulge loops and a 3' or 5' fluorescein label are listed in Table 1. The fluorescein label was added to the ends of the RNA to allow visualization of the products generated by in-line probing (McCann et al. 2011).

Influence of fluorescein tag on hairpin stability

Eight of the hairpins in Table 1 were melted with or without a 3' fluorescein label. The average difference in the stability of the eight hairpins is 0.2 ± 0.3 kcal/mol, a difference within the error of the experiment. Likewise, eight hairpins in Table 1 were melted with or without a 5' fluorescein label. The average difference in the stability of these eight hairpins is 0.0 ± 0.6 kcal/mol. Therefore, in the remainder of the analysis, the presence of a fluorescein label will be ignored.

Influence of the GAAA tetraloop on hairpin stability

The results of the melts of a total of 20 parent hairpins (14 without a fluorescein label and six with a fluorescein label) are presented in Table 1. The free-energy contribution of the hairpin loop can be determined by:

$$\Delta G_{37}^{\circ}(\text{loop}) = \Delta G_{37}^{\circ}(\text{measured for hairpin formation}) - \Delta G_{37}^{\circ}(\text{stem})$$

where $\Delta G_{37}^{\circ}(\text{stem})$ is predicted from the Watson-Crick nearest neighbors (Xia et al. 1998). The hairpin loop values are presented in Table 2 and range from 3.2 to 4.8 kcal/mol with an average value of 4.1 kcal/mol.

Influence of group I and group II single-nucleotide bulge loops embedded in a hairpin stem

Group I and II single-nucleotide bulge loops were inserted into each of the three positions (end, middle, or loop) on either the 5' or 3' side of the hairpin loop (Fig. 1). The influence of the stability of the hairpin was determined by subtraction of the stability of the hairpin with or without the bulge loop:

$$\Delta G_{37}^{\circ}(\text{bulge}) = \Delta G_{37}^{\circ}(\text{bulge hairpin}) - \Delta G_{37}^{\circ}(\text{parent hairpin})$$

The bulge loop values are presented in Table 2. The influence of the group I and II bulges ranges from 0.5 to 2.8 kcal/mol at 37°C. For both cases, the free-energy increment for insertion of the bulge loop varies by position but not by whether the bulge is to the 3' or 5' side of the loop. Therefore, the free-energy increments for insertion of each type of bulge loop were averaged by position, and the averages for group I and group II bulges are presented in Tables 3 and 4, respectively.

The insertion of group I or II bulge loops into a hairpin stem makes the enthalpy of hairpin formation less favorable, as was also observed for duplexes (Znosko et al. 2002; Blose et al. 2007). The insertion of a bulge at the end position makes the enthalpy less favorable by 6.7 ± 3.1 and 6.1 ± 3.1 kcal/mol, respectively, for group I and group II bulges. In contrast, the changes averaged over the middle and loop positions are larger: 16.2 ± 4.6 and 14.8 ± 5.5 kcal/mol, respectively, for group I and group II bulges.

DISCUSSION

Influence of fluorescein tag on hairpin stability

In general, a 3' dangling base provides more stabilization than a 5' dangling base on an RNA helix. In many cases, the increased stabilization caused by a dangling end is as large as the stabilization of a base pair. This has been rationalized as the better stacking of a 3' dangling base on the terminal

