

Non-nearest-neighbor dependence of stability for group III RNA single nucleotide bulge loops

JESSICA L. KENT, MICHAEL D. MCCANN, DANIEL PHILLIPS, BRANDON L. PANARO, GEOFFREY F.S. LIM, and MARTIN J. SERRA¹

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

ABSTRACT

Thirty-five RNA duplexes containing single nucleotide bulge loops were optically melted and the thermodynamic parameters for each duplex determined. The bulge loops were of the group III variety, where the bulged nucleotide is either a AG/U or CU/G, leading to ambiguity to the exact position and identity of the bulge. All possible group III bulge loops with Watson–Crick nearest-neighbors were examined. The data were used to develop a model to predict the free energy of an RNA duplex containing a group III single nucleotide bulge loop. The destabilization of the duplex by the group III bulge could be modeled so that the bulge nucleotide leads to the formation of the Watson–Crick base pair rather than the wobble base pair. The destabilization of an RNA duplex caused by the insertion of a group III bulge is primarily dependent upon non-nearest-neighbor interactions and was shown to be dependent upon the stability of second least stable stem of the duplex. In-line structure probing of group III bulge loops embedded in a hairpin indicated that the bulged nucleotide is the one positioned further from the hairpin loop irrespective of whether the resulting stem formed a Watson–Crick or wobble base pair. Fourteen RNA hairpins containing group III bulge loops, either 3' or 5' of the hairpin loop, were optically melted and the thermodynamic parameters determined. The model developed to predict the influence of group III bulge loops on the stability of duplex formation was extended to predict the influence of bulge loops on hairpin stability.

Keywords: secondary structure prediction; structure probing; thermodynamics

INTRODUCTION

RNAs are intimately involved in a wide variety of biological activities. During the process of gene expression, they serve as informational molecules and part of the decoding machinery. RNAs can function as catalysts, either alone or in conjunction with protein (Altman 1990; Cech 1990; Pace and Brown 1995; Scott 2007). Ribosomal RNA has a direct involvement in protein biosynthesis (Noller et al. 1992; Nissen et al. 2000), and new classes of noncoding RNAs continue to be discovered. The functional diversity of RNA is often predicated on the folding of complex tertiary structures, giving rise to the native, functional fold (Brion and Westhof 1997; Tinoco and Bustamante 1999). Secondary structure, defined as the formation of complementary base pairs, generally forms on a faster timescale (Zarrinkar and Williamson 1994; Woodson 2000) and tends to be more stable than tertiary contacts (Crothers et al. 1974; Banerjee et al. 1993; Onoa et al. 2003). Since tertiary structure arises from preformed secondary structure, accurately determining the secondary structure of an RNA molecule would provide clues to its function

as well as facilitate prediction of the RNA tertiary fold. In addition, links between secondary structure and three-dimensional folding are beginning to be explored (Bailor et al. 2010).

For these reasons, methods are being developed to model RNA secondary structure (Zuker 1989; Michel and Westhof 1990; Gautheret et al. 1993; Mathews and Turner 2002; Tahirov et al. 2002), and there is increased interest in predicting both the secondary and tertiary structure of RNA from sequence. The most commonly used programs for secondary structure prediction, Mfold and RNAstructure (Mathews et al. 1999), use thermodynamic parameters to predict the most stable (and suboptimal) secondary structural folds. The ability of these programs to predict secondary structure depends upon the availability of accurate models to provide the thermodynamic parameters for the various RNA secondary structural motifs. Current structure prediction algorithms using updated thermodynamic parameters correctly predict ~73% of known base pairs (Mathews et al. 1999, 2004). Therefore, it follows that refinement of the input

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.043232.113>.

© 2014 Kent et al. This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

thermodynamic parameters for multiple secondary structure motifs could improve predictions of secondary structures.

One secondary structure motif is the single nucleotide bulge. A single nucleotide bulge occurs when an unpaired nucleotide disrupts regions of contiguous base-pairing. These unpaired nucleotides can participate in a variety of biological events, including protein binding and tertiary structure formation (Harper and Logsdon 1991; Rounseville and Kumar 1992; Rounseville et al. 1996; Klasens et al. 1999). More recently, single nucleotide bulges have been implicated in reverse transcriptase-mediated RNA displacement synthesis as single nucleotide bulges enhance synthesis through stable secondary structures (Lancialt and Champoux 2005). Only a few studies have examined the influence of bulge loops on the thermodynamics of secondary structure formation (Longfellow et al. 1990; Blose et al. 2007; McCann et al. 2011; Lim et al. 2012).

Previously, single nucleotide bulges were subdivided into groups according to the identity of the bulge and adjacent base pairs (Blose et al. 2007). Group I single nucleotide bulge loops were defined as bulge loops where the position of the bulge was unambiguous with a bulged nucleotide that is not identical to either of the neighboring nucleotides. Group II single nucleotide bulge loops were defined as bulge loops where the position of the bulge is ambiguous since the bulge nucleotide is identical to at least one of its neighboring nucleotides. A model was developed to predict the influence of group I and group II single nucleotide bulge loops on the thermodynamics of duplex formation (Blose et al. 2007; McCann et al. 2011). Group III bulge loops have additional sequence and structural complexities. For example, with the oligomer 5'-CGCUGCC/3'GCGCGG, the underlined G on the bottom strand can form a base pair with either the underlined C or U. Depending upon which pair forms, either the C or the U would be the bulged nucleotide and the parent duplex would contain either a Watson-Crick or wobble base pair having different thermodynamic stabilities (Table 1).

In the following study, the complete set of group III single nucleotide bulge loops with Watson-Crick nearest-neighbors has been thermodynamically characterized to improve our ability to predict the stability of RNA duplexes with bulge loops. As observed for group I and II bulges, the free energy increment for the insertion of a group III bulge loop into a duplex was found to be primarily influenced by non-nearest-neighbor interactions (Blose et al. 2007; McCann et al.

2011). The stability of the stem adjacent to the bulge has a direct effect on duplex destabilization caused by the insertion of the bulge. This simple non-nearest-neighbor model was extended to predict the insertion of a group III bulge loop into a hairpin stem.

MATERIALS AND METHODS

RNA synthesis and purification

Most oligomers were synthesized on CPG solid supports (Applied Biosystems 392 DNA/RNA Synthesizer) utilizing phosphoramidites with the 2' hydroxyl protected as the tert-butyl dimethylsilyl ether from Glen Research (Usman et al. 1987; Wincott et al. 1995). 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. Some oligomers were ordered from Dharmacon or Integrated DNA Technologies, and deprotection of the oligomers was carried out using the manufacturers' instructions. The oligomers were then purified as described above. Sample purity was determined through analytical *tlc* or HPLC (C-8) and was >95%.

