Published in final edited form as:

Biochemistry. 2018 April 10; 57(14): 2121–2131. doi:10.1021/acs.biochem.7b01306.

Surprising Sequence Effects on GU Closure of Symmetric 2x2 Nucleotide RNA Internal Loops

Kyle D. Berger^{†,‡}, Scott D. Kennedy^{†,‡}, Susan J. Schroeder[§], Brent M. Znosko^{||}, Hongying Sun^{†,‡}, David H. Mathews^{†,‡}, and Douglas H. Turner^{*,#,‡}

[†]Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642

[‡]Center for RNA Biology, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642

[§]Department of Chemistry, University of Oklahoma, Norman, OK 73019

^{||}Department of Chemistry, Saint Louis University, St. Louis MO 63103

[#]Department of Chemistry, University of Rochester, Rochester, NY 14627

Abstract

GU base pairs are important RNA structural motifs and often close loops. Accurate prediction of RNA structures relies upon understanding the interactions determining structure. The thermodynamics of some two-by-two nucleotide internal loops closed by GU pairs are not well understood. Here, several self-complementary oligonucleotide sequences expected to form duplexes with two-by-two nucleotide internal loops closed by GU pairs were investigated. Surprisingly, NMR revealed that many of the sequences exist in equilibrium between hairpin and duplex conformations. This equilibrium is not observed with loops closed by Watson-Crick pairs. To measure the thermodynamics of some two-by-two nucleotide internal loops closed by GU pairs, non-self-complementary sequences were designed that preclude formation of hairpins. The measured thermodynamics indicate that some internal loops closed by GU pairs are unusually unstable. This instability accounts for the observed equilibria between duplex and hairpin conformations. Moreover, it suggests that future three dimensional structures of loops closed by GU pairs may reveal interactions that unexpectedly destabilize folding.

TOC image

*Corresponding Author: turner@chem.rochester.edu.

Supporting information

The Supporting Information is available free of charge on the ACS Publications website.

The supporting information contains single stranded melts for the non-self-complementary duplexes, a table showing comparisons of thermodynamics for loops with and without a 3' dangling end A, NMR for RNA sequences at 1M NaCl, and additional NMR dilution spectra.

ORCID: Douglas H. Turner: 000-0003-3853-8271

Notes: The authors declare no competing financial interest

	GA	AG	UU	GG	CA	CU	UC	CC	AC	AA
GC	-2.5	-1.4	-0.6	-0.8	1.2	1.3	1.3	1.4	0.8	1.4
CG	-1	-0.7	-0.1	0.6	1.4	1.6	1.7	1.9	1.9	1.5
UA	0.6	1.4	1	2.2	2.3	2.6	3.3	4.9	3.5	2.7
AU	0.6	2.1	0.9	2.1	3.3	3	3.1	3.8	2.8	3
UG	-0.8	2.7	1.9	1	2.5	3.5	4.3	1.6	2	1.1
GU	0.2	1	1.6	1.7	3.7	5.4	4.1	5.1	3.6	5.4
$\Delta G^\circ < 0$ kcal/mol			$\Delta G^\circ \geq 0$ kcal/mol		$\Delta G^\circ \geq 2$ kcal/mol			$\Delta G^\circ \geq 4$ kcal/mol		
$\Delta G^\circ < 0$ kcal/mol			$\Delta G^\circ < 2$ kcal/mol		$\Delta G^\circ < 4$ kcal/mol			$\Delta G^\circ < 6$ kcal/mol		

INTRODUCTION

Much of biology is encoded in DNA sequences. The library of genomic and transcriptomic DNA and RNA sequences is expanding at a phenomenal rate.¹⁻⁴ Translation of coded information into biological insight would be greatly accelerated if secondary and 3D structures of RNA could be accurately predicted from sequence. Most known RNA secondary structures have been defined by sequence comparison and/or chemical mapping, often combined with basic knowledge of RNA thermodynamics.⁵⁻¹⁵

Three dimensional structures of RNA have been determined by X-ray crystallography, cryoelectron microscopy and NMR and are thus harder to determine than secondary structures.¹⁶⁻²³ Accurate modeling of the intermolecular interactions important for RNA folding would hasten determination of both secondary and 3D structure, but such modeling is still challenging.²⁴⁻²⁸ RNAPuzzles, a blind competition to predict RNA 3D structure, demonstrated the importance of accurate RNA secondary structure prediction for 3D prediction and the need for improvements, especially in modeling of noncanonical motifs.^{25, 28} Thus, detailed structural analysis of RNA loops with idiosyncratic stabilities can provide insight into structure-energetics relationships in RNA and contribute to development of *de novo* prediction rules.

RNA loops are important for function²⁹⁻³¹ and for binding potential therapeutics.³²⁻³⁵ Approximating loop stability and probability from sequence is a key to accurate prediction of secondary structure.^{6, 36, 37} Moreover, small loops provide benchmarks for testing computational approaches for predicting RNA thermodynamic stability and 3D structure.^{7, 38-42} Many previous experimental and computational studies have focused on 2×2 nucleotide (nt) internal loops with GA pairs, which typically have two base-base hydrogen bonds.⁴³⁻⁵⁰ This paper reports thermodynamics and NMR spectra for a series of RNA 2×2 nt loops that contain noncanonical pairs unlikely to have two hydrogen bonds and that are closed by GU pairs.

GU pairs often close loops.⁵¹ For example, 24% of 2×2 nt internal loops in one secondary structure database have at least one GU closing pair.⁵² The thermodynamics and structures of internal loops potentially closed by GU pairs are particularly idiosyncratic.^{43-45, 53, 54} For example, the 5'GAGU/3'UGAG loop in (5'GACGAGUGUCA)₂ has a major and minor structure.⁴³ The major structure is novel, with two GG pairs, extruded U's, and an AA stack^{40, 43, 44} and is thus unexpectedly a 4×4 nt loop. The minor structure is also unusual in

that the GU and AG pairs have rare hydrogen bonding patterns.⁵⁵ This contrasts with duplexes having internal loops, 5'UAGG/3'GGAU, 5'UAGA/3'AGAU, 5'AAGU/3'UGAA, 5'CAGG/3'GGAC and 5'GAGC/3'CGAG, which each have a single structure with two imino (cis WC/WC) AG pairs flanked by either wobble UG or Watson-Crick pairs.⁴³ Thus, the eight nucleotides of these internal loops have at least eight base-base hydrogen bonds compared to the four in the major structure of 5'GAGU/3'UGAG.

Thermodynamic measurements on the above loops closed by Watson-Crick pairs^{53, 56-58} indicate that on average the extra hydrogen bond in a GC as compared to an AU pair stabilizes by 1.4 kcal/mol at 37°C, similar to previous estimates.^{59, 60} The stability of the 5'GAGU/3'UGAG loop, however, is not markedly weaker than that of 5'UAGG/3'GGAU, 5'UAGA/3'AGAU, and 5'AAGU/3'UGAA. Evidently, stabilities and structures of internal loops depend on more than number of hydrogen bonds.

Accurate prediction of 3D structures and dynamics for RNA depends on force fields that realistically balance hydrogen bonding with base stacking, along with other interactions.^{26, 42, 61, 62} The structure of the 5'GAGU/3'UGAG internal loop⁴⁴ suggests that 2×2 nt internal loops closed by GU pairs can provide additional benchmarks for testing that balance.⁴⁰ Optical melting experiments and 1D imino proton NMR spectra of the GU closed loops studied here reveal unexpected equilibria and structures that will challenge computations. The results also update thermodynamic parameters for symmetric 2×2 nt loops closed by GU because revised thermodynamic parameters for nearest neighbor canonical pairs containing at least one GU have been used in the calculations.⁶³

MATERIALS AND METHODS

Sequence design

Non-self-complementary duplexes (AB) were designed to minimize the prevalence of hairpin conformation. The “tic-tac-toe” method⁶⁴ was used to minimize base pairing in hairpins (strands A and B) and in self-complementary duplexes (AA and BB).

RNA Preparation

Oligonucleotides were purchased from Dharmacon. RNA was dissolved in either melting buffer (1.0 M NaCl, 20 mM sodium cacodylate, 0.5 mM Na₂EDTA, at pH=7.0) or NMR buffer (90:10 (v/v) H₂O/D₂O, 80 mM NaCl, 20 mM NaH₂PO₄ (self-complementary duplexes) or 50 mM NaH₂PO₄ (non-self-complementary duplexes), 0.05 mM Na₂EDTA, at pH 6.1-6.3). RNA was incubated at 80 °C for 3 min and then slow cooled to room temperature to anneal the duplexes.