TABLE 1. Thermodynamic parameters for hairpin formation^a

oligomer ^b	+fluorescein tag				- fluorescein tag			
	ΔH° (kcal/mol)	ΔS° (cal/K•mol)	ΔG°_{37} (kcal/mol)	T_M (°C)	ΔH° (kcal/mol)	ΔS° (cal/K•mol)	ΔG°_{37} (kcal/mol)	T_M (°C)
Parent Hairpin								
CGAUG <u>AAAA</u> AUCG					-32.2 ± 3.1	-98.3 ± 9.6	-1.7 ± 0.2	54.2
CAGUG <u>AAAA</u> ACUG					-30.7 ± 3.8	-93.7 ± 11.8	-1.7 ± 0.2	54.8
CUAG <u>AAAA</u> CUAG					-31.5 ± 3.4	-96.7 ± 10.5	-1.6 ± 0.3	53.0
CUAUG <u>AAAA</u> CAUAG					-41.3 ± 3.8	-124.4 ± 11.3	-2.7 ± 0.3	58.5
GUAC <u>AAAA</u> GUAC					-38.3 ± 2.0	-115.4 ± 6.2	-2.5 ± 0.1	58.7
CUGUG <u>AAAA</u> ACAG					-35.9 ± 2.7	-109.1 ± 9.2	-1.8 ± 0.1	54.1
GGUC <u>AAAA</u> GACC					-42.0 ± 3.5	-121.1 ± 10.3	-4.4 ± 0.3	73.5
CAUG <u>AAAA</u> CAUG					-42.4 ± 4.4	-130.0 ± 13.3	-2.1 ± 0.3	53.1
CUGAUG <u>AAAA</u> AUCAGF	-34.7 ± 4.6	-103.4 ± 14.1	-2.6 ± 0.3	62.5	-36.7 ± 3.4	-109.7 ± 10.5	-2.6 ± 0.3	61.1
CAUGUG <u>AAAA</u> CAUGF	-35.3 ± 3.3	-103.9 ± 10.1	-3.1 ± 0.3	66.5	-36.8 ± 4.1	-110.5 ± 12.4	-2.5 ± 0.3	59.6
CAUUG <u>AAAA</u> CAAUGF	-41.7 ± 2.9	-125.2 ± 8.7	-2.9 ± 0.2	60.2	-41.6 ± 2.2	-126.2 ± 6.5	-2.5 ± 0.3	56.4
FCAGAU <u>AAAA</u> AUCUG	-30.1 ± 5.1	-88.8 ± 15.9	-2.6 ± 0.2	66.6	-36.5 ± 4.0	-109.2 ± 12.0	-2.6 ± 0.3	61.2
FCUAGUG <u>AAAA</u> CUAG	-32.7 ± 3.9	-97.4 ± 11.9	-2.5 ± 0.2	66.9	-36.5 ± 4.6	-108.9 ± 13.7	-2.7 ± 0.4	61.8
FCAUAGG <u>AAAA</u> CUAUG	-24.8 ± 3.8	-72.3 ± 12.4	-2.4 ± 0.2	69.7	-43.9 ± 4.7	-132.1 ± 14.1	-2.9 ± 0.3	59.2
Group I bulge hairpin								
CGAUG <u>AAAA</u> AUCAGF	-24.8 ± 2.7	-75.9 ± 8.7	-1.2 ± 0.1	53.0				
GGUC <u>AAAA</u> GACACF	-36.5 ± 4.7	-105.0 ± 13.9	-3.9 ± 0.4	74.1				
CAGUG <u>AAAA</u> CAUGF	-18.7 ± 4.1	-59.8 ± 13.1	-0.1 ± 0.3	39.2	-15.1 ± 2.8	-47.2 ± 9.0	-0.4 ± 0.2	45.7
CUAG <u>AAAA</u> CAUGF	-14.3 ± 3.2	-47.3 ± 10.6	0.4 ± 0.2	28.4	-17.6 ± 1.5	-58.8 ± 5.0	0.7 ± 0.1	25.6
FCAGAU <u>AAAA</u> AUCG	-20.7 ± 5.3	-63.4 ± 16.2	-1.0 ± 0.3	53.1	-29.7 ± 1.6	-92.1 ± 5.2	-1.1 ± 0.3	49.2
FCUAGUG <u>AAAA</u> ACAG	-13.7 ± 2.7	-41.8 ± 10.0	-0.7 ± 0.4	53.2	-22.0 ± 5.1	-71.1 ± 17.0	0.1 ± 0.2	36.3
FCAUAGG <u>AAAA</u> CAUG	-15.5 ± 6.2	-49.6 ± 20.5	-0.1 ± 0.5	39.4				
Group II bulge hairpin								
CUGAUG <u>AAAA</u> AUCAAGF	-28.1 ± 5.1	-84.8 ± 16.4	-1.8 ± 0.3	57.8	-28.1 ± 3.3	-85.4 ± 9.8	-1.7 ± 0.3	56.3
GUAC <u>AAAA</u> GUAACF	-29.0 ± 4.3	-88.8 ± 13.6	-1.4 ± 0.4	53.1				
CAUGUG <u>AAAA</u> CAAUGF	-21.4 ± 3.3	-66.4 ± 10.7	-0.8 ± 0.1	48.3	-19.8 ± 2.9	-61.4 ± 9.5	-0.8 ± 0.2	49.3
CUAUGG <u>AAAA</u> CAUAGF	-22.1 ± 6.0	-69.8 ± 19.1	-0.5 ± 0.3	43.9	-21.2 ± 3.5	-68.1 ± 11.2	-0.0 ± 0.1	37.5
FCAAGUG <u>AAAA</u> AUCUG	-22.5 ± 7.0	-67.6 ± 21.9	-1.5 ± 0.3	59.2	-34.9 ± 2.6	-105.7 ± 7.8	-2.1 ± 0.2	56.9
FCUAAGUG <u>AAAA</u> CUAG	-21.6 ± 3.7	-68.3 ± 12.4	-0.4 ± 0.1	43.1	-32.3 ± 12.5	-102.1 ± 39.8	-0.7 ± 0.3	43.1
FCAUAAGG <u>AAAA</u> CUAUG	-12.0 ± 4.1	-38.0 ± 12.6	-0.2 ± 0.5	42.8	-23.8 ± 2.9	-76.5 ± 8.9	-0.1 ± 0.2	42.8