Melting curve and data analysis

For non-self-complementary sequences, individual strand concentrations were calculated from high-temperature single-strand absorbance at 260 and 280 nm using nearest-neighbor extinction coefficients (Borer 1975; Richards 1975). Single strands were then annealed in a 1:1 molar ratio. Optical melting experiments were performed using a Beckman DU 640 spectrophotometer and high-performance temperature controller at 260 or 280 nm. Absorbance changes for oligomers in 1 M NaCl melt buffer (1 M NaCl, 0.01–0.02 M cacodylic acid, 0.5 mM EDTA at pH 7.0) were recorded as a function of temperature from 10–90°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 μM–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 were obtained by two methods: (1) enthalpy and entropy changes from the fits of the individual melting curves were averaged, and (2) plots of the reciprocal melting temperature, T_M^{-1} , versus $\log(C_t/4)$ gave enthalpy and entropy changes (Borer et al. 1974):

$$T_M^{-1} = (2.3 R/\Delta H^0) \log(C_t/4) + (\Delta S^0/\Delta H^0). \quad (1)$$

Here, C_t is the total concentration of oligomer. Parameters derived from the two methods generally agreed within 15%, consistent with the two-state model (Freier et al. 1986; Allawi and SantaLucia 1997). The Gibbs free energy change at 37°C was calculated as

$$\Delta G_{37}^0 = \Delta H^0 - (310.15 \text{ K}) \Delta S^0. \quad (2)$$

Determination of the contribution of bulge loops to duplex thermodynamics

The free energy of duplex formation can be approximated by the nearest-neighbor model (Xia et al. 1998). The free energy contribu-

TABLE 1. Example of group III loop ambiguity

| Sequence | Potential bulge positions | |
|--|---|--|
| GU <u>C</u> GAUCAC CAGU AGUG | GU <u>C</u> GAUCAC CAGU AGUG | GU <u>C</u> GAUCAC CAG UAGUG |
| Parent duplex ($-\Delta G_{37}^0$ kcal/mol) | GU <u>C</u> GAUCAC CAGUAGUG (9.4) | GU <u>C</u> GAUCAC CAGUAGUG (10.4) |
| Stem to left of bulge ($-\Delta G_{37}^0$ kcal/mol) | GU <u>C</u> G CAGU (5.8) | GUC CAG (4.6) |
| Stem to right of bulge ($-\Delta G_{37}^0$ kcal/mol) | UCAC AGUG (6.2) | AUCAC UAGUG (7.4) |

tion of each bulged nucleotide was calculated from the experimental data and the nearest-neighbor model according to Equation 3, where $\Delta G_{37(\text{measured})}^0$ is the experimentally determined value from the melts and

$$\Delta G_{37,\text{bulge}}^0 = \Delta G_{37(\text{measured})}^0 - \Delta G_{37(\text{duplex})}^0 \quad (3)$$

$\Delta G_{37(\text{duplex})}^0$ is calculated from the nearest-neighbor model for the duplex forming the Watson–Crick base pair and not containing the bulge (Table 1). ΔH^0 values are calculated in a similar fashion.

Phylogenetic analysis

A database of phylogenetically determined RNA secondary structures (Cannone et al. 2002) of 305 SSU rRNAs, 169 LSU rRNAs, 16 group I introns RNAs, and seven group II intron RNAs, was searched for group III single nucleotide bulge loops. The loops were characterized by nearest-neighbor base pairs and bulge identity.

In-line probing

Samples for in-line probing were labeled at the 5' (Metkinnen) or 3' (Glen Research) end during synthesis with fluorescein or purchased from IDT. The RNAs were purified as described above except that nonfluorescent *tlc* plates were used and the RNA detected by the fluorescent label. The RNA was incubated in 50 mM Tris (pH 8.3), 20 mM MgCl₂, and 100 mM KCl for 2–4 d at room temperature in the dark (Soukup and Breaker 1999; Regulski and Breaker 2008). Control RNAs were incubated in probing buffer but incubated at 0°C. The cleavage products were analyzed by chromatography on nonfluorescent *tlc* plates (n-propanol:ammonium hydroxide:water, 65:25:10). The migration of the cleavage products were compared to the migration of mixtures of chemically synthesized oligomers of appropriate lengths. The fluorescent products were analyzed with a Kodak image station 400MM Pro.

Statistical analysis

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

RESULTS

The thermodynamics of only two duplexes with group III bulge loops have previously been measured (Znosko et al. 2002). Therefore, to determine the role of the bulge identity and the nearest-neighbors on the stability of group III bulge loops, 33 additional (eight GA, eight AG, 10 CU, and seven UC bulges) duplexes containing a group III single nucleotide bulge loop were prepared and the thermodynamics of duplex formation measured by optical melting. These oligomers, when combined with the previously measured oligomers, give measurements for the complete set of group III single nucleotide bulge loops with Watson–Crick nearest-neighbors.

Thermodynamic parameters for duplex formation by these oligonucleotides are listed in Table 2. The oligonucleotides are

listed by bulge sequence and then in order of decreasing free energy for the respective duplex. Residues in bold are the potential bulge nucleotide. The average deviations in thermodynamic parameter values are 11.2%, 12.5%, and 2.0% for ΔH^0 , ΔS^0 , and ΔG_{37}^0 , respectively. These average deviations are larger than observed for the group I and group II sequences and may indicate additional structural variability and, therefore, slight non-two-state behavior for some of these oligomers.

The ambiguity in the position of the group III bulge loops means that there are two potential parental duplexes (Table 1) with different thermodynamic stability. To determine the free energy contribution of the bulged nucleotide, one of the two potential parental duplexes must be selected. We first predicted the stability of all of the potential parental duplexes for each of the sequences in Table 2. The predicted stability of the duplexes, which included the Watson–Crick base pair, was always more stable than the predicted value for the duplexes with the wobble base pair. The ambiguity of the group III bulge loops can lead to a variety of analysis after trying many combinations; the best agreement between experiment and model (see Discussion) was achieved using the Watson–Crick duplexes as the parental duplex and calculating the free energy contribution of each bulged nucleotide according to Equation 3. This analysis is presented in Table 3. As previously observed (Znosko et al. 2002; Blose et al. 2007; McCann et al. 2011), all bulges destabilize the duplex. The extent of destabilization ranges between 0.5 and 4.9 kcal/mol. The group III pyrimidine single bulges are slightly less destabilizing than the group III purine bulges by, on average, 0.1 kcal/mol. The destabilization caused by the introduction of a single group I pyrimidine bulge into a duplex is 3.7 ± 1.0 kcal/mol and for a purine residue is 3.6 ± 1.0 kcal/mol. Since these values are not statistically different, they can be combined to provide a simple model to predict the stability of a duplex containing a single group III bulge nucleotide (purine or pyrimidine), which is just the average of all of the measured values: 3.6 kcal/mol.

In-line probing takes advantage of the natural instability of RNA to detect single-stranded regions in RNA. Structural flexibility of single-stranded regions allows the RNA to examine additional structural conformations, some of which produce an in-line arrangement of the 2' alkoxide and the 5'-O of the phosphodiester, leading to a favorable transition state and cleavage of the RNA strand at the site of flexibility. In-line probing was used to resolve the ambiguity of the group III bulge nucleotide. Because of the small size of our RNA oligomers, analytical *tlc* was used to resolve the products of in-line cleavage. Figure 1 displays the results of one such experiment. The two hairpins in this example have the group III bulge loop on the 3' side of the hairpin loop, and they differ by the orientation of the bulge loop, either 5'-GA or 5'-AG. In both cases, the control reactions show no cleavage products. We previously have shown that cleavage occurs to the 3' side of the bulged nucleotide (McCann et al. 2011); therefore, we would expect formation of either a dimer or trimer

depending upon which nucleotide (A or G) was bulged. In both cases, the predominant cleavage produces the dimer product, indicative of the fact that the nucleotide further from the hairpin loop is the bulged nucleotide irrespective of whether a Watson–Crick or wobble base pair forms. Table 4 summarizes the result of in-line cleavage analysis for seven pairs of hairpins with embedded group III bulge loops; in all cases, the major cleavage site occurs after the nucleotide further from the hairpin loop. These results are consistent whether the bulge loop is on the 5′ or 3′ side of the hairpin loop.