UV Melting

Melting curves for duplexes and for each single strand used for non-self-complementary duplexes were measured at 280 nm with a Beckman Coulter DU 640 spectrophotometer. Oligonucleotide concentrations ranged between 1 μM and 200 μM as calculated from absorbance at 80 °C and using extinction coefficients from Borer⁶⁵ and Richards.⁶⁶ Data collection began at 12 °C and was ramped to about 80 °C at a rate of 1 °C/min.

Melting Analysis

Thermodynamic parameters (ΔH° and ΔS°) were extracted from melting curves by curve fitting. Melting curves were fit to a two-state model using MeltWin as previously described.⁶⁷ In addition, thermodynamic parameters were determined by measuring the concentration dependence of melting temperature, T_M , and plotting $1/T_M$ versus $\ln(C_T/a)$:

$$1/T_M = (R/\Delta H^\circ)\ln(C_T/a) + \Delta S^\circ/\Delta H^\circ \quad (1)$$

Here, a is 1 or 4, respectively, for self-complementary and non-self-complementary duplexes. Melting was deemed to be two-state if the ΔH° obtained by the concentration dependence of T_M and by curve fitting agreed within 15%.

With the exception of 5'GCGUGCCUUGCG/3'CGCAUCCGACGC, the single strands either had no cooperative melting or an apparent T_M lower than the heteroduplexes (Figures S1-S6). The two strands for 5'GCGUGCCUUGCG/3'CGCAUCCGACGC yielded melting curves with clear cooperative transitions and similar T_M values (Figure S4). In this case, 1D NMR imino proton spectra and 2D NOESY spectra at 10 and 25 °C and 0.1 M Na⁺ revealed that the expected heterodimer was more than 90% of the sample. The remainder was in self-complementary duplexes composed of each individual strand. The melting curves at 1 M Na⁺ for the mixture of 5'CGCAGCCUACGC/3'GCGUCCGUGCG were treated as a simple two state transition. Agreement by better than 15% for ΔH° values derived from eq 1 and by curve fitting suggest the self-complementary duplexes had little effect on the measured thermodynamics.

Calculation of internal loop thermodynamics

Previously, two different methods have been used to calculate thermodynamic increments for internal loops. When thermodynamics for a duplex with the same Watson-Crick/GU pairs but no internal loop were available, then those values were subtracted from measured values for the duplex and then adjusted for the Watson-Crick or GU nearest neighbor missing from the duplex with the internal loop.^{68, 69} If a relevant all Watson-Crick/GU duplex was not available, then published nearest neighbor parameters^{36, 70} were subtracted from the measured values of the duplex, $\Delta G^\circ_{\text{duplex}}$, containing the internal loop. For duplexes with terminal GC pairs, the equation is:

$$\Delta G^\circ_{\text{loop}} = \Delta G^\circ_{\text{duplex}} - \sum \Delta G^\circ_{\text{NN}} - \Delta G^\circ_{\text{init}} - \Delta G^\circ_{\text{sym}} \quad (2)$$

Here, $\sum \Delta G^\circ_{\text{NN}}$ is the sum of the nearest neighbor parameters for Watson-Crick and GU pairs, $\Delta G^\circ_{\text{init}}$ is the free energy increment for duplex initiation, and $\Delta G^\circ_{\text{sym}} = -T \Delta S^\circ_{\text{sym}}$ accounts for the symmetry difference between self-complementary ($\Delta S^\circ_{\text{sym}} = -1.4$ cal/K mol) and non-self-complementary duplexes ($\Delta S^\circ_{\text{sym}} = 0$ cal/K mol).⁷¹ For consistency here, eq 2 has been used for all 2×2 nt internal loops, including those previously measured. Additionally, updated values for nearest neighbor GU pairs have been used.⁶³ Thus, some loop ΔH° , ΔS° , and ΔG° values differ from those reported in previous publications.

As an example of the calculations, the G° increment at 37 °C (eq 2) for formation of the loop, 5'GCAU/3'UACG, in the duplex, (CGGGCAUCCG)₂, was calculated as:

$$\Delta G^\circ_{37}(5'GCAU/3'UACG) = -6.74 - 2(-2.36) - 2(-3.26) - 2(-1.80) - 4.09 - 0.43 \\ = 3.58 \text{ kcal/mol at } 37^\circ\text{C}.$$

(3)

Here, -6.74 kcal/mol is the measured G°_{37} of duplex formation. The other numbers are nearest neighbor parameters taken from Xia et al. and J. Chen et al.^{63, 70} The calculated value of 3.58 kcal/mol for loop formation is more stable than the 6.00 kcal/mol previously reported.⁷² The difference is that the previous calculation subtracted the measured G°_{37} of -11.2 kcal/mol for (5'CGGGUCCG)₂⁷³ and added the best value of 1.29 kcal/mol for the (5'GU/3'UG) middle nearest neighbor available at that time.³⁶ Use of the measured value for (5'CGGGUCCG)₂, however, neglects the fact that the 5'GGUC/3'CUGG quartet is unusually stable^{63, 73} but absent in (5'CGGGCAUCCG)₂.

For sequences that formed only hairpins, the loop was treated as a hairpin loop of four nucleotides because we are unaware of a GXYU loop with the GU pair in a wobble conformation.⁷⁴⁻⁷⁷ This is consistent with previous calculations of hairpin thermodynamics.^{78, 79}

NMR

NMR spectra were measured with either a 500 MHz or 600 MHz Varian Inova spectrometer. RNA concentrations for NMR were ~ 1 mM unless otherwise noted. One-dimensional imino proton spectra were obtained using a 1-1-spin-echo pulse⁸⁰ with excitation maximum adjusted to the imino region. Two-dimensional NOESY spectra were obtained using WATERGATE water suppression.^{81, 82}

RESULTS

Studies started with self-complementary sequences⁷² designed to form symmetric duplexes with a central 2×2 nt internal loop. As discussed below, however, there were conditions in which NMR showed that the sequences can also form a hairpin, thus revealing an unexpected conformational pliability for GU pairs.

Self-Complementary Oligonucleotides

5'CGGGCAUCCG forms a duplex between 0.13 and 1 mM oligonucleotide—A plot of $1/T_M$ vs $\ln C_T$ for 5'CGGGCAUCCG at 1 M Na⁺ was linear with a positive slope between oligonucleotide concentrations of 12.5 and 140 μ M, but turned over at lower concentrations (Figure 1). A fit of the linear portion of the $1/T_M$ vs $\log C_T$ plot in Figure 1 to eq 1 provided the thermodynamic parameters listed in Table 1 for formation of the duplex, (5'CGGGCAUCCG)₂. The thermodynamic values are within experimental error of those previously reported (Figure 1 caption and Table 2).⁷²

One-dimensional NMR at 0.1 and 1 M Na⁺ revealed the presence of one conformation at 1 mM and 100 μM oligonucleotide, respectively, but more than one conformation at 20 μM oligonucleotide and 0.1 M Na⁺ (Figure 2 and Figure S7). Presence of a hairpin conformation at low RNA concentration could explain the non-linear portion of the 1/T_M vs ln C_T plot in Figure 1. The 1D NMR spectrum at 100 μM oligonucleotide and 1 M Na⁺ (Figure S7) had imino proton resonances equivalent to those seen at 1 mM oligonucleotide and 0.1 M Na⁺. Thermodynamic experiments are typically done at 1 M Na⁺ where duplex formation is favored relative to hairpin.

5'CGGGCUUCCG forms a hairpin at 1 mM oligonucleotide—Surprisingly, replacing the central A of 5'CGGGCAUCCG with a U resulted in a 1/T_M vs ln C_T plot with melting temperature essentially independent of oligonucleotide concentration at 1 M Na⁺ (Figure 3). This is consistent with a hairpin to random coil transition. Imino proton NMR spectra at 50 μM and 1 mM of 5'CGGGCUUCCG at 0 °C and 0.1 M Na⁺ each have six clear resonances (Figure 4a). A 2D NOESY spectrum (150 ms mixing time) confirmed three CG pairs. G4H1 was confirmed by a strong imino (11.04 ppm) to amino cross-peak. A positive ROESY (10 ms mixing time) cross-peak between 11.04 and 13.49 ppm for G4H1 indicated slow conformational exchange. This chemical shift difference is consistent with a hairpin loop having two conformations. One possible minor conformation is a “kissing hairpin” with two GC pairs^{22, 84, 85} and a stabilizing 3' dangling end U (Figure S8).^{71, 86-88}

Fitting of the optical melting curves for 5'CGGGCUUCCG to a unimolecular transition from hairpin to random coil provided the thermodynamic parameters in Table 3. The thermodynamics of hairpin formation have been studied for sequences with related loops, 5'GGACGCUUGUCC, 5'GGACUUCGGUCC, and 5'GGCUUCGGCC.^{78, 79} The first sequence has a CG instead of GC pair closing the loop and the last two sequences have the loop sequence including closing pair in reverse compared to 5'CGGGCUUCCG. These hairpin results are also listed in Table 3, with the loop parameters updated with current nearest neighbor parameters for the Watson-Crick paired stems.⁷⁰

5'GGCGAAUGCC forms a hairpin at 1 mM oligonucleotide—Replacing the central C of 5'CGGGCAUCCG with an A while maintaining a stem of three GC pairs resulted in a concentration independent melting temperature for 5'GGCGAAUGCC at 1 M Na⁺ (Figure 3). This is consistent with a hairpin to random coil transition. NMR spectra at 10 °C, 0.1 M Na⁺ and 50 μM or 1 mM oligonucleotide have four clear resonances, consistent with a single major conformation having three CG pairs and a G or U resonance at 10.6 ppm (Figure 4b). The concentration independent NMR is also consistent with a hairpin conformation. Fitting of the optical melting curves to a unimolecular transition from hairpin to random coil provided the thermodynamic parameters in Table 3.