^aSolutions are 1.0 M NaCl, 10 mM sodium cacodylate, 0.5 mM EDTA (pH 7.0).^bNucleotides in bold are the potential bulge residues. Hairpin loops are underlined.
(F) Fluorescein label.

base pair (Freier et al. 1983). Similar influences have been observed when 2-aminopurine is the dangling base. The enhanced stabilization of the 3' dangling base was shown to result from the greater proportion of time that the 3' dangling base is stacked on the terminal base pair (Liu et al. 2008). The 3' fluorescein labeled hairpins were, on average, only 0.2 kcal/mol more stable than the unlabeled hairpins, and the 5'-fluorescein-labeled hairpins did not generate any additional stabilization. It was somewhat surprising that

the 3'-terminal fluorescein did not significantly stabilize the hairpins examined in Table 1. This may be due to the nature of the xanthene ring of the fluorescein label or the linkage (3-aminopropyl)-O-propyl used to couple the fluorescein to the RNA.

Influence of the GAAA tetraloop on hairpin stability

All of the hairpin loops in this investigation were designed with a stable GNRA tetraloop, specifically, GAAA. The free-

TABLE 2. Free-energy increment for secondary structural motifs^a

oligomer ^b	+fluorescein tag	- fluorescein tag
	ΔG°_{37} (kcal/mol)	ΔG°_{37} (kcal/mol)
Hairpin Loop		
CGAUG <u>AAAA</u> AUCG		3.7
CAGUG <u>AAAA</u> ACUG		4.3
CUAGG <u>AAAC</u> UAG		3.9
CUAUGG <u>AAAC</u> CAUAG		3.9
GUACG <u>AAAG</u> UAC		3.3
CUGUG <u>AAAC</u> ACAG		4.2
GGUCG <u>AAAG</u> ACC		3.5
CAUGG <u>AAAC</u> CAUG		3.2
CUGAUG <u>AAAA</u> AUCAGF	4.6	4.6
CAUGUG <u>AAAA</u> ACAUGF	4.0	4.6
CAUUGG <u>AAAC</u> CAUUGF	3.3	3.8
FCAGAUG <u>AAAA</u> AUCUG	4.6	4.6
FCUAGUG <u>AAAC</u> ACUAG	4.8	4.6
FCAUAGG <u>AAAC</u> CAUAG	4.2	3.7
Group I bulge loop		
CGAUG <u>AAAA</u> AUCAGF	0.5	
GGUCG <u>AAAG</u> ACACF	0.5	
CAGUG <u>AAAA</u> ACAUGF	1.6	1.3
CUAGG <u>AAAC</u> CAUAGF	1.2	0.9
FCAGAUG <u>AAAA</u> AUCG	0.7	0.6
FCUAGUG <u>AAAC</u> ACAG	1.1	1.7
FCAUAGG <u>AAAC</u> CAUG	2.0	
Group II bulge loop		
CUGAUG <u>AAAA</u> AUCAAGF	0.8	0.9
GUACG <u>AAAG</u> UAACF	1.1	
CAUGUG <u>AAAA</u> ACAUGF	2.3	1.7
CUAUGG <u>AAAC</u> CAUAGF	2.2	2.7
FCAAGAUG <u>AAAC</u> AUCUG	1.1	0.5
FCUAAGUG <u>AAAC</u> ACUAG	2.1	2.0
FCAUAAGG <u>AAAC</u> CAUAG	2.2	2.8

^aSolutions are 1.0 M NaCl, 10 mM sodium cacodylate, 0.5 mM EDTA (pH 7.0).
^bNucleotides in bold are the potential bulge residues. Hairpin loops are underlined.
(F) Fluorescein label.

energy increment of the loops varied between 3.7 and 4.6 kcal/mol. A simple nearest neighbor model to predict the stability of the RNA hairpin tetraloop has recently been

described (Sheehy et al. 2010). The model predicts that the GAAA tetraloop would have a free-energy increment of 3.7 kcal/mol, which is slightly less than the average hairpin loop value measured here, 4.1 kcal/mol. The hairpins in Table 1 have two different hairpin closing base pairs, either UA or GC. The hairpin loops closed by UA, on average, had a higher free-energy increment than the hairpin loops closed by GC, 4.3 versus 3.8 kcal/mol, respectively. This difference may represent a non-nearest-neighbor interaction (Sheehy et al. 2010) or may be related to the stacking of the first mismatch on the closing base pair, as exemplified by the stacking of a mismatch on a terminal base pair of a duplex (Vecenie et al. 2006). The difference in stability between a $\begin{pmatrix} 5'UG \\ 3'AA \end{pmatrix}$ and a $\begin{pmatrix} 5'GG \\ 3'CA \end{pmatrix}$ terminal mismatch is 0.4 kcal/mol, almost exactly the difference observed in the free-energy increment for the two hairpin loops (Hickey and Turner 1985; Sugimoto et al. 1987; Sheehy et al. 2010). Further investigation will be necessary to determine the origins of the stability differences of the hairpin loops.