The thermodynamics for hairpin formation for the hairpins that were analyzed by in-line probing are presented in Table 5, along with the thermodynamics for the corresponding parental hairpins. As with group III bulge loops inserted into duplexes, the insertion of the bulge loop into a hairpin stem destabilizes hairpin formation. The influence of the stability of the hairpin was determined by subtraction of the stability of the hairpin with or without the bulge loop:

$$\Delta G^0_{37(\text{bulge})} = \Delta G^0_{37(\text{bulgehairpin})} - \Delta G^0_{37(\text{parenthairpin})}. \quad (4)$$

With group III bulge loops, there are two potential parental stems, one with the Watson–Crick base pair and the other with the wobble base pair. The bulge loop values are presented in Table 6. The influence of the group III bulges ranges from 1.3–3.0 kcal/mol at 37°C for the Watson–Crick parental hairpins and 0.3–2.7 kcal/mol at 37°C for the wobble parental hairpins. The free energy increment for inserting a group III bulge loop does not appear to depend upon whether the bulge is on the 5′ or 3′ side of the hairpin loop. Nor does it appear to depend upon whether the bulge sequence is 5′-AG or 5′-GA.

The enthalpy nearest-neighbor parameters for single nucleotide bulge loops are also presented in Table 3 for duplex and Table 5 for hairpin formation. In nearly all cases, the introduction of a bulge decreases the enthalpy of hairpin formation.

TABLE 2. Thermodynamic parameters for duplex formation^a

| oligomers ^b | T _M ⁻¹ vs log C _T plots | | | | average of curve fits | | | |
|--|--|--------------|----------------------------------|-------------------------------------|-----------------------|--------------|----------------------------------|-------------------------------------|
| | −ΔH° (kcal/mol) | −ΔS° (eu) | −ΔG° ₃₇ (kcal/mol) | T _M ^c (°C) | −ΔH° (kcal/mol) | −ΔS° (eu) | −ΔG° ₃₇ (kcal/mol) | T _M ^c (°C) |
| GA | | | | | | | | |
| GCUC GA CCUG CGAG UGGAC | 90.9 | 257.3 | 11.1 | 53.5 | 86.5 | 243.7 | 11.0 | 53.7 |
| CAGUCC GA GUC GUCAG UCAG | 58.3 | 160.4 | 8.6 | 48.2 | 60.7 | 167.9 | 8.6 | 48.0 |
| CAUGU GA CUAC ^d GUACA UGAUG | 82.8 | 238.3 | 8.9 | 46.0 | 82.6 | 237.8 | 8.9 | 46.0 |
| CAGUCC GAGC GUCAG UCG | 67.2 | 192.3 | 7.6 | 42.1 | 62.3 | 176.3 | 7.7 | 42.7 |
| GCUC GACGAC CGA UGCUG | 55.6 | 155.4 | 7.4 | 41.8 | 61.4 | 174.2 | 7.3 | 41.1 |
| CAU GAGAGC GUA UCUCG | 47.1 | 128.9 | 7.2 | 41.2 | 49.3 | 136.0 | 7.2 | 41.1 |
| GAU GAGAGC CUA UCUCG | 64.8 | 185.3 | 7.3 | 40.9 | 64.9 | 185.5 | 7.4 | 41.0 |
| GUC GAUCAC CAG UAGUG | 57.2 | 163.9 | 6.4 | 36.3 | 43.0 | 117.2 | 6.7 | 38.0 |
| GUC GAUGAC CA UACUG | 40.1 | 111.5 | 5.5 | 29.3 | 40.9 | 113.7 | 5.6 | 30.4 |
| AG | | | | | | | | |
| GAC AGCCAG CUGU GGUC | 64.5 | 182.1 | 8.0 | 44.4 | 68.2 | 193.6 | 8.2 | 44.6 |
| CAG AGCAGC GUCU GUCG | 50.2 | 137.0 | 7.7 | 44.2 | 52.4 | 143.7 | 7.8 | 44.6 |
| GUCGU AGC CAGCAU G | 44.5 | 120.1 | 7.2 | 42.0 | 47.6 | 129.9 | 7.3 | 42.0 |
| UG AGUGUA ACU ACAGU | 56.5 | 160.7 | 6.6 | 37.5 | 61.2 | 176.4 | 6.5 | 36.9 |
| GACAU AGUC CUGUAU AG | 39.0 | 106.0 | 6.2 | 34.0 | 32.8 | 84.8 | 6.5 | 36.8 |
| CAUG AGUGC GUACU ACG | 63.5 | 184.2 | 6.3 | 36.0 | 63.8 | 185.3 | 6.4 | 36.1 |
| CAUC AGUGC GUAGU ACG | 67.7 | 198.9 | 6.0 | 34.5 | 65.2 | 190.9 | 6.0 | 34.5 |
| CAAU AGUGC GUUAU ACG | 61.4 | 184.9 | 4.1 | 25.0 | 61.0 | 183.7 | 4.1 | 25.0 |
| CU | | | | | | | | |
| CGUGCG CUA GCACGCG U | 67.7 | 176.1 | 13.0 | 70.1 | 62.8 | 161.5 | 12.7 | 71.0 |
| GAG CUGGUC CUCG CCAG | 60.5 | 164.3 | 9.3 | 53.2 | 68.3 | 188.32 | 9.9 | 52.9 |
| GCAC CUGAGG CGUG CUCC | 68.9 | 191.1 | 9.6 | 51.4 | 81.9 | 231.4 | 9.6 | 51.3 |
| GCAC CUGCAG ^d CGUG CGUC | 57.5 | 156.0 | 9.1 | 51.7 | 64.1 | 176.4 | 9.4 | 51.3 |
| GACG CUAGC CUGCG UCG | 93.9 | 269.3 | 10.4 | 50.4 | 88.2 | 251.6 | 10.2 | 50.3 |
| CGAC CUGCAG GCUG CGUC | 61.5 | 171.2 | 8.4 | 46.5 | 62.6 | 174.4 | 8.5 | 47.2 |
| CUG CUGUCG GACG CAGC | 68.3 | 191.9 | 8.8 | 47.6 | 80.3 | 229.5 | 9.1 | 47.2 |
| GGUA CUAAGC CCAUG UUCG | 78.0 | 223.3 | 8.8 | 46.3 | 70.1 | 198.1 | 8.6 | 46.6 |
| CAGA CUAGC GUCUG UCG | 62.3 | 175.2 | 8.0 | 44.5 | 66.0 | 186.9 | 8.0 | 44.2 |
| GCUC CUGUG CGAG CAC | 74.0 | 215.7 | 7.1 | 39.3 | 68.0 | 196.7 | 7.0 | 39.3 |
| GCUC CUAUG CGAG UAC | 53.8 | 151.7 | 6.8 | 35.7 | 64.1 | 185.3 | 6.7 | 36.2 |