5'CCUGUCUAGG, 5'GGCGACUGCC, 5'CCAUUCGUGGA, and 5'GUCGCCUGAC each form an equilibrium mixture of duplex and hairpin—For ~1 mM RNA samples of 5'CCUGUCUAGG, 5'GGCGACUGCC, 5'CCAUUCGUGGA, and 5'GUCGCCUGAC, 1D imino proton NMR spectra at 0.1 M Na⁺ have more resonances than expected for a single conformation (Figures 4c,d, S9, S10). Moreover, the relative areas of resonances change with oligonucleotide concentration (Figure 4c,d, S9, S10). Increasing

Na⁺ to 1 M for 5'CCUGUCUAGG and 5'CCAUUCGUGGA also changed relative intensities of resonances (Figures S11 and S12). For 5'CCAUUCGUGGA, the chemical shifts of 10.3, 11.3, and 11.95 ppm for the imino resonances retained at 50 μ M RNA (Figure S10) are in a similar pattern as those reported for the UUCG tetraloop in 5'GGCACUUCGGUGCC.⁷⁷ Evidently, 5'CCUGUCUAGG, 5'GGCGACUGCC, 5'CCAUUCGUGGA, and 5'GUCGCCUGAC, exist in equilibria between duplex and hairpin.

5'CGGACAUCCG, 5'GAGACUUCUC and 5'GCAUCUGC form duplexes at 0.1 mM oligonucleotide and 1 M NaCl—To check for possible duplex/hairpin equilibria in loops with AU closure, imino proton spectra at 1 M Na⁺ were measured for 5'CGGACAUCCG, 5'GAGACUUCUC, and 5'GCAUCUGC (Figure 5). No extra imino proton resonances were detected, consistent with only duplex formation. Moreover, for 5'CGGACAUCCG and 5'GAGACUUCUC, respectively, one and two resonances with chemical shifts above 13.6 ppm were detected, consistent with the number of expected types of Watson-Crick AU pairs in the duplexes. For 5'GCAUCUGC, a 2D spectrum in water revealed a single resonance representative of the expected Watson-Crick AU pairs as identified by a UH3/AH2 NOE. The chemical shift of the UH3 was 13.6 ppm as expected for the proposed duplex. To our knowledge, a hairpin loop of the form 5'AXYU, with the AU forming a Watson-Crick pair and XY being any combination of A, C, G, and U, has never been deposited in the PDB.⁸⁹ Previously, no extra resonances were observed for 5'GCAAAUGC, 5'GGAACUCC, and 5'GGACAUCC.⁸³ Evidently, the duplex/hairpin type equilibria or pure hairpin observed for some loops closed by GU pairs are unlikely to occur with AU closure.

The thermodynamics of duplex formation for (5'GAGACUUCUC)₂ and (5'GCAUCUGC)₂ have been reported.⁷² Thermodynamics for (5'CGGACAUCCG)₂ were measured (Table 1 and Figure S13) to provide values for the 5'ACAU/3'UACA loop. The loop ΔG°_{37} of 3.82 kcal/mol is less favorable than the 2.67 kcal/mol derived from the thermodynamics of duplex formation for (5'GGACAUCC)₂, which had melting temperatures lower by about 20 °C (Tables 1 and 2).⁸³ Low melting temperatures make it difficult to fit lower baselines of melting curves.

Non-Self-Complementary Oligonucleotides that form Duplexes

To avoid equilibria between duplex and hairpin, non-self-complementary oligonucleotides were used to measure thermodynamics for several internal loops closed by GU pairs. Sequences and 1D NMR spectra are shown in Figures 6 and S14. Thermodynamic results are listed in Tables 1 and 2.

DISCUSSION

Insights into intermolecular interactions important for RNA folding can facilitate interpretation of the sequences in RNA transcriptomes. Small model systems provide benchmarks able to bridge experimental results with computations. Here, thermodynamic results are presented for seven 2×2 nt internal loops closed by GU pairs and two hairpin loops.

GU pairs are common in RNA secondary structure and often close internal loops.^{51, 52} Thus, the thermodynamics for closure of loops by GU pairs is important for prediction of secondary structure from sequence by a variety of methods, including sequence comparison,^{5, 9, 10, 15} single sequence with or without chemical modification restraints^{8, 13, 36, 90} and/or NMR constraints.⁹¹

The G°_{37} values for the internal loops studied here in duplexes without 3' dangling A's range from 3.6 to 5.4 kcal/mol (Table 2). This range is noteworthy because every difference of 1.4 kcal/mol changes an equilibrium constant by 10-fold at 37 °C. Thus, the thermodynamic results reported here can provide benchmarks for computational methods⁴² that predict the thermodynamics of RNA loops on the basis of fundamental interactions.

The range of G°_{37} values reported here for internal loops is smaller than the range of 0.4 to 6.0 kcal/mol previously reported.⁷² This difference is largely due to extracting loop parameters using only nearest neighbor parameters,^{63, 70} and elimination of the unexpected equilibria between internal loops and hairpin loops as revealed by NMR.

The (5' GACU)₂ loop was studied in two duplexes. One had terminal CG pairs and the other differed by having them replaced with 3' dangling end A's (Tables 1 and 2 and Figure S14). The respective values for G°_{37} of the loops are 3.6 and 2.2 kcal/mol. Evidently, the nature of termini can affect the contribution of an internal loop to duplex stability. A related non-nearest neighbor effect has previously been observed for bulged loops measured in the presence or absence of 3' dangling end A's.⁹² Clearly, the nearest neighbor model is sometimes a rough approximation when loops are involved. In contrast, it is an excellent model when only canonical base pairs are present^{63, 70} and even in some other cases with 2×2 internal loops (Table S1). Future structural and computational studies may provide insight into fundamental principles determining the applicability of nearest neighbor approximations.

Table 2 summarizes thermodynamics for the internal loops studied here. All the loops fall into category 1, defined as 2×2 nt loops “unlikely to have a mismatch with two hydrogen bonds between bases”.⁹³ For this type of loop closed by GU or AU pairs, parameters determined by Christiansen and Znosko⁵² from a dataset of 97 internal loops with 2×2 nts predict a loop G°_{37} of 3.0±0.2 kcal/mol. For the loops in Table 2, the G°_{37} values for (5' GAAU)₂, (5' GCUU)₂, (5' GCCU)₂, (5' GUCU)₂, and (5' UUCG)₂ may differ from 3.0 kcal/mol by more than experimental error and may reflect non-nearest neighbor effects. Thus, their 3D structures may provide insight into key interactions determining stability.

One surprise for the self-complementary sequences studied here was that under conditions tested, the sequences with GU pairs exist in hairpin/duplex equilibria or were only hairpins. For sequences with hairpin/duplex equilibria, melting curves used to measure thermodynamics do not represent a two-state transition from duplex to single strand. Thus, six non-self-complementary duplexes with GU closed self-complementary internal loops were studied to provide revised internal loop thermodynamics (Tables 1 and 2). Table 4 provides a summary of available values of G°_{37} , H° and S° for all symmetric 2×2 nt internal loops as calculated with eq. 1.

The self-complementary sequence, 5'CGGGCUUCCG, formed only a hairpin (Figures 3 and 4a). The G°_{37} (Table 3) calculated for the hairpin loop, GGCUUC, is +4.8 kcal/mol whereas that calculated for a related loop, CGCUUG, is +3.4 kcal/mol using published experimental data⁷⁸ for 5'GGACGCUUGUCC and nearest neighbor parameters for Watson-Crick pairs.⁷⁰ The average of the two values, 4.1 kcal/mol, is the same as that predicted from parameters of Mathews et al.⁶ when the GU pair in the loop is approximated as an AC pair. As known previously,^{6, 78, 79} the CUUCGG hairpin loop is unusually stable. As expected, the reverse loop sequence, GGCUUC, is less stable (Table 3).