Influence of group I single-nucleotide bulge loops embedded in a hairpin stem

In the development of the current model to predict the influence of bulge loops on stability of RNA duplexes (Bloese et al. 2007; McCann et al. 2011), we investigated insertion of the bulge into different locations in duplexes of different lengths. In no instance was the bulge inserted next to a terminal base pair. This motif was omitted because it was assumed that a bulge located next to a terminal base pair would prevent the terminal base pair from forming. This would influence the thermodynamics of duplex formation by interfering with the formation of the terminal base pair in addition to the influence of inserting the bulge. It was, therefore, expected that insertion of the bulge at the 3'- and 5'-end positions would have a large impact on the stability of the hairpin. Surprisingly, the insertion of the bulge at the end positions showed a very small effect, on average decreasing the stability of the hairpin by 0.6 ± 0.1 kcal/mol. We measured the thermodynamics of duplex formation for two duplexes with terminal group I single-nucleotide bulge loops; the results are presented in Table 5. The influence of the terminal bulge was only 0.7 ± 0.1 kcal/mol per bulge loop. It is possible that the bulged nucleotide could disrupt the terminal base pair and form instead a mismatch with a single-nucleotide overhang. For the first duplex in Table 5, the $\begin{pmatrix} 5'C()G \\ 3'GAC \end{pmatrix}$ end could form $\begin{pmatrix} 5'CG \\ 3'GAC \end{pmatrix}$ instead. The nearest neighbor free-energy value of $\begin{pmatrix} 5'CG \\ 3'GC \end{pmatrix}$ is -2.36 kcal/mol, and the nearest neighbor value of the $\begin{pmatrix} 5'CG \\ 3'GAC \end{pmatrix}$ terminal mismatch is -1.4 kcal/mol.

TABLE 3. Thermodynamics for group I single nucleotide bulges in a hairpin stem

5' CA GU ^G _A ^a GUACA ^A								
			Stem only		Stem and first mismatch		Stem and hairpin loop	
	Potential end stem	ΔG_{37}° ^b kcal/mol	Potential hairpin stem	ΔG_{37}° kcal/mol	Potential hairpin stem	ΔG_{37}° kcal/mol	Potential hairpin stem	ΔG_{37}° kcal/mol
	CA GU	-2.1	GU CA	-2.2	GU ^G CA _A	-3.3	GU ^G _A CA ^A	+0.8
ΔG_{37}° bulge ^c								
	M	P		P		P		P
end	0.6 ± 0.1	0.7		2.5		3.1		0.5
middle	1.5 ± 0.3	1.8		1.9		2.5		0.3
loop	2.1 ± 0.2	2.5		0.7		1.3		-0.9

^aExample hairpin.^bValues predicted using nearest neighbor model for duplex stability (Xia et al. 2001) and hairpin loop.^cValues predicted using Equation 1 for insertion of a bulge loop into the hairpin stem. (M) Measured (average ± standard deviation); (P) predicted as described in the text.

The difference 0.9 kcal/mol is almost exactly the change in free energy for the insertion of the bulge into the duplex. A similar analysis for the hairpins in Table 1, however, results in an average difference between the terminal base pair and the terminal mismatch of 1.2 ± 0.2 kcal/mol, significantly different from the 0.6 ± 0.1 kcal/mol that we observe for the insertion of the bulge next to the terminal base pair. It is also possible that a terminal bulge may be in equilibrium between the terminal base pair and the terminal mismatch structures. If the equilibrium constant is 1 so that the two species have equal concentrations, then this would make the free energy of the bulge appear more favorable by $-RT \ln 2$ or 0.4 kcal/mol at 37°C. While we cannot rule out formation of the terminal mismatch, the small effect of the bulge adjacent to a terminal base pair for the hairpins suggests that the flexibility of the end of a helix, either in a duplex or hairpin stem, allows for effective base-stacking and H-bond formation in spite of the presence of a bulge loop. This effect is also manifest in the enthalpy change associated with the insertion of a bulge loop adjacent to a terminal base pair. The average enthalpy change for

the insertion of the bulge at the end positions for the hairpins in Table 1 is 6.7 kcal/mol, which is significantly smaller than the enthalpy change observed for inserting a bulge into an interior position of a duplex, 14.7 kcal/mol (McCann et al. 2011). Similarly, the average enthalpy change due to the terminal group I single-nucleotide bulge loops in the duplexes presented in Table 5 is only 4.8 ± 1.2 kcal/mol per bulge loop.