TABLE 2. Continued

| | | | | | | | | |
|----------------------------------|------|-------|------|------|-------|-------|------|------|
| UC | | | | | | | | |
| GAG UC UAC CUC GAGUG | 58.5 | 158.1 | 9.5 | 53.5 | 66.7 | 183.0 | 9.9 | 53.5 |
| CAG UC GAGC GUC GCUCG | 82.1 | 231.4 | 10.3 | 52.1 | 95.3 | 272.4 | 10.8 | 51.5 |
| ACUGCA UC G UGACGU GC | 56.4 | 152.9 | 8.9 | 50.8 | 61.12 | 167.7 | 9.1 | 50.5 |
| CAGA UC GAGC GUCU GUCG | 67.6 | 192.3 | 8.0 | 43.6 | 70.0 | 199.8 | 8.0 | 43.7 |
| GU UC GUCG CAU GCAGC | 53.8 | 148.9 | 7.6 | 43.5 | 59.2 | 166.5 | 7.6 | 42.7 |
| GCA UC GUG CGU GACAC | 74.2 | 213.9 | 7.9 | 42.7 | 86.7 | 253.9 | 7.9 | 42.1 |
| GCG UC AG CGC GUC | 44.3 | 121.8 | 6.6 | 37.2 | 42.8 | 116.9 | 6.6 | 37.3 |

^aSolutions are 1.0 M NaCl, 10–20 mM sodium cacodylate, 0.5 mM EDTA (pH 7).

^bNucleotide in bold are potential bulge residues. Top sequence is written 5'→3' and bottom stand is written 3'→5'.

^cCalculated at 10⁻⁴ M oligomer concentration.

^dZnosko et al. (2002). The average deviations in thermodynamic parameters are 8.5%, 9.2%, and 2.2% for ΔH^0 , ΔS^0 , and ΔG^0_{37} , respectively.

tion between the influence of the bulge and the adjacent stem. The best correlation was observed when we graphed the free energy of the bulge determined using the Watson–Crick parental strand and the second least stable stem of the four potential stem (Table 1). Note that sometimes the second least stable stem included a wobble base pair. The results of this analysis are presented in Figure 2. The line for the group III data is not significantly different than the group I and group II data. The data from the group I, II, and III bulge loops can therefore be combined to generate a relationship to predict the thermodynamic influence of all bulge loops. That relationship is given in the equation below:

$$\Delta G^0_{37 \text{ bulge loop}} = -0.51 \Delta G^0_{37 \text{ less stable stem}} + 0.85, \quad (5)$$

The average enthalpy contribution for group III single nucleotide bulge loops is 15.2 kcal/mol for insertion of a bulge into a duplex and 13.6 kcal/mol for insertion of a bulge into a hairpin stem.

DISCUSSION

Non-nearest-neighbor influences on the thermodynamics for group III bulge loop on duplex formation

The main determinant of the thermodynamic influence of group I and II bulge loops on duplex formation was shown to be due to non-nearest-neighbor influences (Blose et al. 2007; McCann et al. 2011). Specifically, the stability of the stem affects the extent to which insertion of the bulge loop destabilized the duplex. To determine if group III bulge loops are also influenced by the stability of the stems adjacent to the bulge, we examined how the stability of the adjacent stem influences the thermodynamics of group III bulge loops insertion into a duplex. The conformational ambiguity of the group III bulge loops allows for two possible sets of duplexes (Table 1) and, therefore, four possible stems. In addition, the influence of the bulge depends upon which of the potential parental strands is selected (Table 1). Many combinations were tested, most of which gave a poor correla-

where $\Delta G^0_{37 \text{ less stable stem}}$ is the less stable stem for group I bulge loops and the second least stable stem for group II and group III bulge loops. While this represents a simple model to predict the influence of bulge loops on the stability of duplex formation, because of the non-nearest-neighbor

TABLE 3. Thermodynamic parameters and natural occurrence for group III bulges

| bulge sequence | $\Delta G^0_{37 \text{ bulge}}^a$ (kcal/mol) | $\Delta G^0_{37 \text{ bulge}}^a$ predicted (kcal/mol) | $\Delta H^0_{\text{bulge}}^a$ (kcal/mol) | % occurrence ^b | bulge sequence | $\Delta G^0_{37 \text{ bulge}}^a$ (kcal/mol) | $\Delta G^0_{37 \text{ bulge}}^a$ predicted (kcal/mol) | $\Delta H^0_{\text{bulge}}^a$ (kcal/mol) | % Occurrence ^b |
|----------------|---|--|---|------------------------------|----------------|---|--|---|------------------------------|
| CGAG G UG | 4.5 | 4.8 | -0.6 | 34.7 | ACUA UG U | 4.7 4.3 ^c | 4.1 3.9 | 4.9 14.7 ^c | 4.4 |
| CGAG G UC | 4.9 4.7 | 4.0 3.6 | 27.6 9.8 | 0.8 | ACUG UG C | 4.4 4.6 4.1 ^d | 4.4 4.8 4.8 | 15.8 22.1 11.2 ^d | 4.0 |
| CGAU G UA | 4.0 | 3.8 | 17.2 | 1.6 | GCUA CG U | 3.5 0.5 | 3.9 1.0 | -12.3 8.3 | 10.8 |
| UGAC A UG | 4.5 3.8 ^d | 4.5 4.5 | 19.7 3.9 ^d | 1.2 | GCUG CG C | 3.9 4.7 | 4.4 4.8 | 9.0 21.6 | 1.6 |
| UGAG A UC | 3.2 3.3 | 2.7 2.8 | 25.1 9.4 | 0.0 | UCUA AG U | 2.3 | 2.8 | 10.7 | 0.0 |
| UGAU A UA | 1.8 | 2.2 | 19.0 | 2.8 | UCUG AG C | 3.4 | 3.4 | -7.8 | 3.6 |
| CAGC GU G | 4.3 | 4.0 | 11.9 | 1.6 | AUCA U GU | 4.3 | 4.0 | 9.4 | 5.6 |
| CAGU GU A | 4.2 | 4.1 | 6.1 | 0.4 | AUCG U GC | 3.4 3.1 | 3.1 2.1 | 18.2 15.9 | 2.8 |
| GAGC CU G | 4.6 | 4.0 | 26.8 | 0.8 | AUCU U GA | 4.3 | 4.0 | 1.7 | 1.6 |
| GAGU CU A | 2.8 3.9 | 2.7 4.1 | 9.4 10.3 | 0.0 | GUCA C GU | 2.7 | 3.0 | 13.4 | 8.4 |
| UAGC AU G | 1.5 | 2.0 | 16.9 | 0.8 | GUCG C GC | 3.4 | 4.1 | -1.5 | 3.2 |
| UAGU AU A | 3.9 2.3 | 4.4 2.5 | 7.4 30.6 | 1.6 | GUCU C GA | 3.0 | 3.1 | 20.5 | 7.2 |

^aValues calculated as described in text.

^bPercentage of group III bulge loops with listed sequence.

^cMultiple values represent multiple measurements of the same bulge loop sequence in different duplex stems.

^dZnosko et al. (2002). Total number of single nucleotide group III bulge sequences is 249.

