Stability of RNA Hairpins Closed by Wobble Base Pairs†

Matthew R. Giese, Kelly Betschart, Taraka Dale, Cheryl K. Riley, Carrie Rowan, Kimberly J. Sprouse, and Martin J. Serra*

Department of Chemistry, Allegheny College, Meadville, Pennsylvania 16335

Received August 19, 1997; Revised Manuscript Received November 18, 1997

ABSTRACT: Thermodynamic parameters are reported for hairpin formation in 1 M NaCl by RNA sequences of the type GGXAN$_m$AYCC, where XY is the wobble base pair, GU or UG, and the underlined loop sequences are three to eight nucleotides. A nearest-neighbor analysis indicates the free energy of loop formation is dependent upon loop size and closing base pair. Hairpin loops closed by UG base pairs are on average 1.3 kcal/mol less stable than hairpins closed by GU base pairs. The hairpin loops closed by UG have approximately the same stability as hairpin loops closed by AU/UA base pairs, while the loops closed by GU are approximately 0.7 kcal/mol more stable than hairpins loops closed by GC/CG base pairs. These results, combined with the model previously developed [Serra et al. (1997) *Biochemistry 36*, 4844] to predict the stability for hairpin loops closed by Watson−Crick base pairs, allow for the following model to predict the stability of hairpin loops: $\Delta G^\circ_{37L}(n) = \Delta G^\circ_{37iL}(n) + \Delta G^\circ_{37mm} + 0.6$ (if closed by AU, UA, or UB) − 0.7 (if closed by GU) − 0.7 (if first mismatch is GA or UU except for loops closed by GU). Here, $\Delta G^\circ_{37iL}(n)$ is the free energy increment for initiating a loop of n nucleotides with a CG or GC pair, and ΔG°_{37mm} is the free energy for the interaction of the first mismatch with the closing base pair. For hairpin loops of $n = 4-9$, $\Delta G^0_{37iL}(n)$ is 4.9, 5.0, 5.0, 5.0, 4.9, and 5.5 kcal/mol, respectively. For hairpin loops of $n = 3$, $\Delta G^\circ_{37L}(3) = +4.8 + 0.6$ (if closed by AU, UA, or UG) kcal/mol. Thermodynamic parameters for hairpin formation in 1 M NaCl for 13 naturally occurring RNA hairpin sequences closed by wobble base pairs are reported. The model provides good agreement for both T_M and ΔG°_{37} for most hairpins studied. Thermodynamic values for five terminal mismatches adjacent to wobble base pairs are also reported.

The rules governing the folding of biological molecules into their active three-dimensional structure based upon primary sequence are beginning to be explored. Secondary structure prediction is the first step in determining the three-dimensional structure of a biological molecule. RNA is an excellent candidate to study and gain an understanding of the molecular forces responsible for this folding. RNA has a limited molecular vocabulary, the four nucleotides, and a small number of secondary structural motifs. Several methods for RNA structure determination are currently used. Phylogenetic analysis (*1*, *2*) and chemical and enzymatic probes (*3*, *4*) are used to map RNA secondary structural features. X-ray diffraction and NMR spectroscopy have been used to determine the structures of tRNA, ribozymes, and a number of motifs found in larger RNAs (*5*−*8*).

To complement these structural studies, a number of methods are being developed to model RNA structure (*9*, *10*). Several methods for predicting RNA secondary structure based upon free energy minimization techniques are in use (*11*−*13*). These methods can predict many of the structural features of RNAs, but none can yet predict secondary structure with better than about 70% reliability (*13*, *14*). Further refinements in these techniques require better understanding of the role of sequence on the stability of different RNA structural motifs.

Hairpins are an important RNA structural motif. Nearly 70% of *Escherichia coli* ribosomal RNA nucleotides are found in hairpin structures. We have previously presented a model to predict RNA hairpin stability for hairpins closed by Watson−Crick base pairs (*15*). However, RNA hairpins are often closed by wobble GU or UG base pairs. For example, nearly 20% (14 of 80) of *Escherichia coli* rRNA hairpins are closed with wobble base pairs. We report here the thermodynamic parameters for hairpin formation with loops closed by GU or UG base pairs and a model to predict their stability.

MATERIALS AND METHODS

RNA Synthesis and Purification. Oligomers were synthesized on solid support using the phosphoramidite approach (*16*, *17*). After ammonia and fluoride deprotection, the crude oligomer was purified by preparative TLC (*n*-propanol/ammonium hydroxide/water 55:35:10) and Sep-Pak C18 (Waters) chromatography. Purities were checked by either analytical C8 HPLC chromatography or analytical TLC and were greater than 95%.