Thermodynamic parameters for the (5'GCUU)₂, (5'GUCU)₂, and (5'GCCU)₂ internal loops were obtained from non-self-complementary duplexes (Table 2). The loop G°_{37} values are very unfavorable at 5.4, 4.1, and 5.1 kcal/mol, respectively. This may reflect unfavorable stacking interactions, backbone distortion, and/or competition with water that leads to reduced hydrogen bonding between G and U as compared to other internal loops. In contrast, the self-complementary sequence, 5'CGGGCUUCCG, forms a hairpin that may be stabilized by non-nearest neighbor interactions previously detected for hairpins closed by a GU pair with the 5'G directly preceded by two G residues.^{36, 95}

The non-self-complementary duplex, 5'GAGUAAUGAC/3'CUCGAAGCUG, has a measured G°_{37} of -5.2 kcal/mol.⁵³ With the nearest neighbor parameters used here,^{63, 70} the G°_{37} increment for the non-self-complementary internal loop, 5'UAAU/3'GAAG, is 3.13 kcal/mol. This is close to the average of 3.25 kcal/mol for the values of 1.1 and 5.4 kcal/mol in Table 4 for the (UAAG)₂ and (GAU)₂ loops. This suggests that the difference between 1.1 and 5.4 kcal/mol may be due to differences of interactions in the 5'UA/3'GA and 5'GA/3'UA nearest neighbors rather than differences in the 5'AA/3'AA stacks.

The duplex, (5'GACGAGUGUCA)₂, unexpectedly exhibits an equilibrium between two 3D conformations,⁴³ and both are novel.^{44, 55} In contrast, replacing the potential GU pairs with AU to give (GACAAGUGUCA)₂ results in a single conformation with two imino AG pairs, as also seen for sequences with the 5'AG/3'GA loop closed by UA, GC, or CG Watson-Crick pairs.^{43, 46} Evidently, small motifs expected to have GU pairs can have structures useful as benchmarks for predicting 3D structure.⁴⁰

The structural heterogeneity of potential GU pairs closing loops presumably reflects the properties of the individual bases. G has the highest dipole moment,⁹⁶ the most atoms capable of hydrogen bonding, and a large surface area, which favors stacking. In contrast, U has the least propensity to stack⁹⁷⁻¹⁰² and a dipole moment roughly half that of G. These properties allow many different patterns of hydrogen bonding for GU pairs.¹⁰³ Moreover, the range (1.1 to 5.4 kcal/mol) in G°_{37} values for 2×2 nt loops closed by GU or UG and with noncanonical pairs having less than two base-base hydrogen bonds is larger than the range (2.3 to 4.9 kcal/mol) for those closed by AU or UA pairs or by GC or CG pairs (0.8 to 1.9 kcal/mol) (Table 4).

In general, the unexpected presence of hairpins in solutions of the self-complementary sequences studied here may be partially a consequence of unfavorably constrained stacking forced by duplex formation. The results reported here suggest that experimental structures

and thermodynamics of loops closed by GU pairs can provide stringent benchmarks for testing the accuracy of RNA force fields in balancing hydrogen bonding, stacking, and backbone interactions. In particular, loops with unusual thermodynamics may also have unusual structures. The experiments reported here were conducted between pH 6 and 7, the pH range most important physiologically. Results at different pHs and/or in the presence of Mg^{2+} could provide additional challenges for computations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Jonathan L. Chen for helping to demonstrate RNA melting methods and analysis. The authors also thank Andrew D. Kauffmann for experimental advice and helpful conversation.

Funding: This work was supported by NIH grants R01GM22939 to D. H. T, R01GM076485 to D. H. M., and 2R15GM085699-03 to B. M. Z. Additional support for K. D. B. was provided by the NIH training grant T32GM68411-10.

References

1. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009; 10:57–63. [PubMed: 19015660]
2. Ding Y, Tang Y, Kwok CK, Zhang Y, Bevilacqua PC, Assmann SM. In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature.* 2014; 505:696–700. [PubMed: 24270811]
3. Talkish J, May G, Lin Y, Woolford JL Jr, McManus CJ. Mod-seq: high-throughput sequencing for chemical probing of RNA structure. *Rna.* 2014; 20:713–720. [PubMed: 24664469]
4. Rouskin S, Zubradt M, Washietl S, Kellis M, Weissman JS. Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature.* 2014; 505:701–705. [PubMed: 24336214]
5. Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, Bateman A, Finn RD, Petrov AI. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.* 2018; 46:D335–D342. [PubMed: 29112718]
6. Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci U S A.* 2004; 101:7287–7292. [PubMed: 15123812]
7. Lorenz R, Bernhart SH, Honer Zu Siederdisen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. ViennaRNA Package 2.0. *Algorithms Mol Biol.* 2011; 6:26. [PubMed: 22115189]
8. Hajdin CE, Bellaousov S, Huggins W, Leonard CW, Mathews DH, Weeks KM. Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. *Proc Natl Acad Sci U S A.* 2013; 110:5498–5503. [PubMed: 23503844]
9. Fu Y, Sharma G, Mathews DH. Dynalign II: common secondary structure prediction for RNA homologs with domain insertions. *Nucleic Acids Res.* 2014; 42:13939–13948. [PubMed: 25416799]
10. Tan Z, Fu Y, Sharma G, Mathews DH. TurboFold II: RNA structural alignment and secondary structure prediction informed by multiple homologs. *Nucleic Acids Res.* 2017; 45:11570–11581. [PubMed: 29036420]
11. Havgaard JH, Gorodkin J. RNA structural alignments, part I: Sankoff-based approaches for structural alignments. *Methods Mol Biol.* 2014; 1097:275–290. [PubMed: 24639164]
12. Asai K, Hamada M. RNA structural alignments, part II: non-Sankoff approaches for structural alignments. *Methods Mol Biol.* 2014; 1097:291–301. [PubMed: 24639165]