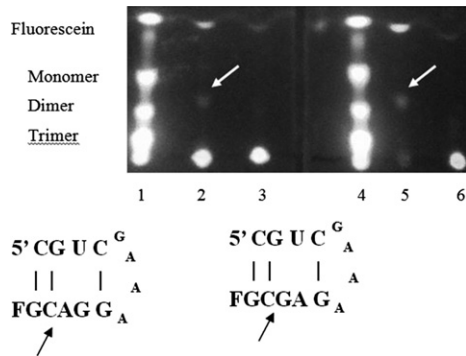


FIGURE 1. *tlc* analysis of in-line probing of group III bulge loops. (Lanes 1,4) Size markers: monomer, dimer, and trimer, each are labeled on the 3' end with fluorescein. (Lane 2) Cleavage of group III hairpin sequence given in *lower left* of figure. (Lane 3) Control lane, hairpin incubated at 0°C in probing buffer. (Lane 5) Cleavage of group III hairpin, sequence given in *lower right* of figure. (Lane 6) Control lane, hairpin incubated at 0°C in probing buffer. Hairpin figures indicate position of bulge and cleavage site.

influences, it cannot be incorporated directly into the current RNA secondary structure prediction programs.

Example calculations

Using the example duplex given in Table 1, the measured free energy for the duplex is 6.4 kcal/mol at 37°C (Table 2). The stability of the parent duplex forming the Watson–Crick base pair has a predicted stability using the nearest- neighbor model of 10.4 kcal/mol at 37°C. The free energy increment for the insertion of the bulge is then taken as the difference between the measured and predicted values or 4.0 kcal/mol. This represents the measured value presented in Table 3. The predicted value for the free energy increment for insertion of the group III bulge loops on duplex formation is given by Equation 5. For this example, the second least stable stem of the four potential stems is 5'-GUCG/3'-CAGU which has a predicted stability of −5.8 kcal/mol at 37°C. Using the −5.8 kcal/mol value in Equation 5 predicts that the free energy increment for insertion of the group III bulge should be 3.8 kcal/mol at 37°C, a difference of 0.2 kcal/mol from the measured value. Similar calculations were done for all of the group III bulge loops in Table 3. The average difference between the measured and predicted values for the group III bulge loops was 0.41 kcal/mol. Twenty-two of the 36 measured group III bulge loops had predicted values within 0.5 kcal/mol of the measured values, and only two had predicted values that were >1 kcal/mol. The two were 5'-CAGUCGAGC/3'-GUCAGUCG and 5'-ACUGCAUCG/3'-UGACGUGC; in both instances, the bulge loop is near the duplex terminus. In both cases, the predicted value is less than the measured free energy increment. It is possible that with these instances, the duplexes are forming structure other than the bulge loop; the entire end of the duplex may be frayed or the ends may form alternate non-Watson–Crick base pairs. There are

four other examples where the bulge is located near the terminus of the duplex and the free energy increments are predicted well; therefore, it is not obvious why the two examples above are predicted less well. Previous models to predict the free energy increment have used a constant value or 3.8 kcal/mol for the free energy increment for the insertion of a bulge loop into a duplex (Mathews et al. 1999). Using this model, the average difference between the measured and predicted values is 0.78 kcal/mol. Thus the model presented in Equation 5 improves the prediction of the free energy increment for insertion of a group III bulge loops into a duplex by, on average, 0.3 kcal/mol over the previous model.

Structural determination of the group III bulged nucleotide using in-line probing

The results of the in-line cleavage of the hairpins listed in Table 4 show that the major fluorescent product was the fragment expected if the bulged nucleotide was the one further from the hairpin loop, irrespective of whether the resulting hairpin forms a Watson–Crick or wobble base pair. These results are in agreement with those obtained for group II bulge loops, where it was observed that the bulge nucleotide was also the one further from the hairpin loop (McCann et al. 2011). Thus it appears that structural constraints are more

TABLE 4. Summary of in-line cleavage results

| | Percentage cut at 3' side guanosine | Percentage cut at 3' side adenosine |
|---------------------------------|-------------------------------------|-------------------------------------|
| 5' CGUCG ^a FGCAGG | | >95 |
| 5' CGUC FGCGAG | >95 | |
| 5' GAUAA FCUAGUU | | >95 |
| 5' GAUAA FCUGAUU | 75 | 25 |
| 5' CAUG FGUAGC | | >95 |
| 5' CAUG FGUGAC | >95 | |
| | | |
| 5' FCUAGUA GAU AU | | >95 |
| 5' FCUGAUA GAU AU | >95 | |
| 5' FGCAGG CGUC | | >95 |
| 5' FGCGAG CGUC | >95 | |
| 5' FGCAGGA CGUCU | 20 | 80 |
| 5' FGCGAGA CGUCU | >95 | |
| 5' FGUAGC CAUG | | >95 |
| 5' FGUGAC CAUG | 80 | 20 |

^aRepresents GAAA tetraloop.

TABLE 5. Thermodynamic parameters for hairpin formation in 1 M NaCl^a

| RNA Hairpin ^b | $-\Delta H^\circ$ kcal/mol | $-\Delta S^\circ$ (eu) | $-\Delta G^\circ_{37}$ (kcal/mol) | T_M (°C) |
|---------------------------|-------------------------------|---------------------------|--------------------------------------|---------------|
| parent | | | | |
| CGUCGAAAGAGCGF | 24.6±3.1 | 75.6±9.3 | 0.8±0.2 | 48.0 |
| CGUCGAAAGGACGF | 34.3±3.1 | 105.7±9.6 | 1.5±0.4 | 51.0 |
| CGUCGAAAGACG | 42.2±3.2 | 125.9±9.7 | 3.2±0.2 | 61.8 |
| CGUCGAAAGGCG | 40.7±3.6 | 121.5±11.8 | 3.0±0.4 | 62.1 |
| GAUAAGAAAUUAGUCF | 30.3±3.4 | 97.2±10.8 | 0.1±0.1 | 38.4 |
| GAUAAGAAAUUGAUCF | no observable transition | | | |
| GAUAAGAAAUUAUCF | 43.6±5.5 | 135.7±17.0 | 1.5±0.2 | 47.8 |
| GAUAAGAAAUUGUCF | 46.3±7.7 | 148.0±24.2 | 0.4±0.3 | 39.7 |
| CAUGGAAACAGUGF | 28.1±3.5 | 90.5±10.6 | 0.1±0.3 | 37.6 |
| CAUGGAAACGAUGF | 18.5±4.3 | 59.0±13.5 | 0.1±0.3 | 39.8 |
| CAUGGAAACAUG | 42.4±4.4 | 130.0±13.3 | 2.1±0.3 | 53.1 |
| CAUGGAAACGUG | 31.8±2.2 | 101.4±7.5 | 0.4±0.1 | 41.0 |
| FCUAGUAGAAAUUAG | 28.1±3.5 | 90.5±10.6 | 0.1±0.3 | 37.6 |
| FCUGAUAGAAAUUAG | 34.8±12.9 | 113.4±42.0 | -0.4±0.2 | 33.5 |
| FCUAUAGAAAUUAG | 45.2±19.5 | 140.6±63.2 | 2.0±0.4 | 52.6 |
| FCUGUAGAAAUUAG | non-two state | | | |
| FGCAGGGAACUGC | 19.7±3.5 | 64.9±10.6 | -0.5±0.3 | 29.9 |
| FGCGAGGGAACUGC | 20.4±3.2 | 64.1±9.8 | 0.6±0.4 | 45.9 |
| GCAGGGAACUGC | 30.7±4.6 | 91.0±14.1 | 2.5±0.4 | 64.2 |
| GCGGGAACUGC | 34.2±3.7 | 103.0±11.9 | 2.2±0.2 | 58.5 |
| FGCAGGAGAAUCUGC | 25.8±2.9 | 75.7±9.0 | 2.3±0.4 | 67.1 |
| FGCGAGAGAAUCUGC | 31.3±9.1 | 97.1±28.5 | 1.2±0.4 | 49.1 |
| FGCAGAGAAUCUGC | 39.6±8.4 | 115.8±25.4 | 3.7±0.6 | 68.6 |
| FGCGGAGAAUCUGC | 37.1±5.7 | 109.4±20.0 | 3.2±0.7 | 66.2 |
| FGUAGCGAAAGUAC | 24.6±5.5 | 78.6±17.4 | 0.2±0.3 | 39.4 |
| FGUGACGAAAGUAC | 22.7±3.7 | 71.5±11.8 | 0.5±0.2 | 44.1 |
| GUACGAAAGUAC ^c | 38.3±2.0 | 115.4±6.2 | 2.5±0.1 | 58.7 |
| GUGCGAAAGUAC | 30.8±2.2 | 93.9±8.0 | 1.6±0.3 | 54.5 |