Melting Curves and Data Analysis. The buffer for the melting studies was 1.0 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7. Single-stranded extinction coefficients were calculated from the extinction coefficients

† This work is supported by an award of Research Corporation, NIH Grant GM-49429, and Merck/AAAS Undergraduate Science Research Program.

* Author to whom correspondence should be addressed.

S0006-2960(97)02050-3 CCC: $15.00
Published on Web 01/13/1998

for dinucleotide monophosphate and nucleosides (*18, 19*). Strand concentrations were determined from high-temperature absorbance at 280 nm. Absorbance versus temperature curves were measured at 280 nm with a heating or cooling rate of 1.0 °C min^{-1}, on a Perkin-Elmer Lambda 2S or Beckman DU 640 spectrophotometer as described previously (*20*). Oligomer concentrations were varied over at least a 40-fold range between 1 mM and 10 μM.

Absorbance versus temperature profiles were fit to a two-state model with sloping base lines by using a nonlinear least-squares program (*21*), adapted for unimolecular hairpin transitions. Thermodynamic parameters for hairpin formation were obtained from the averages of the fits of the individual melting curves. Melting temperatures and thermodynamic parameters were concentration-independent as expected for hairpin formation.

RESULTS

The set of RNA hairpins of loop size three to eight (excluding six) with wobble base pair closures was prepared, and the thermodynamics of hairpin formation was measured by optical melting. The first mismatch, AA, was chosen to facilitate comparison with previously studied hairpins (*15, 20, 22*). Typical melting curves are shown in the supporting information (see paragraph at end of paper regarding supporting information), and the measured thermodynamics for hairpin formation in 1 M NaCl are listed in Table 1. The free energy for folding, ΔG°_{37}, of the RNA hairpins varies with loop size and closing base pair, GU or UG.

Four hairpins in the original series (loop size 3, 4, and 5 closed by GU and loop size 5 closed by UG) did not melt in a two-state unimolecular manner but rather had concentration-dependent melting temperatures. Therefore, four additional hairpins were included in this study. The GU hairpin with loop of three has an AAA loop, the GU loop of four has the stem

GCG
• • °
CGU

and the loops of five have the sequence AUUUA.

The stability of an RNA hairpin can be dissected into its two structural motifs, the double-helical stem and the loop. The thermodynamic parameters for loop formation (ΔH°_L, ΔS°_L, and ΔG°_{37L}) can be calculated by subtracting the double-helical stem contribution from the hairpin value. For example, the free energy for loop formation at 37 °C, ΔG°_{37L}, can be calculated from $\Delta G^\circ_{37L} = \Delta G^\circ_{37}$ (measured for hairpin formation) $-\Delta G^\circ_{37}$ (stem). The results of this analysis are presented in Table 1.

Hairpin loops of three are modeled as independent of loop sequence with no interaction between the closing base pair and first mismatch (*15, 23*). Therefore, ΔG°_{37iL} is equal to ΔG°_{37L} for hairpin loops of three, where ΔG°_{37iL} is the free energy for initiating a hairpin loop. As observed previously for hairpin loops of six (*22*), the hairpin loop of three closed by UG ($\Delta G^\circ_{37L} = 5.9$ kcal/mol) is less stable than the loop closed by GU ($\Delta G^\circ_{37L} = 4.6$ kcal/mol).

Previous studies (*15, 20*) have shown that the stability of hairpin loops larger than three depends on the stacking interactions, ΔG°_{37mm}, of the first mismatch with the closing base pair. The ΔG°_{37mm} is derived from measurements of model duplexes with terminal mismatches. Terminal mismatches of GU base pairs have not previously been measured. To determine the ΔG°_{37mm} for GU base pairs with several terminal mismatches, a set of duplexes, listed in Table 2, was prepared and the thermodynamics of duplex formation was measured by optical melting. The plots of reciprocal melting temperature (T_M^{-1}) vs log C_t are shown in the supporting information. Enthalpy and entropy changes obtained from these plots are presented in Table 2. The parameters derived from the averages of the individual melting curves are also presented in Table 2. There is good agreement between the two sets of values consistent with two-state melting behavior. The thermodynamic parameters for the terminal mismatch on a GU/UG base pair can be calculated from the thermodynamics of helix formation of the duplex containing the terminal mismatch and the thermodynamics of helix formation of the core. For example, $\Delta\Delta G^\circ$ for the terminal AA mismatch on a GU base pair is $(^1/_2)[\Delta G^\circ(\text{AUCCGGGA}) - \Delta G^\circ(\text{UCCGGG})]$. The results of this analysis are listed in Table 3.