13. Wu Y, Shi B, Ding X, Liu T, Hu X, Yip KY, Yang ZR, Mathews DH, Lu ZJ. Improved prediction of RNA secondary structure by integrating the free energy model with restraints derived from experimental probing data. *Nucleic Acids Res.* 2015; 43:7247–7259. [PubMed: 26170232]
14. Cheng CY, Kladwang W, Yesselman JD, Das R. RNA structure inference through chemical mapping after accidental or intentional mutations. *Proc Natl Acad Sci U S A.* 2017; 114:9876–9881. [PubMed: 28851837]
15. Pace, NR., Thomas, BC., Woese, CR. Probing RNA Structure, Function, and History by Comparative Analysis. In: Gesteland, RF.Cech, TR., Atkins, JF., editors. *The RNA World.* Cold Spring Harbor Laboratory Press; 1999. p. 113-141.
16. Ferre-D'Amare AR, Doudna JA. RNA folds: insights from recent crystal structures. *Annu Rev Biophys Biomol Struct.* 1999; 28:57–73. [PubMed: 10410795]
17. Doherty EA, Doudna JA. Ribozyme structures and mechanisms. *Annu Rev Biophys Biomol Struct.* 2001; 30:457–475. [PubMed: 11441810]
18. Furtig B, Richter C, Wohnert J, Schwalbe H. NMR spectroscopy of RNA. *Chembiochem.* 2003; 4:936–962. [PubMed: 14523911]
19. Gu X, Mooers BH, Thomas LM, Malone J, Harris S, Schroeder SJ. Structures and Energetics of Four Adjacent G.U Pairs That Stabilize an RNA Helix. *J Phys Chem B.* 2015; 119:13252–13261. [PubMed: 26425937]
20. Amunts A, Brown A, Bai XC, Llacer JL, Hussain T, Emsley P, Long F, Murshudov G, Scheres SHW, Ramakrishnan V. Structure of the yeast mitochondrial large ribosomal subunit. *Science.* 2014; 343:1485–1489. [PubMed: 24675956]
21. Liu Z, Gutierrez-Vargas C, Wei J, Grassucci RA, Ramesh M, Espina N, Sun M, Tutuncuoglu B, Madison-Antenucci S, Woolford JL Jr, Tong L, Frank J. Structure and assembly model for the *Trypanosoma cruzi* 60S ribosomal subunit. *Proc Natl Acad Sci U S A.* 2016; 113:12174–12179. [PubMed: 27791004]
22. Miyazaki Y, Irobalieva RN, Tolbert BS, Smalls-Mantey A, Iyalla K, Loeliger K, D'Souza V, Khant H, Schmid MF, Garcia EL, Telesnitsky A, Chiu W, Summers MF. Structure of a conserved retroviral RNA packaging element by NMR spectroscopy and cryo-electron tomography. *J Mol Biol.* 2010; 404:751–772. [PubMed: 20933521]
23. Natchiar SK, Myasnikov AG, Kratzat H, Hazemann I, Klaholz BP. Visualization of chemical modifications in the human 80S ribosome structure. *Nature.* 2017; 551:472–477. [PubMed: 29143818]
24. Sripakdeevong, P., Beauchamp, K., Das, R. Why Can't We Predict RNA Structure At Atomic Resolution?. In: Leontis, NB., Westhof, E., editors. *RNA 3D Structure Analysis and Prediction.* Springer; 2012. p. 43-65.
25. Miao Z, Adamiak RW, Blanchet MF, Boniecki M, Bujnicki JM, Chen SJ, Cheng C, Chojnowski G, Chou FC, Cordero P, Cruz JA, Ferre-D'Amare AR, Das R, Ding F, Dokholyan NV, Dunin-Horkawicz S, Kladwang W, Krokhotin A, Lach G, Magnus M, Major F, Mann TH, Masquida B, Matelska D, Meyer M, Peselis A, Popenda M, Purzycka KJ, Serganov A, Stasiewicz J, Szachniuk M, Tandon A, Tian S, Wang J, Xiao Y, Xu X, Zhang J, Zhao P, Zok T, Westhof E. RNA-Puzzles Round II: assessment of RNA structure prediction programs applied to three large RNA structures. *Rna.* 2015; 21:1066–1084. [PubMed: 25883046]
26. Ditzler MA, Otyepka M, Sponer J, Walter NG. Molecular dynamics and quantum mechanics of RNA: conformational and chemical change we can believe in. *Acc Chem Res.* 2010; 43:40–47. [PubMed: 19754142]
27. Cruz JA, Blanchet MF, Boniecki M, Bujnicki JM, Chen SJ, Cao S, Das R, Ding F, Dokholyan NV, Flores SC, Huang L, Lavender CA, Lisi V, Major F, Mikolajczak K, Patel DJ, Philips A, Puton T, Santalucia J, Sijeniyi F, Hermann T, Rother K, Rother M, Serganov A, Skorupski M, Soltysinski T, Sripakdeevong P, Tuszynska I, Weeks KM, Waldsich C, Wildauer M, Leontis NB, Westhof E. RNA-Puzzles: a CASP-like evaluation of RNA three-dimensional structure prediction. *Rna.* 2012; 18:610–625. [PubMed: 22361291]
28. Miao Z, Adamiak RW, Antczak M, Batey RT, Becka AJ, Biesiada M, Boniecki MJ, Bujnicki JM, Chen SJ, Cheng CY, Chou FC, Ferre-D'Amare AR, Das R, Dawson WK, Ding F, Dokholyan NV, Dunin-Horkawicz S, Geniesse C, Kappel K, Kladwang W, Krokhotin A, Lach GE, Major F, Mann TH, Magnus M, Pachulska-Wieczorek K, Patel DJ, Piccirilli JA, Popenda M, Purzycka KJ, Ren A,

- Rice GM, Santalucia J Jr, Sarzynska J, Szachniuk M, Tandon A, Trausch JJ, Tian S, Wang J, Weeks KM, Williams B 2nd, Xiao Y, Xu X, Zhang D, Zok T, Westhof E. RNA-Puzzles Round III: 3D RNA structure prediction of five riboswitches and one ribozyme. *Rna*. 2017; 23:655–672. [PubMed: 28138060]
29. Wyatt, JR., Tinoco, II. RNA Structural Elements. In: Gesteland, RF., Atkins, JF., editors. *The RNA World*. Cold Spring Harbor Laboratory Press; 1993. p. 465-496.
30. Bevilacqua PC, Blose JM. Structures, kinetics, thermodynamics, and biological functions of RNA hairpins. *Annu Rev Phys Chem*. 2008; 59:79–103. [PubMed: 17937599]
31. Atkins, JF., Gesteland, RF., Cech, T. *RNA Worlds: from life's origins to diversity in gene regulation*. Cold Spring Harbor Laboratory Press; New York: 2011.
32. Davidson A, Leeper TC, Athanassiou Z, Patora-Komisarska K, Karn J, Robinson JA, Varani G. Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. *Proc Natl Acad Sci U S A*. 2009; 106:11931–11936. [PubMed: 19584251]
33. Velagapudi SP, Gallo SM, Disney MD. Sequence-based design of bioactive small molecules that target precursor microRNAs. *Nat Chem Biol*. 2014; 10:291–297. [PubMed: 24509821]
34. Shortridge MD, Varani G. Structure based approaches for targeting non-coding RNAs with small molecules. *Curr Opin Struct Biol*. 2015; 30:79–88. [PubMed: 25687935]
35. Childs-Disney JL, Disney MD. Approaches to Validate and Manipulate RNA Targets with Small Molecules in Cells. *Annu Rev Pharmacol Toxicol*. 2016; 56:123–140. [PubMed: 26514201]
36. Mathews DH, Sabina J, Zuker M, Turner DH. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol*. 1999; 288:911–940. [PubMed: 10329189]
37. Sloma MF, Mathews DH. Exact calculation of loop formation probability identifies folding motifs in RNA secondary structures. *RNA*. 2016; 22:1808–1818. [PubMed: 27852924]
38. Turner DH. Fundamental interactions in RNA: Questions answered and remaining. *Biopolymers*. 2013; 99:1097–1104. [PubMed: 23733402]
39. Sripakdeevong P, Kladwang W, Das R. An enumerative stepwise ansatz enables atomic-accuracy RNA loop modeling. *Proc Natl Acad Sci U S A*. 2011; 108:20573–20578. [PubMed: 22143768]
40. Sripakdeevong P, Cevce M, Chang AT, Erat MC, Ziegeler M, Zhao Q, Fox GE, Gao X, Kennedy SD, Kierzek R, Nikonowicz EP, Schwalbe H, Sigel RK, Turner DH, Das R. Structure determination of noncanonical RNA motifs guided by (1)H NMR chemical shifts. *Nat Methods*. 2014; 11:413–416. [PubMed: 24584194]
41. Phan A, Mailey K, Saeki J, Gu X, Schroeder SJ. Advancing viral RNA structure prediction: measuring the thermodynamics of pyrimidine-rich internal loops. *Rna*. 2017; 23:770–781. [PubMed: 28213527]
42. Smith LG, Zhao J, Mathews DH, Turner DH. Physics-based all-atom modeling of RNA energetics and structure. *Wiley Interdisciplinary Reviews: RNA*. 2017; 8:e1422–n/a.
43. Hammond NB, Tolbert BS, Kierzek R, Turner DH, Kennedy SD. RNA internal loops with tandem AG pairs: the structure of the 5'GAGU/3'UGAG loop can be dramatically different from others, including 5'AAGU/3'UGAA. *Biochemistry*. 2010; 49:5817–5827. [PubMed: 20481618]
44. Kennedy SD, Kierzek R, Turner DH. Novel conformation of an RNA structural switch. *Biochemistry*. 2012; 51:9257–9259. [PubMed: 23134175]
45. Tolbert BS, Kennedy SD, Schroeder SJ, Krugh TR, Turner DH. NMR structures of (rGUGAGGCU)₂ and (rGCGGAUGCU)₂: probing the structural features that shape the thermodynamic stability of GA pairs. *Biochemistry*. 2007; 46:1511–1522. [PubMed: 17279616]
46. Wu M, SantaLucia J Jr, Turner DH. Solution structure of (rGGCAGGCC)₂ by two-dimensional NMR and the iterative relaxation matrix approach. *Biochemistry*. 1997; 36:4449–4460. [PubMed: 9109652]
47. SantaLucia J Jr, Turner DH. Structure of (rGGCGAGCC)₂ in solution from NMR and restrained molecular dynamics. *Biochemistry*. 1993; 32:12612–12623. [PubMed: 8251479]
48. Aytenfisu AH, Spasic A, Seetin MG, Serafini J, Mathews DH. Modified Amber Force Field Correctly Models the Conformational Preference for Tandem GA pairs in RNA. *J Chem Theory Comput*. 2014; 10:1292–1301. [PubMed: 24803859]