^aMeasurements were made in 1.0 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na₂EDTA (pH 7). Errors in ΔH° , ΔS° , and ΔG° are standard deviations.

^bHairpin loop sequences are underlined, and potential bulges are shown in bold.

^cLim et al. (2012).

critical to determining the bulged nucleotide than the stability of the corresponding stem. No cleavage was detected due to fraying of the ends of the hairpin stem. This may be due to the fact that while a bulge confers conformational flexibility on a duplex, the preferred conformation is favorable for in-line cleavage (Portmann et al. 1996).

Thermodynamics of group III bulge loops embedded within a hairpin stem

Table 5 lists the thermodynamics for hairpin formation of the 14 hairpins examined by in-line probing. Also included in Table 5 are the thermodynamics of hairpin formation of the parent hairpins, including both the Watson–Crick and wobble base paired stems. Some of the hairpins in Table 5 have a fluorescein tag at either the 5' or 3' end. We have previously shown that the fluorescein tag does not affect the thermodynamics of hairpin formation (Lim et al. 2012), so we will not consider it further in our analysis. Again, there is a multitude of ways to analyze the influence of a group III bulge on the thermodynamics of hairpin formation due to the ambiguity of the bulged nucleotide. The influence of the bulge loop on the thermodynamics of hairpin formation

was determined using Equation 4 for both the Watson–Crick and wobble parental hairpins.

Table 6 displays the results of this analysis. The in-line probing results show that the bulged nucleotide was the one further from the hairpin loop, so we selected this nucleotide as the bulge, irrespective of whether or not this led to the formation of a wobble base pair. For group II bulge loops imbedded in a hairpin stem, the stability of the hairpin stem distal from the hairpin loop gave the best prediction of the influence of the bulge on the thermodynamics of hairpin formation (Lim et al. 2012). Therefore, we chose a similar model, using the stability of the distal stem and Equation 5, to predict the influence of the group III bulge loops on hairpin stability. The predicted values for the insertion of a group III bulge loop into a hairpin stem also are listed in Table 6. The best agreement between the measured and predicted values is obtained using the measured values for the parental hairpin with the Watson–Crick base pair rather than the wobble parental strand and selecting the parental hairpin based upon the in-line probing results (nucleotide further from the hairpin loop was selected as the bulged base).

For example, using the two hairpins shown in Figure 1, CGUCGAAAGCAGCG and CGUCGAAAGCGACG (where underlined letters represent the hairpin loop, and bold letters are the potential bulge nucleotides), the parental hairpin in both cases would be modeled as CGUCgaaGCACG (irrespective of the in-line probing results). The stability of this hairpin is -3.2 kcal/mol at 37°C (Table 5). The measured values for the stability of the two hairpins containing the group III bulge are -0.8 and -1.5 kcal/mol at 37°C, respectively (Table 5). The measured values for the insertion of the bulge loops would be 2.4 and 1.7 kcal/mol at 37°C for the first and second hairpin, respectively (Table 6). The distal stem for both hairpins would be 5'-CG/3'-GC; therefore, the predicted value for the insertion of the bulge loop into the hairpin stem would be 2.1 kcal/mol, using Equation 5 and the stability of 5'-CG/3'-GC nearest-neighbors (-2.36 kcal/mol) (Table 6).

Using this model, there is good agreement between the measured and predicted values with an average difference of 0.4 kcal/mol. Therefore, the model developed for the influence of bulge loop insertion into a duplex motif can be extended to predict the influence of the insertion of group III single nucleotide bulge loops into the stem of a hairpin, if the stability of the parental hairpin is chosen with the stem forming the Watson–Crick base pair and the ΔG° (stem) is calculated for the distal stem.

Enthalpic contributions of group III single nucleotide bulge loops on duplex and hairpin formation

Since the free energy increment for the insertion of a group III bulge loops was dependent upon the stability of the stem adjacent to the bulge, we investigated the influence of stem stability on the enthalpy of bulge insertion. As previously observed for the group I and II bulge loops, the least-squares fit

TABLE 6. Free energy increment for bulge loops^a

| oligomer ^b | Measured ^c Parent with A-U Base Pair | Measured ^c Parent with G-U Base Pair | Predicted ^d |
|--------------------------|---|---|---|
| | ΔG_{37}° bulge (kcal/mol) | ΔG_{37}° bulge (kcal/mol) | ΔG_{37}° bulge (kcal/mol) |
| 3' end | | | |
| CGUC <u>GAA</u> AGCGF | 2.4 | 2.2 | 2.1 |
| CGUC <u>GAA</u> GGACGF | 1.7 | 1.5 | 2.1 |
| GAUA <u>GAA</u> UUAGUCF | 1.4 | 0.3 | 1.8 |
| GAUA <u>GAA</u> UUGAUCF | | | 1.8 |
| CAUG <u>GAA</u> CAGUGF | 2.0 | 0.3 | 1.7 |
| CAUG <u>GAA</u> CGAUGF | 2.0 | 0.3 | 1.7 |
| 5' end | | | |
| FCUAGUAG <u>AA</u> UAUAG | 1.9 | | 1.7 |
| FCUGAUAG <u>AA</u> UAUAG | 2.4 | | 1.7 |
| FGCAGG <u>GAA</u> CUGC | 3.0 | 2.7 | 2.6 |
| FGCGAG <u>GAA</u> CUGC | 1.9 | 1.6 | 2.6 |
| FGCAGGAG <u>AA</u> UCUGC | 1.3 | 0.8 | 2.6 |
| FGCGAGAG <u>AA</u> UCUGC | 2.5 | 2.0 | 2.6 |
| FGUAGCG <u>AA</u> GUAC | 2.3 | 1.4 | 1.8 |
| FGUGACG <u>AA</u> GUAC | 2.0 | 1.1 | 1.8 |

^aMeasurements were made in 1.0 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na₂EDTA (pH 7).

^bHairpin loop sequences are underlined, and potential bulge nucleotides are shown in bold.