The average free-energy increment for inserting a bulge into the middle position of the hairpin stem is 1.5 kcal/mol and for inserting a bulge at the loop position is 2.1 kcal/mol. The model developed to predict the free-energy increment of bulge loops on duplex formation (Eq. 1) indicates that the stability of the less stable stem influences the extent of the destabilization. For duplexes, the stability of the two stems can be estimated using the nearest neighbor parameters for duplex formation (Xia et al. 1998). For hairpins, it is not obvious how to determine the thermodynamic stability of the stems adjacent to the bulged nucleotide because it is not clear how to

TABLE 4. Thermodynamics for group II single nucleotide bulges in a hairpin stem

5' CA U GU ^G _A ^a GUAACA ^A								
			Stem only		Stem and first mismatch		Stem and hairpin loop	
Bulged nucleotide	Potential end stem	ΔG_{37}° ^b kcal/mol	Potential hairpin stem	ΔG_{37}° kcal/mol	Potential hairpin stem	ΔG_{37}° kcal/mol	Potential hairpin stem	ΔG_{37}° kcal/mol
5' Adenosine	CAU GUA	-3.2	GU CA	-2.2	GU ^G CA _A	-3.3	GU ^G _A CA ^A	+0.8
3' Adenosine	CA GU	-2.1	UGU ACA	-4.3	UGU ^G ACA _A	-5.4	UGU ^G _A ACA ^A	-1.3
ΔG_{37}° bulge ^c								
	M	P		P		P		P
end	0.8 ± 0.2	0.7		3.6		4.2		2.0
middle	2.0 ± 0.2	1.8		3.0		3.9		1.7
loop	2.4 ± 0.2	2.5		1.8		2.7		0.5

^aExample hairpin.^bValues predicted using nearest neighbor model for duplex stability (Xia et al. 2001) and hairpin loop.^cValues predicted using Equation 1 for insertion of a bulge loop into the hairpin stem. Note, only the bottom row was used for prediction (see Discussion).

(M) Measured (average ± standard deviation); (P) predicted as described in the text.

TABLE 5. Thermodynamic parameters for duplex formation^a

oligomers ^b	T_M^{-1} vs $\log C_T$ plots				Average of curve fits			
	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	$-\Delta G_{37}^\circ$ (kcal/mol)	T_M^c (°C)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	$-\Delta G_{37}^\circ$ (kcal/mol)	T_M^c (°C)
C A GUGCAC G G CACGUG A C	64.8	174.0	10.8	63.6	65.6	176.2	10.9	64.0
C A GGUACC G G CCAUGG A C	67.0	179.8	11.3	65.2	67.8	181.8	11.4	65.4
CGUGCACG ^d GCACGUGC	76.2	206.0	12.3					
CGGUACCG ^d GCCAUGGC	74.9	201.2	12.5					

^aSolutions are 1.0 M NaCl, 10 mM sodium cacodylate, 0.5 mM EDTA (pH 7).^bNucleotide in bold is the bulge residue. Top sequence is written 5' → 3'.^cCalculated at 10^{-4} M oligomer concentration.^dPredicted using nearest neighbor values (Xia et al. 1998).

determine the effect of the hairpin loop on the stability of the adjacent stem. For example, the stem between the bulge and the hairpin loop could be considered as just the base-paired nucleotides or could include the stabilization due to the first mismatch, or also could include the destabilization due to the hairpin loop. Table 3 lists the possible stems and thermodynamic stabilities for the hairpin shown in the table. Using these possible stem stabilities and Equation 1, the influence of the bulge can be predicted based on the expected possible stem combinations. Similar calculations were done for each of the group I hairpins in Table 1, and the predicted values for the insertion of a bulge adjacent to stems of these stabilities were determined. The predicted values are presented in Table 3. The best predictor of the free energy increment for the insertion of a group I bulge loop into a hairpin stem is to use the stability of the “distal stem,” the stem without the hairpin loop. The measured average values and predicted values are 0.6 and 0.7, 1.5 and 1.8, and 2.1 and 2.5 kcal/mol, respectively, for the end, middle, and loop positions of the bulge. None of the other potential hairpin stems gives good agreement with the measured values. Therefore, for the hairpins in Table 1, the model developed for the influence of bulge loop insertion into a duplex motif can be used to predict the influence of the insertion of bulge loops into the stem of a hairpin.

The bulge effect on the enthalpy of hairpin formation at the middle and end positions is 15.0 and 17.6 kcal/mol, respectively. Neither of these values is significantly different from the effect of bulge loop insertion on the enthalpy of duplex formation, 14.7 kcal/mol (McCann et al. 2011). The enthalpy value can be used in conjunction with the free energy and enthalpy parameters for other nearest neighbor motifs to determine the stability of RNA structures at temperatures other than 37°C (Lu et al. 2006).