49. Morgado CA, Svozil D, Turner DH, Sponer J. Understanding the role of base stacking in nucleic acids. MD and QM analysis of tandem GA base pairs in RNA duplexes. *Phys Chem Chem Phys*. 2012; 14:12580–12591. [PubMed: 22722325]
50. Yildirim I, Stern HA, Sponer J, Spackova N, Turner DH. Effects of Restrained Sampling Space and Nonplanar Amino Groups on Free-Energy Predictions for RNA with Imino and Sheared Tandem GA Base Pairs Flanked by GC, CG, iGiC or iCiG Base Pairs. *J Chem Theory Comput*. 2009; 5:2088–2100. [PubMed: 20090924]
51. Gautheret D, Konings D, Gutell RR. G.U base pairing motifs in ribosomal RNA. *Rna*. 1995; 1:807–814. [PubMed: 7493326]
52. Christiansen ME, Znosko BM. Thermodynamic characterization of tandem mismatches found in naturally occurring RNA. *Nucleic Acids Res*. 2009; 37:4696–4706. [PubMed: 19509311]
53. Schroeder SJ, Turner DH. Thermodynamic stabilities of internal loops with GU closing pairs in RNA. *Biochemistry*. 2001; 40:11509–11517. [PubMed: 11560499]
54. Znosko BM, Kennedy SD, Wille PC, Krugh TR, Turner DH. Structural features and thermodynamics of the J4/5 loop from the *Candida albicans* and *Candida dubliniensis* group I introns. *Biochemistry*. 2004; 43:15822–15837. [PubMed: 15595837]
55. Spasic A, Kennedy SD, Needham L, Manoharan M, Kierzek R, Turner DH, Mathews D. Molecular Dynamics Correctly Models the Unusual Major Conformation of the GAGU RNA Internal Loop and with NMR Reveals an Unusual Minor Conformation. *RNA*. 2018
56. Walter AE, Wu M, Turner DH. The stability and structure of tandem GA mismatches in RNA depend on closing base pairs. *Biochemistry*. 1994; 33:11349–11354. [PubMed: 7537087]
57. SantaLucia J Jr, Kierzek R, Turner DH. Effects of GA mismatches on the structure and thermodynamics of RNA internal loops. *Biochemistry*. 1990; 29:8813–8819. [PubMed: 2271557]
58. Xia T, McDowell JA, Turner DH. Thermodynamics of nonsymmetric tandem mismatches adjacent to G.C base pairs in RNA. *Biochemistry*. 1997; 36:12486–12497. [PubMed: 9376353]
59. Turner DH, Sugimoto N, Kierzek R, Dreiker SD. Free energy increments for hydrogen bonds in nucleic acid base pairs. *Journal of the American Chemical Society*. 1987; 109:3783–3785.
60. Freier SM, Sugimoto N, Sinclair A, Alkema D, Neilson T, Kierzek R, Caruthers MH, Turner DH. Stability of XGCGCp, GCGCYp, and XGCGCYp helices: an empirical estimate of the energetics of hydrogen bonds in nucleic acids. *Biochemistry*. 1986; 25:3214–3219. [PubMed: 3730357]
61. Vangaveti S, Ranganathan SV, Chen AA. Advances in RNA molecular dynamics: a simulator's guide to RNA force fields. *Wiley Interdisciplinary Reviews: RNA*. 2017; 8:n/a–n/a.
62. Turner DH, Sugimoto N, Freier SM. RNA structure prediction. *Annu Rev Biophys Biophys Chem*. 1988; 17:167–192. [PubMed: 2456074]
63. Chen JL, Dishler AL, Kennedy SD, Yildirim I, Liu B, Turner DH, Serra MJ. Testing the nearest neighbor model for canonical RNA base pairs: revision of GU parameters. *Biochemistry*. 2012; 51:3508–3522. [PubMed: 22490167]
64. Schroeder SJ, Turner DH. Optical melting measurements of nucleic acid thermodynamics. *Methods Enzymol*. 2009; 468:371–387. [PubMed: 20946778]
65. Borer P. Optical properties of nucleic acids, absorption and circular dichroism spectra. In: Fasman, GD., editor. *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*. 3rd. CRC Press; Cleveland, OH: 1975. p. 589-595.
66. Richards, EG. Use of tables in calculation of absorption, optical rotatory dispersion and circular dichroism of polyribonucleotides. In: Fasman, GD., editor. *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*. 3rd. CRC Press; Cleveland, OH: 1975. p. 589-595.
67. McDowell JA, Turner DH. 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*. 1996; 35:14077–14089. [PubMed: 8916893]
68. Gralla J, Crothers DM. Free energy of imperfect nucleic acid helices. 3. Small internal loops resulting from mismatches. *J Mol Biol*. 1973; 78:301–319. [PubMed: 4747633]
69. SantaLucia J Jr, Kierzek R, Turner DH. Stabilities of consecutive A.C, C.C, G.G, U.C, and U.U mismatches in RNA internal loops: Evidence for stable hydrogen-bonded U.U and C.C.+ pairs. *Biochemistry*. 1991; 30:8242–8251. [PubMed: 1714301]

70. Xia T, SantaLucia J Jr, Burkard ME, Kierzek R, Schroeder SJ, Jiao X, Cox C, Turner DH. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry*. 1998; 37:14719–14735. [PubMed: 9778347]
71. Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Neilson T, Turner DH. Improved free-energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci U S A*. 1986; 83:9373–9377. [PubMed: 2432595]
72. Christiansen ME, Znosko BM. Thermodynamic characterization of the complete set of sequence symmetric tandem mismatches in RNA and an improved model for predicting the free energy contribution of sequence asymmetric tandem mismatches. *Biochemistry*. 2008; 47:4329–4336. [PubMed: 18330995]
73. He L, Kierzek R, SantaLucia J, Walter AE, Turner DH. Nearest-neighbor parameters for G.cntdot.U mismatches: 5'GU3'/3'UG5' is destabilizing in the contexts CGUG/GUGC, UGUA/AUGU, and AGUU/UUGA but stabilizing in GGUC/CUGG. *Biochemistry*. 1991; 30:11124–11132. [PubMed: 1718426]
74. Serganov A, Huang L, Patel DJ. Structural insights into amino acid binding and gene control by a lysine riboswitch. *Nature*. 2008; 455:1263–1267. [PubMed: 18784651]
75. Auerbach T, Mermershtain I, Davidovich C, Bashan A, Belousoff M, Wekselman I, Zimmerman E, Xiong L, Klepacki D, Arakawa K, Kinashi H, Mankin AS, Yonath A. The structure of ribosome-lankacidin complex reveals ribosomal sites for synergistic antibiotics. *Proc Natl Acad Sci U S A*. 2010; 107:1983–1988. [PubMed: 20080686]
76. Ding F, Lu C, Zhao W, Rajashankar KR, Anderson DL, Jardine PJ, Grimes S, Ke A. Structure and assembly of the essential RNA ring component of a viral DNA packaging motor. *Proc Natl Acad Sci U S A*. 2011; 108:7357–7362. [PubMed: 21471452]
77. Nozinovic S, Furtig B, Jonker HR, Richter C, Schwalbe H. High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA. *Nucleic Acids Res*. 2010; 38:683–694. [PubMed: 19906714]
78. Antao VP, Tinoco I Jr. Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic Acids Res*. 1992; 20:819–824. [PubMed: 1371866]
79. Dale T, Smith R, Serra MJ. A test of the model to predict unusually stable RNA hairpin loop stability. *Rna*. 2000; 6:608–615. [PubMed: 10786851]
80. Sklená V, Bax A. Spin-echo water suppression for the generation of pure-phase two-dimensional NMR spectra. *Journal of Magnetic Resonance (1969)*. 1987; 74:469–479.
81. Grzesiek S, Bax A. The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements. *Journal of the American Chemical Society*. 1993; 115:12593–12594.
82. Piotto M, Saudek V, Sklenar V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR*. 1992; 2:661–665. [PubMed: 1490109]
83. Wu M, McDowell JA, Turner DH. A periodic table of symmetric tandem mismatches in RNA. *Biochemistry*. 1995; 34:3204–3211. [PubMed: 7533535]
84. Kim CH, Tinoco I Jr. A retroviral RNA kissing complex containing only two G.C base pairs. *Proc Natl Acad Sci U S A*. 2000; 97:9396–9401. [PubMed: 10931958]
85. Li PT, Bustamante C, Tinoco I Jr. Unusual mechanical stability of a minimal RNA kissing complex. *Proc Natl Acad Sci U S A*. 2006; 103:15847–15852. [PubMed: 17043221]
86. Freier SM, Burger BJ, Alkema D, Neilson T, Turner DH. Effects of 3' dangling end stacking on the stability of GGCC and CCGG double helices. *Biochemistry*. 1983; 22:6198–6206.
87. Freier SM, Alkema D, Sinclair A, Neilson T, Turner DH. Contributions of dangling end stacking and terminal base-pair formation to the stabilities of XGGCCp, XCCGGp, XGGCCYp, and XCCGGYp helices. *Biochemistry*. 1985; 24:4533–4539. [PubMed: 4063336]
88. Burkard ME, Kierzek R, Turner DH. Thermodynamics of unpaired terminal nucleotides on short RNA helices correlates with stacking at helix termini in larger RNAs. *J Mol Biol*. 1999; 290:967–982. [PubMed: 10438596]
89. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H,