^cMeasured using Equation 4.

^dPredicted as described in text.

of data did not have a slope significantly different from zero (data not shown), suggesting that the average enthalpy value (15.2 kcal/mol for duplex insertion and 13.8 kcal/mol for hairpin) is a reasonable approximation for the enthalpic energy contribution for insertion of a group III bulge loop. The enthalpic contributions due to a group III bulge loops are not statistically different than the value obtained for group I or group II bulge loops (16.5 kcal/mol), so all of the enthalpic values for group I, II, and III bulge loops can be combined for an average value of 16.0 kcal/mol. The enthalpic values can be used in conjunction with the free energy parameters to determine the stability of RNA structures at temperatures other than 37°C (Lu et al. 2006).

Phylogenetic analysis of group III single nucleotide bulge loop

The database examined in this study contained 249 group III single nucleotide bulge loops. Group III bulge loops are less numerous than the group I (3520) or group II (768) bulge loops. The frequencies of occurrence of the group III single

nucleotide bulge loops are listed in Table 3. Over one-third of the group III bulge loops have the sequence (CG AC/GUG). This bulge includes the highly conserved A at position 397 (*Escherichia coli* numbering) of the small ribosomal RNA. Since the stability of group III single nucleotide bulge loops is independent of the identity of the bulge, there is no correlation between the thermodynamic contribution of the bulge and its frequency of occurrence. Therefore, the section of naturally occurring bulge nucleotides must be related to factors other than stability.

The distribution of bulge loops embedded within a hairpin stem was then examined. We selected for bulge loops embedded in the stems of hairpin loops from 3–8 nt in length and 1–6 bp from the hairpin loop. A total of 212 (or 88%) group III bulge loops match these criteria. Since >70% of the ribosomal RNAs are found in hairpin structures, it is not surprising that a high proportion of the group III bulge loops are present in hairpin stems. Of the 212 group III bulge loops, 163 of them are either AG or GA (data not shown). The bulge loops were then characterized as having the bulged base either nearer or further from the hairpin loop. The majority, 120 of the 212, had the nucleotide further from the hairpin loop as the bulged nucleotide.

Almost 90% of the phylogenetically determined group III bulge loops formed Watson–Crick base pairs; only 24 of the bulge loops formed wobble base pairs. These results are interesting given that the in-line probing results suggest a marked

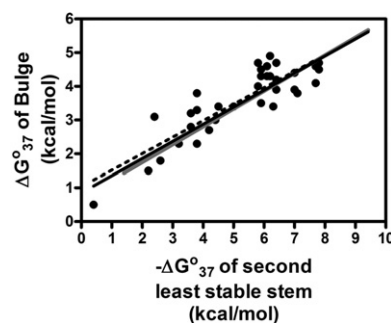


FIGURE 2. Plot of free energy change for group III bulge loop formation, ΔG_{37}° bulge, versus free energy of second least stable duplex stem, as described in text. Group III bulge indicated by black circle. Lines are the least-squares fit of the data points (group I and II, gray solid line [Bloese et al. 2007]; black group III, black dotted line; all, black solid line).

preference for the group III bulged nucleotide to be the nucleotide further from the hairpin loop irrespective of whether the resulting hairpin formed the Watson–Crick or Wobble base pair. Factors beyond thermodynamics are clearly important in bulge loop selection and structure in native structures.

Naturally occurring group III bulge loops

An interesting bulge loop in the context of a hairpin stem is position 2402 of the large ribosomal subunit. The secondary and tertiary structures of the bulge loop are shown in Figure 3 (Schuwirth et al. 2005; Selmer et al. 2006). In *Thermus thermophilus*, the bulged nucleotide, a cytosine residue, is 2 bp from the hairpin loop. The analogous position in *E. coli* is a uracil residue, which is also located 2 bp from the hairpin loop. In both cases, the bulge loops are of the group IV type, having sequence characteristics of both group II and III bulges. That is, there is ambiguity as the bulged nucleotide could be represented by any of the 4 nt on the 5' side of the hairpin loop. Both bulge loops are positioned in an extrahelical orientation and are pointing into an “open” pocket where there is no discernible interaction with other components of the ribosome. However, they point in opposite directions with the bulged nucleotide in *T. thermophilus* pointing away from the hairpin loop, while the *E. coli* bulge is located alongside the hairpin loop. It is possible that, in fact, the bulge loops are interacting with some unresolved structural component of the ribosome that leads to the specificity of the bulged nucleotide identity and position. Alternately, the bulged nucleotide may be directed by the wobble base pair next to the bulged position in *T. thermophilus* rRNA.

A second example of a naturally occurring group III bulge loop is position 70 in the RNase P RNA of *E. coli*. In the native RNase P, the bulge is localized in a duplex. Schmitz and Tinoco (2000) has examined the structure of the bulged loop in the context of a hairpin stem as shown in Figure 3.

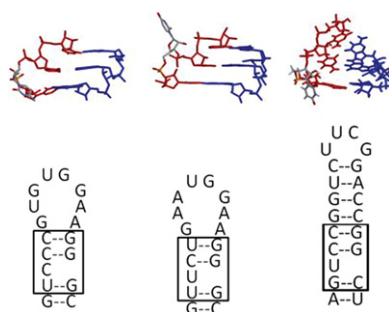


FIGURE 3. Secondary and tertiary structure of naturally occurring bulge loops. Position 2402 in the large ribosomal subunit RNA, (left) *T. thermophilus*, (center) *E. coli*, and (right) position 70 in *E. coli* RNase P RNA. Boxed nucleotides in the secondary structure diagrams are shown in the three-dimensional presentations. The bulged nucleotide is shown in CPK color and the nearest-neighbor pairs in red (same strand as the bulge) and blue (opposite strand). Extracted from PDB files 2AW4, 2J03, and 1F7E.

The bulged nucleotide is a uracil residue and represents the base located further from the hairpin loop. In this context, the bulge is positioned away from the hairpin loop. Interestingly, if the cytosine residue of the bulge is mutated to uracil, producing a group II bulge, the bulge undergoes a rearrangement, and now the base closer to the hairpin loop becomes the bulged nucleotide.

These examples emphasize the structural flexibility imparted into the RNA when a bulge is introduced. Further investigation will be needed to understand the forces responsible for the selection of the correct bulged nucleotide and its orientation in a complex structural motif.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for many helpful comments that greatly improved the manuscript. This work was supported by National Science Foundation grant MCB-0744631 and the Paul E. and Mildred L. Hill Funds of Allegheny College.

Received October 30, 2013; accepted February 21, 2014.