Uhlenbeck (Groebe and Uhlenbeck 1989) previously measured the thermodynamic parameters for hairpin formation of two group I bulge loops in the context of a hairpin stem and the parent hairpin. The measured free-energy increment of the insertion of the bulge was 3.4 or 3.5 kcal/mol for an A or G bulge, respectively. Using the model developed above, the predicted value would be 6.4 kcal/mol for both bulge loops. The inability to predict the influence on stability of these bulge loops may be due to the longer stem lengths (eight) versus the 4-bp stems used in this investigation, or the influence of the unpaired terminal bases on both the 3' and 5' sides of their hairpins. It may also be due to the difficulty in accurately measuring the thermodynamics of the parent hairpin, which in their study had a melting temperature of 94.3°C. Current prediction models (Vecenie et al. 2006) predict a melting temperature for the parent hairpin of 92.8°C but predict that the free energy for hairpin formation is -13.4 kcal/mol at 37°C versus the measured free energy of -11.1 kcal/mol. Using the predicted free energy for the parent hairpin leads to a value of 5.7 or 5.8 kcal/mol for the insertion of the bulge loop into the hairpin stem, close to the predicted value of 6.5 kcal/mol. Hairpin loops used in this investigation were all of the GNRA tetraloop class; therefore, further investigations will be necessary to determine the extent to which the model developed above for predicting bulge loop stability in the context of a hairpin loop will be applicable.

Influence of group II single-nucleotide bulge loops embedded in a hairpin stem

A similar analysis was completed for the group II hairpins in Table 1. Again, the influence of the bulge was position-dependent but did not depend on whether the bulge was to

the 3' or 5' side of the hairpin loop. The average free-energy increment for the insertion of a group II single-nucleotide bulge loop at the end position was 0.8 kcal/mol, and the average enthalpy change was 6.7 kcal/mol. Because of the positional ambiguity of the bulge, it is not obvious how to determine the thermodynamic stability of the stems adjacent to the bulged nucleotide. Table 4 lists the possible stems and thermodynamic stabilities for a representative hairpin. Using these possible stem stabilities and Equation 1, it is possible to predict the influence of the bulge based on each of the expected possible stem combinations. In our structural studies of group II bulge loops, we observed, for the middle and loop positions, that the bulged nucleotide further from the hairpin loop was, in fact, the bulged nucleotide. Therefore, we focused our analysis on the lower row of possible stem stabilities in Table 4, which would result if the nucleotide further from the hairpin loop was, in fact, the bulged nucleotide. Using these possible stem stabilities and Equation 1, the influence of the bulge can be predicted based on the expected possible stem combinations. Similar calculations were done for each of the group II hairpins in Table 1, and the predicted values for the insertion of a bulge adjacent to stems of these stabilities were determined. The predicted values are presented in Table 4. As with the group I hairpins, the best predictor of the free-energy increment for the insertion of a group II bulge loop into a hairpin stem is to use the stability of the "distal stem," the stem without the hairpin loop. The measured average values and predicted values are 0.8 and 0.7, 2.0 and 1.8, and 2.4 and 2.5 kcal/mol, respectively, for the end, middle, and loop bulges. None of the other potential hairpin stems gives good agreement with the measured values. Therefore, the model developed for the influence of bulge loop insertion into a duplex motif can be used also to predict the influence of the insertion of group II single-nucleotide bulge loops into the stem of a hairpin, if the ΔG° (stem) is calculated for the distal stem.

A group II bulge loop is found in a hairpin stem of the SAM-1 riboswitch. The bulge is an A, and the hairpin loop is a GAAA tetraloop, identical to the loop used in this study. The bulge nucleotide is in the binding site for SAM and undergoes a conformational change upon SAM binding. In the free riboswitch, the two adenosine residues (45 and 46) form a base triple with uridine 57. While in the SAM bound form, A45 and U57 form a base pair, and A46 forms a base triple with C47 and G56. In neither case is the bulge nucleotide unpaired, but rather forms two distinct base triples. One important consequence of a bulge loop in addition to its influence on the thermodynamics of structure formation is the structural flexibility it imparts on an RNA helix. Bulge loops previously have been found to assume three distinct structural orientations, intercalated, extrahelical, and side-by-side (base triple) (Znosko et al. 2002).

Conclusions

Since the free-energy increment for the insertion of a bulge loop into either a duplex or bulge stem is dominated by non-nearest-neighbor influences, the model developed to predict bulge loop stability cannot be directly incorporated into current methods for RNA secondary structure determination (e.g., Mfold). An iterative approach may be necessary, whereby an average value for bulge loop insertion is used initially to develop a set of RNA structures whose free energy is then refined using the model determined above. This may also be necessary for some hairpin structures that have been shown also to exhibit non-nearest-neighbor influences (Sheehy et al. 2010).