- Westbrook JD, Zardecki C. The Protein Data Bank. *Acta Crystallogr D Biol Crystallogr*. 2002; 58:899–907. [PubMed: 12037327]
90. Spasic A, Assmann SM, Bevilacqua PC, Mathews DH. Modeling RNA secondary structure folding ensembles using SHAPE mapping data. *Nucleic Acids Res*. 2018; 46:314–323. [PubMed: 29177466]
91. Chen JL, Bellaousov S, Tubbs JD, Kennedy SD, Lopez MJ, Mathews DH, Turner DH. Nuclear Magnetic Resonance-Assisted Prediction of Secondary Structure for RNA: Incorporation of Direction-Dependent Chemical Shift Constraints. *Biochemistry*. 2015; 54:6769–6782. [PubMed: 26451676]
92. Longfellow CE, Kierzek R, Turner DH. Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides. *Biochemistry*. 1990; 29:278–285. [PubMed: 2322546]
93. Shankar N, Xia T, Kennedy SD, Krugh TR, Mathews DH, Turner DH. NMR reveals the absence of hydrogen bonding in adjacent UU and AG mismatches in an isolated internal loop from ribosomal RNA. *Biochemistry*. 2007; 46:12665–12678. [PubMed: 17929882]
94. Burkard ME, Xia T, Turner DH. Thermodynamics of RNA internal loops with a guanosine-guanosine pair adjacent to another noncanonical pair. *Biochemistry*. 2001; 40:2478–2483. [PubMed: 11327869]
95. Giese MR, Betschart K, Dale T, Riley CK, Rowan C, Sprouse KJ, Serra MJ. Stability of RNA hairpins closed by wobble base pairs. *Biochemistry*. 1998; 37:1094–1100. [PubMed: 9454601]
96. Kollman, P. Theoretical Methods. In: Bloomfield, VA, Crothers, DM., Tinoco, II., editors. *Nucleic Acids: Structures, Properties, and Functions*. University Science Books; 2000. p. 223–258.
97. Condon DE, Kennedy SD, Mort BC, Kierzek R, Yildirim I, Turner DH. Stacking in RNA: NMR of Four Tetramers Benchmark Molecular Dynamics. *J Chem Theory Comput*. 2015; 11:2729–2742. [PubMed: 26082675]
98. Lee CH. Conformational studies of 13 trinucleoside bisphosphates by 360-MHz ¹H-NMR spectroscopy. 1. Ribose protons. *Eur J Biochem*. 1983; 137:347–356. [PubMed: 6317391]
99. Lee CH, Tinoco I. Conformation studies of 13 trinucleoside diphosphates by 360 MHz PMR spectroscopy. a bulged base conformation: I. base protons and H1' protons. *Biophysical Chemistry*. 1980; 11:283–294. [PubMed: 16997249]
100. Chen H, Meisburger SP, Pabit SA, Sutton JL, Webb WW, Pollack L. Ionic strength-dependent persistence lengths of single-stranded RNA and DNA. *Proc Natl Acad Sci U S A*. 2012; 109:799–804. [PubMed: 22203973]
101. Richards EG, Flessel CP, Fresco JR. Polynucleotides. VI. Molecular properties and conformation of polyribouridylic acid. *Biopolymers*. 1963; 1:431–446.
102. Inners LD, Felsenfeld G. Conformation of polyribouridylic acid in solution. *J Mol Biol*. 1970; 50:373–389. [PubMed: 5476918]
103. Leontis NB, Stombaugh J, Westhof E. The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res*. 2002; 30:3497–3531. [PubMed: 12177293]

Figure 1.

Plot of $1/T_M$ vs $\log C_T$ for 5'CGGGCAUCCG at 1 M Na^+ . Error bars are $\pm 1^\circ\text{C}$. The linear fit (shown with the line) to oligonucleotide concentrations between 12.5 and 140 μM gave values for H° , S° , and G°_{37} of -38.0 kcal/mol, -100.9 eu, and -6.74 kcal/mol, respectively. Values are similar to those reported previously: -39.5 kcal/mol, -106.3 eu, and -6.49 kcal/mol, respectively.⁷²

Figure 2. Imino proton NMR spectra of 1 mM and 0.02 mM 5'CGGGCAUCCG in 0.1 M Na⁺ at 5 °C. The additional peaks in the low concentration spectrum are likely due to a small amount of hairpin.

Figure 3.

Plots of $1/T_M$ vs $\log C_T$ at 1 M Na^+ for 5'CGGGCUUCCG (blue), 5'GGCGAAUGCC (green), 5'GCGUGCUUUGCG/3'CGCAUUCGACGC (orange), and 5'GCUGAAUACG/3'CGAUAAAGUGC (black). Error bars are $\pm 1^\circ\text{C}$.

Figure 4.

Imino proton NMR spectra in 0.1 M Na⁺ for: (a) 5'-CGGGCUUCCG; a 2D spectrum contains an exchange peak for G4 at 11.04/13.49 ppm, revealing two conformations in slow exchange, (b) 5'-GGCGAAUGCC, (c) 5'-CCUGUCUAGG; two resonances are overlapped at 13.3 ppm, (d) 5'-GUCGCCUGAC. A 2D spectrum for sequence d reveals that the peak around 12.3 ppm contains two resonances for G8. In general, red and blue resonance assignments refer to hairpins and duplexes, respectively, as determined from 2D spectra.

Figure 5. Imino proton NMR spectra for 0.1 mM RNA in 1 M NaCl at 5 °C for duplexes with AU flanked 2×2 nt loops. (a) (5'CGGACA UCCG)₂, (b) (5'GAGACU UCUC)₂, and (c) (5'GCAUC UGC)₂.

Figure 6.

Imino proton NMR spectra at 0.1 M Na⁺ and ~1 mM duplex for:

- (a) 5'-GCGUGCUUUGCG/3'-CGCAUUCGACGC,
- (b) 5'-GCGUGUCUUGCG/3'-CGCAUCUGACGC,
- (c) 5'-GCGUGAAUAGCG/3'-CGCAUAAGUCGC,
- (d) 5'-GCGUGACUAGCG/3'-CGCAUCAGUCGC,
- (e) 5'-GCGUUUCGUGCG/3'-CGCAGCUUACGC, and

f) 5' GCGUGCCUUGCG/3' CGCAUCCGACGC.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Thermodynamic Parameters for Duplex Formation in 1 M NaCl.

Sequences	T_M^{-1} vs $\ln(C_T/n)$ plots				Average of melt curve fits			
	$-H^\circ$ (kcal/mol)	$-S^\circ$ (eu)	$-G^\circ_{37}$ (kcal/mol)	T_m ($^\circ\text{C}$) ^b	$-H^\circ$ (kcal/mol)	$-S^\circ$ (eu)	$-G^\circ_{37}$ (kcal/mol)	T_m ($^\circ\text{C}$) ^b
(5'CGGGCAUCCG) ₂	38.0±4.6	100.9±14.5	6.74±0.23	45.9	40.6±6.1	108.8±19.3	6.83±0.13	46.1
5'GCGUGACUACGACGCGCAUCAGUCGC5'	80.0±8.4	229.0±26.4	8.95±0.31	46.7	77.6±11.7	221.2±36.8	8.96±0.31	47.0
5'GCUGACUACGACGCAUUCGACGC5'	77.2±3.5	225.0±11.3	7.40±0.04	40.6	88.2±6.5	260.4±20.7	7.45±0.13	40.3
5'GCGUGAAUAGCGCGCAUAAAGUCGC5'	46.2±5.4	125.7±17	7.19±0.24	41.5	46.3±12.0	125.9±38.7	7.28±0.20	42.1
5'GCGUGCUUUGCGCGCAUUCGACGC5'	51.0±5.0	141.4±16.1	7.18±0.15	41.0	50.1±8.5	138.2±27.9	7.26±0.15	41.6
5'GCGUGUCUUGCGCGCAUCUGACGC5'	64.1±7.7	179.6±21.4	8.43±0.25	46.2	73.3±15.3	208.3±48.5	8.71±0.33	46.4
5'GCGUGCCUUGCGCGCAUCCGACGC5'	44.1±4.6	118.3±14.8	7.42±0.23	43.4	51.5±7.4	142.0±23.5	7.49±0.33	42.8
5'GCGUUUCGUGCGCGCAGCUUACGC5'	77.7±4.6	221.7±14.3	8.93±0.14	46.9	84.5±9.3	243.2±29.0	9.11±0.34	46.8
(5'CGGACAUCCG) ₂	70.7±6.9	203.5±21.7	7.60±0.18	45.7	64.9±11.9	184.8±33.3	7.57±0.40	46.3
(5'GGACAUCC) ₂ ^a	37.1	106.7	4.03	23.8	55.4	167.0	3.92	27.4

^aWu et al. 1995.⁸³ Error limits are omitted because of difficulty fitting data with low melting temperatures.