The free energy for initiating a hairpin loop of n nucleotides ($n > 3$) can be calculated from the free energy of loop formation, $\Delta G^\circ_{3iL}(n)$ and the free energy for the stacking interactions of the first mismatch on the closing base pair, ΔG°_{37mm}: $\Delta G^\circ_{37L}(n) = \Delta G^\circ_{37iL}(n) + \Delta G^\circ_{37mm}$. The measured ΔG°_{37iL} for hairpin loops larger than three can be determined by subtracting ΔG°_{37mm} from $\Delta G^\circ_{37L.}$ The results of this analysis are presented in Table 1.

The ΔG°_{37iL} for hairpin loops of size 4−8 closed by GU base pairs are 4.6, 3.7, 3.2, 4.3, and 2.9 kcal/mol, respectively. For hairpin loops of size 4−8 closed by UG base pairs, the ΔG°_{37iL} values are 6.0, 5.2, 5.3, 6.0, and 5.7 kcal/mol, respectively. The hairpin loops closed by GU are on average 1.6 kcal/mol more stable that the hairpin loops closed by UG base pairs. To test the generality of the conclusions from this work, thermodynamic parameters were measured for hairpin loops that either occur naturally in small and large subunit rRNA or represent random sequences. Some of the hairpin loops were chosen with "unusually stable" GA and UU first mismatches (*15, 20*). These results are listed in Table 4.

DISCUSSION

We have previously developed a simple model to predict the stability of RNA hairpin loops (*15*). The free energy increment of hairpin loops of $n > 3$ closed by Watson−Crick base pairs was given by

$$\Delta G^\circ_{37L}(n) = \Delta G^\circ_{37iL}(n) + \Delta G^\circ_{37mm} + 0.6 \text{ (if closed by AU or UA)} \quad -0.7 \text{ (if first mismatch is GA or UU)} \quad (1)$$

$\Delta G^\circ_{37iL}(n)$, the free energy for initiating a loop of n (4−9) nucleotides is 4.9, 5.0, 5.0, 5.0, 4.9, and 5.5 kcal/mol, respectively. ΔG°_{37mm} is the free energy for the stacking interactions of the first mismatch with the closing base pair, +0.6 represents the penalty in free energy for closing a hairpin with an AU (or UA) instead of a GC (or CG) pair, and −0.7 is an additional stabilization for hairpins with GA or UU first mismatches.

Table 1: Thermodynamic Parameters for Hairpin Formation in 1.0 M NaCl

RNA Hairpin	T_M °C (pred)	ΔH° kcal/mol (pred)	ΔS° eu (pred)	$\Delta G°_{37}$ kcal/mol (pred)	$\Delta H°_L$[a] kcal/mol	$\Delta S°_L$[a] eu	$\Delta G°_{37L}$[a] kcal/mol (pred)	$\Delta G°_{37iL}$[b] kcal/mol
GGG A / CCU A	31.8 (30.1)	-18.4±2.9 (-18.5)	-60.4±9.2 (61.0)	+0.3±0.2 (+0.5)	+0.1	-14.9	+4.6 (+4.8)	+4.6
GGU A / CCG A	18.2 (32.2)	-14.2±1.9 (-21.1)	-48.6±6.2 (-69.1)	+0.9±0.2 (+0.4)	+7.5	+5.0	+5.9 (+5.4)	+5.9
GCG A U / CGU A U	41.6 (46.7)	-20.6±2.3 (-21.4)	-65.4±7.6 (-66.9)	-0.3±0.1 (-0.7)	-3.3	-24.3	+4.3 (+3.9)	+4.6
GGU A U / CCG A U	34.7 (42.4)	-20.3±3.1 (-24.8)	-66.0±10.4 (-78.6)	+0.2±0.2 (-0.3)	+1.4	-12.4	+5.2 (+4.7)	+6.0
GGG A U U / CCU A U	42.6 (42.5)	-35.5±2.5 (-22.6)	-112.2±3.4 (-71.6)	-0.6±0.2 (-0.3)	-17.0	-66.7	+3.7 (+4.0)	+4.0
GGU A U U / CCG A U	43.5 (41.2)	26.5±0.5 (-24.8)	83.6±2.2 (-78.9)	-0.6±0.3 (-0.2)	-4.8	-30.0	+4.4 (+4.8)	+5.2
GGG A U A[c] / CCU A U A	50.8 (42.5)	-34.3±2.7 (-22.6)	-106.0±8.2 (-71.6)	-1.4±0.2 (-0.3)	-15.8	-60.5	+2.9 (+4.0)	+3.2
GGU A U A[c] / CCG A U A	42.8 (41.2)	-27.3±0.1 (-24.8)	-86.5±0.5 (-78.9)	-0.5±0.1 (-0.2)	-5.6	-32.5	+4.5 (+4.8)	+5.3
AGGU A U A[c] / UCCG A U A	56.6 (57.1)	-35.4±0.9 (-32.4)	-107.0±3.0 (-98.1)	-2.2±0.1 (-1.9)	-6.1	-34.2	+4.5 (-4.8)	+5.3
GGG A U A / CCU A U A	39.0 (42.5)	-28.2±1.5 (-22.6)	-90.2±4.8 (-71.6)	-0.3±0.1 (-0.3)	-9.7	-44.7	+4.0 (+4.0)	+4.3
GGU A U A / CCG A U A	35.3 (41.2)	-30.3±2.4 (-24.8)	-99.5±7.5 (-78.9)	+0.2±0.3 (-0.2)	-8.6	-45.9	+5.2 (+4.8)	+6.0
GGG A U A A / CCU A U A A	53.1 (43.8)	-32.6±3.5 (-22.6)	-99.8±9.4 (-71.3)	-1.7±0.3 (-0.4)	-14.1	-54.3	+2.6 (+3.9)	+2.9
GGG A U A A / CCU A U A A	35.3 (42.4)	-29.5±2.4 (-24.8)	-95.2±8.7 (-78.6)	-0.1±0.1 (-0.3)	-7.8	-41.6	+4.9 (+4.7)	+5.7