MATERIALS AND METHODS

RNA synthesis and purification

Most oligomers were synthesized on CPG solid supports (Applied Biosystems 392 DNA/RNA Synthesizer) using phosphoramidites with the 2'-hydroxyl protected as the *tert*-butyl dimethylsilyl ether from Glen Research (Usman et al. 1987; Wincott et al. 1995) or purchased from IDT. Fluorescein labels were added during synthesis. The 3' end was labeled using 3'-(6-FAM) CPG (Glen Research), and the 5' end was labeled using 5' fluorescein phosphoramidite (Metkinnen Chemistry). Oligomers underwent ammonia and fluoride deprotection, and the crude sample was purified using preparative *tlc* (*n*-propanol:ammonium hydroxide:water, 55:35:10) and Sep-Pak C18 (Waters) chromatography. Sample purity was determined through analytical *tlc* or HPLC (C-18) and was >95%.

Melting curve and data analysis

Optical melting experiments were performed using a Perkin-Elmer Lambda 25 equipped with a PTP-6 temperature controller. Absorbance changes for oligomers in 1 M NaCl melt buffer (1 M NaCl, 10 mM cacodylic acid, 0.5 mM EDTA at pH 7.0) were recorded as a function of temperature from 95°C to 10°C at a rate of 1°C/min, as described previously (Serra et al. 1994). The experiment was repeated at 10 varying sample concentrations to give at least a 50-fold concentration range (10 μ M–1 mM) for each sample. Absorbance-versus-temperature profiles were fit to a two-state model with sloping base lines using a nonlinear least squares program (McDowell and Turner 1996). Thermodynamic parameters for duplex formation (enthalpy and entropy) were obtained from the fits of the individual melting curves. Each hairpin melting curve displayed a single transition with concentration-independent melting temperatures, indicative of unimolecular hairpin transitions.

The Gibbs free energy change at 37°C was calculated as:

$$\Delta G_{37}^\circ = \Delta H^\circ - (310.15) \Delta S^\circ.$$

Statistical analysis

Statistical analysis of the data was done using the statistical software available with GraphPad Prism and GraphPad Instat.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation Grant # MCB-0744631 and the Paul E. and Mildred L. Hill Fund of Allegheny College.

Received May 23, 2011; accepted December 11, 2011.