REFERENCES

- Allawi HT, SantaLucia J Jr. 1997. Thermodynamics and NMR of internal G-T mismatches in DNA. *Biochemistry* **36**: 10581–10594.
- Altman S. 1990. Ribonuclease P. *Postscript J Biol Chem* **265**: 20053–20056.
- Bailor MH, Sun X, Al-Hashimi HM. 2010. Topology links RNA secondary structure with global conformation, dynamics, and adaptation. *Science* **327**: 202–206.
- Banerjee AR, Jaeger JA, Turner DH. 1993. Thermal unfolding of a group I ribozyme: the low temperature transition is primarily a disruption of tertiary structure. *Biochemistry* **32**: 153–163.
- Blose J, Manni M, Klapac K, Stranger-Jones Y, Zyra A, Sim V, Griffith C, Long J, Serra MJ. 2007. Sequence dependence of stability for RNA bulge loops based on the complete set of group I single nucleotide bulge loops. *Biochemistry* **46**: 15123–15135.
- Borer PN. 1975. In *Handbook of biochemistry and molecular biology: nucleic acids* (ed. Fasman GD), p. 589. CRC Press, Cleveland, OH.
- Borer P, Dengler B, Tinoco I Jr. 1974. Stability of ribonucleic acid double-stranded helices. *J Mol Biol* **86**: 843–853.
- Brion P, Westhof E. 1997. Hierarchy and dynamics of RNA folding. *Annu Rev Biophys Biomol Struct* **26**: 113–137.
- Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, Feng B, Lin N, Madabusi LV, Muller KM, et al. 2002. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* **3**: 2.
- Cech TR. 1990. Self splicing of group I introns. *Annu Rev Biochem* **59**: 543–568.
- Crothers DM, Cole PE, Hibers CW, Schulman RG. 1974. The molecular mechanism of thermal unfolding of Escherichia coli formylmethionine transfer RNA. *J Mol Biol* **87**: 63–88.
- Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Neilson T, Turner DH. 1986. Improved free-energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci* **83**: 9373–9377.
- Gautheret D, Major F, Cedergren R. 1993. Modeling the three-dimensional structure of RNA using discrete nucleotide conformation sets. *J Mol Biol* **229**: 1049–1064.
- 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.

- Klasens BI, Thiesen M, Virtanen A, Berkhout B. 1999. The ability of HIV-1 AAUAAA signal to bind polyadenylation factors is controlled by local RNA structure. *Nucleic Acids Res* **27**: 446–454.
- Lancialt C, Champoux JJ. 2005. Effects of unpaired nucleotides within HIV-1 genomic secondary structures on pausing and strand transfer. *J Biol Chem* **280**: 2413–2423.
- Lim GFS, Merz GE, McCann MD, Gruskiewicz JM, Serra MJ. 2012. Stability of single nucleotide bulge loops embedded in a GAAA RNA hairpin stem. *RNA* **18**: 807–814.
- Longfellow CE, Kierzek R, Turner DH. 1990. Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides. *Biochemistry* **29**: 278–285.
- 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, Turner DH. 2002. Dynalign: an algorithm for finding the secondary structure common to two RNA sequences. *J Mol Biol* **317**: 191–203.
- Mathews DH, Sabina J, Zuker M, Turner DH. 1999. Expanded sequence dependence of thermodynamic parameters improves 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 GFS, 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.
- Michel F, Westhof E. 1990. Modeling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol* **216**: 585–610.
- Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. 2000. The structural basis of ribosomal activity in peptide bond synthesis. *Science* **289**: 920–930.
- Noller HF, Hoffarth V, Zimmick L. 1992. Unusual resistance of peptidyl transferase to protein extraction. *Science* **265**: 3587–3590.
- Onoa B, Dumont S, Liphardt J, Smith SB, Tinoco I Jr, Bustamante C. 2003. Identifying kinetic barriers to mechanical unfolding of the *T. thermophila* ribozyme. *Science* **299**: 1892–1895.
- Pace NR, Brown JW. 1995. Evolutionary perspective on the structure and function of ribonuclease P, a ribozyme. *J Bacteriol* **177**: 1919–1928.
- Portmann S, Grimm S, Workman C, Usman N, Egli M. 1996. Crystal structures of an A-form duplex with single-adenosine bulges and a conformational basis for site-specific RNA self-cleavage. *Chem Biol* **3**: 173–184.
- Regulski EE, Breaker RR. 2008. In-line probing analysis of riboswitches. *Methods Mol Biol* **419**: 53–67.
- Richards EG. 1975. In *Handbook of biochemistry and molecular biology: nucleic acids* (ed. Fasman GD), p. 197. CRC Press, Cleveland, OH.
- Rounseville MP, Kumar A. 1992. Binding of a host cell nuclear protein to the stem region of human immunodeficiency virus type 1 *trans*-activation-responsive RNA. *J Virol* **66**: 1688–1694.
- Rounseville MP, Lin HC, Agbottah E, Shukla RR, Rabson AB, Kumar A. 1996. Inhibition of HIV-1 replication in viral mutants with altered TAR RNA stem structures. *Virology* **216**: 411–417.
- Schmitz M, Tinoco I Jr. 2000. Solution structure and metal-ion binding of the P4 element from bacterial RNase P RNA. *RNA* **6**: 1212–1225.
- Schuwirth BS, Borovinskaya MA, Hau CW, Zhang W, Vila-Sanjurjo A, Holton JM, Cate JH. 2005. Structures of the bacterial ribosome at 3.5 Å resolution. *Science* **310**: 827–834.
- Scott WG. 2007. Ribozymes. *Curr Opin Struct Biol* **17**: 280–286.
- Selmer M, Dunham CM, Murphry FV, Weilbaumer A, Petry S, Kelly AC, Weir JR, Ramakrishnan V. 2006. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**: 1935–1942.
- 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.
- Soukup GA, Breaker RR. 1999. Relationship between internucleotide linkage geometry and the stability of RNA. *RNA* **5**: 1308–1325.
- Taher F, Gouy M, Regnier M. 2002. Automatic RNA secondary structure prediction with a comparative approach. *Comput Chem* **26**: 521–530.
- Tinoco I Jr, Bustamante C. 1999. How RNA folds. *J Mol Biol* **293**: 271–281.
- Usman N, Ogilvie KK, Jiang M-Y, Cedergren R. 1987. 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.
- 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.
- Woodson SA. 2000. Recent insights on RNA folding mechanisms from catalytic RNA. *Cell Mol Life Sci* **57**: 796–808.
- 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.
- Zarrinkar PP, Williamson JR. 1994. Kinetic intermediates in RNA folding. *Science* **265**: 918–927.
- Znosko BM, Silvestri SB, Volkman H, Boswell B, Serra MJ. 2002. Thermodynamic parameters for an expanded nearest-neighbor model for the formation of RNA duplexes with single nucleotide bulges. *Biochemistry* **41**: 10406–10417.
- Zuker M. 1989. On finding all suboptimal foldings of an RNA molecule. *Science* **244**: 48–52.



RNA

A PUBLICATION OF THE RNA SOCIETY

Non-nearest-neighbor dependence of stability for group III RNA single nucleotide bulge loops

Jessica L. Kent, Michael D. McCann, Daniel Phillips, et al.

RNA 2014 20: 825-834 originally published online April 17, 2014

Access the most recent version at doi:[10.1261/ma.043232.113](https://doi.org/10.1261/ma.043232.113)

References

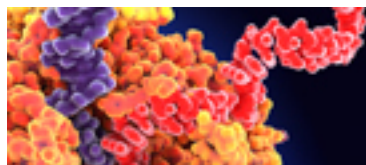
This article cites 47 articles, 16 of which can be accessed free at:
<http://rnajournal.cshlp.org/content/20/6/825.full.html#ref-list-1>

Creative Commons License

This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

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](#).



Use CRISPRmod for targeted modulation of endogenous gene expression to validate siRNA data



To subscribe to *RNA* go to:

<http://rnajournal.cshlp.org/subscriptions>