^bMelting temperatures are at 0.1 mM total oligonucleotide.

Table 2

Thermodynamic Parameters for Internal Loops.

Sequence	G°_{37} (kcal/mol)	H° (kcal/mol)	S° (eu)
$(5'CGGGCAUCCG)_2$	3.58 3.83 ^a	20.5 19.0 ^a	54.5 48.9 ^a
5'GCGUGACUAGCGCGCAUCAGUCGC5'	3.62	-8.8	-39.9
5'GCUGACUACGAACGAUCAGUGC5'	2.19	-15.1	-55.7
5'GCGUGAAUAGCGCGCAUAAGUCGC5'	5.38	25.1	63.4
5'GCGUGCUUUGCGCGCAUUCGACGC5'	5.37	23.2	57.4
5'GCGUGUCUUGCGCGCAUCUGACGC5'	4.13	10.06	19.1
5'GCGUGCCUUGCGCGCAUCCGACGC5'	5.14	30.1	80.5
5'GCGUUUCGUGCGCGCAGCUUACGC5'	4.30	9.35	16.3
$(5'CGGACAUCCG)_2$	3.82	-1.40	-16.8
$(5'GGACAUCC)_2^b$	2.67	10.95	26.7

^aChristiansen & Znosko (2008).⁷² Loop parameters were calculated with Watson-Crick nearest neighbor parameters from Xia et al.⁷⁰ and J. Chen et al.⁶³

^bWu et al. 1995.⁸³ Loop parameters were calculated with Watson-Crick nearest neighbor parameters from Xia et al.⁷⁰ and J. Chen.⁶³

Table 3

Thermodynamic Parameters for Hairpins in 1 M NaCl.^a

sequence	T _M ^a (°C)	Experiment			Loop Formation ^b		
		G° ₃₇	H°	S°	G° ₃₇	H°	S°
		kcal/mol	kcal/mol	eu	kcal/mol	kcal/mol	eu
5'GGCCGA <u>A</u> UGCC	64.9	-2.47	-29.86	-88.31	4.21	-1.59	-18.71
5'CGGGGCUUCCG	49.6	-0.86	-21.99	-68.14	4.76	2.04	-8.74
5'GGACGCUUGUCC ^a	70.9	-4.4 ⁷⁸	-45.0 ⁷⁸	-130.8 ⁷⁸	3.4	-7.77	-36.1
5'GGACUUCCGGUCC ^a	76.4	-6.3 ⁷⁸	-55.9 ⁷⁸	-159.9 ⁷⁸	1.6	-18.67	-65.2
GGCUUCGGCC	72.2	-4.1 ⁷⁹	-40.1 ⁷⁹	-116.1 ⁷⁹	2.6	-11.8	-46.5

^aT_M is calculated from H° and S°.

^bLoop parameters for Antao and Tinoco⁷⁸ were recalculated with Watson-Crick nearest neighbor parameters of Xia et al.⁷⁰

Table 4
Nearest Neighbor Thermodynamics for 2×2 Symmetric Internal Loops at 37 °C in 1 M NaCl^a

G° (kcal/mol at 37°C)									
	GA	AG	UU	GG	CA	CU	UC	CC	AA
GC	-2.5 ^b	-1.4 ^b	-0.6 ^c	-0.8 ^d	1.2 ^c	1.3 ^c	1.3 ^d	1.4 ^d	1.4 ^c
CG	-1.0 ^e	-0.7 ^{e,f}	-0.1 ^g	0.6 ^g	1.4 ^g	1.6 ^g	1.7 ^g	1.9 ^g	1.5 ^{g,e}
UA	0.6 ^b	1.4 ^h	1.0 ^c	2.2 ⁱ	2.3 ^c	2.6 ^c	3.3 ^c	4.9 ^d	2.7 ^c
AU	0.6 ^b	2.1 ^h	0.9 ^c	2.1 ⁱ	3.3 ^{c,j}	3.0 ^d	3.1 ^d	3.8 ^d	3.0 ^c
UG	-0.8 ^b	2.7 ^h	1.9 ^c	1.0 ^d	2.5 ^d	3.5 ^d	4.3 ^j	1.6 ^d	2.0 ^d
GU	0.2 ^b	(1.0) ^{h,k}	1.6 ^d	1.7 ^d	3.7 ^{d,j}	5.4 ^j	4.1 ^j	5.1 ^j	5.4 ^j
H° (kcal/mol)									
	GA	AG	UU	GG	CA	CU	UC	CC	AA
GC	-27.5 ^b	-19.2 ^b	-18.1 ^c	-20.5 ^d	-5.1 ^c	-19.4 ^c	-10.9 ^d	10.8 ^d	-12.3 ^c
CG	-11.6 ^e	-17.0 ^{e,f}	-13.9 ^g	-24.1 ^g	-7.2 ^g	3.0 ^g	0.8 ^g	1.8 ^g	-5.9 ^g
UA	-17.8 ^b	-20.6 ^h	-11.0 ^c	-1.6 ⁱ	7.9 ^c	2.1 ^c	8.3 ^c	10.0 ^d	18.5 ^d
AU	-15.5 ^b	-7.1 ^h	-18.1 ^c	-6.9 ⁱ	4.8 ^{c,j}	1.9 ^d	11.8 ^d	12.5 ^d	11.3 ^c
UG	-20.1 ^b	-21.8 ^h	-0.3 ^c	15.3 ^d	-9.7 ^d	29.4 ^d	9.4 ^j	10.8 ^d	7.1 ^d
GU	-14.2 ^b	(-10.4) ^{h,k}	10.5 ^d	4.9 ^d	19.8 ^{d,j}	23.2 ^j	10.1 ^j	30.1 ^j	-8.8 ^j
S° (cal/(K°mol))									
	GA	AG	UU	GG	CA	CU	UC	CC	AA
GC	-80.6 ^b	-57.4 ^b	-56.3 ^c	-63.6 ^d	-20.2 ^c	-66.9 ^c	-39.4 ^d	30.1 ^d	-42.2 ^c
CG	-34.0 ^e	-52.7 ^{e,f}	-44.3 ^g	-79.7 ^g	-27.5 ^g	4.6 ^g	-2.7 ^g	-0.3 ^g	-24.9 ^g
UA	-59.2 ^b	-70.9 ^h	-38.6 ^c	-12.3 ⁱ	17.8 ^c	-1.8 ^c	16.1 ^c	16.4 ^d	48.2 ^d
AU	-52.0 ^b	-29.4 ^h	-61.3 ^c	-29.0 ⁱ	4.9 ^{c,j}	-3.6 ^d	28.0 ^d	28.0 ^d	27.2 ^c
UG	-62.1 ^b	-78.8 ^h	-7.1 ^c	46.1 ^d	-39.3 ^d	83.5 ^d	16.3 ^j	29.6 ^d	16.7 ^d
GU	-46.4 ^b	(-36.8) ^{h,k}	28.5 ^d	10.1 ^d	51.7 ^{d,j}	57.4 ^j	19.1 ^j	80.5 ^j	-39.9 ^j

^aLeft most column lists base pairs closing loop and top row lists non-canonical pairs in loop. For example, bottom loop in right most column is 5'GAAU/3'UAAG (G°37=5.4 kcal/mol). Bold values are averages from measurements on more than one duplex.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

^bWalter et al. (1994),⁵⁶

^cWu et al. (1995),⁸³

^dChristiansen & Znosko (2008),⁷²

^eSantaLucia et al. (1990),⁵⁷

^fXia et al. (1997),⁵⁸

^gSantaLucia et al. (1991),⁶⁹

^hSchroeder & Turner (2001),⁵³

ⁱBurkard et al. (2001),⁹⁴

^jThis work.

^kUnpublished NMR shows that the sequence melted⁵³ forms a duplex rather than a hairpin. Although there may be multiple conformations for the loop,^{43, 44, 55} the values given assume a single conformation closed by wobble GU pairs.