REFERENCES

- Andronescu M, Condon A, Hoos HH, Mathews DH, Murphy KP. 2007. Efficient parameter estimation for RNA secondary structure prediction. *Bioinformatics* **23**: i19–i28.
- Blose JM, Manni ML, Klapac KA, Stranger-Jones Y, Zyra AC, Sim V, Griffith CA, Long JD, Serra MJ. 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.
- Cevic M, Thibaudeau C, Plavec J. 2010. NMR structure of the *let-7* miRNA interacting with the site LCS1 of *lin-41* mRNA from *Caenorhabditis elegans*. *Nucleic Acids Res* **38**: 7814–7821.
- Climie SC, Friesen JD. 1987. Feedback regulation of the *rplJL-rpoBC* ribosomal protein operon of *Escherichia coli* requires a region of mRNA secondary structure. *J Mol Biol* **198**: 371–381.
- Freier SM, Burger BJ, Alkema D, Neilson T, Turner DH. 1983. Effects of 3' dangling end stacking on the stability of GGCC and CCGG double helices. *Biochemistry* **22**: 6198–6206.
- Gott JM, Wilhelm LJ, Uhlenbeck OC. 1991. RNA binding properties of the coat protein from bacteriophage GA. *Nucleic Acids Res* **19**: 6499–6503.
- Groebe DR, Uhlenbeck OC. 1989. Thermal stability of RNA hairpins containing a four-membered loop and a bulge nucleotide. *Biochemistry* **28**: 742–747.
- Gutell RR. 1993. Collection of small subunit (16S- and 16S-like) ribosomal RNA structures. *Nucleic Acids Res* **21**: 3051–3054.
- Gutell RR, Gray MW, Schnare MN. 1993. A compilation of large subunit (23S and 23S-like) ribosomal RNA structures: 1993. *Nucleic Acids Res* **21**: 3055–3074.
- Harper JW, Logsdon NJ. 1991. Refolded HIV-1 tat protein protects both bulge and loop nucleotides in TAR RNA from ribonucleolytic cleavage. *Biochemistry* **30**: 8060–8066.
- Hickey DR, Turner DH. 1985. Effects of terminal mismatches on RNA stability: Thermodynamics of duplex formation for ACCGGGp, ACCGGAp, and ACCGGCp. *Biochemistry* **24**: 3987–3991.
- Lilley DM. 1995. Kinking of DNA and RNA by base bulges. *Proc Natl Acad Sci* **92**: 7140–7152.
- Liu JD, Zhao L, Xia T. 2008. The dynamic structural basis of differential enhancement of conformational stability of 5'- and 3'-dangling ends in RNA. *Biochemistry* **47**: 5962–5975.
- Lu ZJ, Turner DH, Mathews DH. 2006. A set of nearest neighbor parameters for predicting the enthalpy change of RNA secondary structure formation. *Nucleic Acids Res* **34**: 4912–4924.
- Mathews DH, Sabina J, Zuker M, Turner DH. 1999. Expanded sequence dependence of thermodynamic parameters provides improved prediction of RNA secondary structure. *J Mol Biol* **288**: 911–940.
- Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH. 2004. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci* **101**: 7287–7292.
- McCann MD, Lim GF, Manni ML, Estes J, Klapac KA, Frattini GD, Knarr RJ, Gratton JL, Serra MJ. 2011. Non-nearest-neighbor dependence of the stability for RNA group II single-nucleotide bulge loops. *RNA* **17**: 108–119.
- McDowell JA, Turner DH. 1996. Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: Solution structure of (rGAGGUCUC)₂ by two-dimensional NMR and simulated annealing. *Biochemistry* **35**: 14077–14089.
- Parker R, Siliciano PG, Guthrie C. 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **49**: 229–239.
- Peattie DA, Douthwaite S, Garrett RA, Noller HF. 1981. A “bulged” double helix in a RNA–protein contact site. *Proc Natl Acad Sci* **78**: 7331–7335.
- Schmelzer C, Schweyen RJ. 1986. Self-splicing of group II introns in vitro: Mapping of the branch point and mutational inhibition of lariat formation. *Cell* **46**: 557–565.
- Serra MJ, Axenson TJ, Turner DH. 1994. A model for the stabilities of RNA hairpins based on a study of the sequence dependence of stability for hairpins of six nucleotides. *Biochemistry* **33**: 14289–14296.
- Sheehy JP, Davis AR, Znosko BM. 2010. Thermodynamic characterization of naturally occurring RNA tetraloops. *RNA* **16**: 417–429.
- Sugimoto N, Kierzek R, Turner DH. 1987. Sequence dependence for the energetics of terminal mismatches in ribooligonucleotides. *Biochemistry* **26**: 4559–4562.
- Turner DH, Sugimoto N, Freier SM. 1988. RNA structure prediction. *Annu Rev Biophys Chem* **17**: 167–192.
- Usman N, Ogilvie KK, Jiang MV, Cedergren R. 1987. The automated chemical synthesis of long oligoribonucleotides using 2'-O-silylated ribonucleoside 3'-O-phosphoramidites on a controlled-pore glass support: Synthesis of a 43-nucleotide sequence similar to the 3'-half molecule of an *Escherichia coli* formylmethionine tRNA. *J Am Chem Soc* **109**: 7845–7854.
- Vecenie CJ, Morrow CV, Zyra A, Serra MJ. 2006. A study of the sequence dependence of stability of RNA hairpin molecules with six nucleotide loops. *Biochemistry* **45**: 1400–1407.
- Wincott F, DiRenzo A, Shaffer C, Grimm S, Tracz D, Workman C, Sweedler D, Gonzalez C, Scaringe S, Usman N. 1995. Synthesis, deprotection, analysis and purification of RNA and ribozymes. *Nucleic Acids Res* **23**: 2677–2684.
- Woese CR, Gutell RR. 1989. Evidence for several higher order structural elements in ribosomal RNA. *Proc Natl Acad Sci* **86**: 3119–3122.
- Xia T, SantaLucia J Jr, Burkard ME, Kierzek R, Schroeder SJ, Jiao X, Cox C, Turner DH. 1998. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* **37**: 14719–14735.
- Xia T, Mathews DH, Turner DH. 2001. Thermodynamics of RNA secondary structure formation. In *RNA* (ed. DG Soll), pp. 21–48. Elsevier, New York.
- Znosko BM, Silvestri SB, Volkman H, Boswell B, Serra MJ. 2002. Thermodynamic properties for an expanded nearest-neighbor model for the formation of RNA duplexes with single nucleotide bulges. *Biochemistry* **41**: 10406–10417.



RNA

A PUBLICATION OF THE RNA SOCIETY

Stability of single-nucleotide bulge loops embedded in a GAAA RNA hairpin stem

Geoffrey F.S. Lim, Gregory E. Merz, Michael D. McCann, et al.

RNA 2012 18: 807-814 originally published online February 16, 2012

Access the most recent version at doi:[10.1261/rna.028308.111](https://doi.org/10.1261/rna.028308.111)

References

This article cites 31 articles, 6 of which can be accessed free at:
<http://rnajournal.cshlp.org/content/18/4/807.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *RNA* go to:
<http://rnajournal.cshlp.org/subscriptions>