[a] Calculated from equations equivalent to $\Delta G°_{37L} = \Delta G°_{37}$ (hairpin formation) − $\Delta G°_{37}$ (pred. stem). [b] Calculated from $\Delta G°_{37L} = \Delta G°_{37iL} + \Delta G°_{37mm}$, where $\Delta G°_{37mm}$ is derived from measurements of model duplexes with terminal mismatches (*23*) except for loop of three where $\Delta G°_{37iL} = \Delta G°_{37L}$. [c] Reference 22.

Only five RNA hairpins closed by wobble base pairs have been previously studied, all with loop size of six (*20*, *22*). These studies found that hairpin loops closed by GU were more stable when the first mismatch was AA, while loops with UG base pair closure were more stable with GA (unusually stable) first mismatches. It was not clear whether the differences in stability of these hairpin loops was due to differences in the stability of the hairpin base closures, the interactions of the closing base pairs with the first mismatches, or both. To determine the interaction of the closing wobble base pair with the first mismatches, duplexes with wobble base pairs adjacent to AA and GA terminal mismatches, were studied. The stability of the AA and GA terminal mismatches on the GU/UG base pair are approximately 0.5 kcal/mol less stable than previously predicted (*24*). Their stability closely approximates that of the 3′-terminal dangling nucleotide as observed for terminal mismatches on Watson−Crick base pairs (*25*).

Table 2: Thermodynamic Parameters of Duplex Formation[a]

	T_M^{-1} vs log C_t plots				average of curve fits			
oligomers	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal/mol)	T_M[b] (°C)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal/mol)	T_M[b] (°C)
AGCGUA	34.3	97.8	4.0	22.7	32.2	90.2	4.2	21.8
AUCCGGGA	54.4	149.8	7.9	50.8	51.7	141.4	7.8	49.4
AGGCGCUG	52.8	140.0	9.4	60.9	57.6	154.3	9.8	56.3
AUGGCCGG	54.2	143.4	9.7	62.2	54.5	144.1	9.8	59.6
GGCCGGUG	61.1	162.2	10.8	65.6	60.1	159.1	10.8	61.7
	reference duplexes							
GCGU[c]	27.0	79.4	2.4					
UCCGGG[d]	47.7	129.8	7.4					
GGCGCU[d]	56.4	154.7	8.4					
UGGCCG[d]	53.0	143.3	8.5					
GCCGGU[d]	58.2	158.1	9.2					

[a] Solutions are 1 M NaCl, 10 mM sodium cacodylate, and 0.5 mM EDTA, pH 7. [b] Calculated at 10^{-4} M oligomer concentration. [c] Predicted (*23*). [d] Reference *28*.

Table 3: Thermodynamic Parameters for Terminal Mismatches (1 M NaCl)[a]

	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal/mol)
UA • GA	3.7	9.2	0.8
UG • GA	−3.1	−11.2	0.5
UG • GG	1.5	2.1	0.8
GA • UA	3.4	10	0.3
GG • UA	0.6	0	0.6

[a] Thermodynamic parameters are calculated as illustrated in text.

The free energy measurements listed in Tables 1 and 4, when combined with previous measurements (*20, 22*), allow more reliable values for predicting loop stability for hairpins closed by wobble base pairs. Table 5 lists the average $\Delta G°_{37iL}(n)$ for hairpin loops closed by GU or UG base pairs. These results allow the development of a simple model for predicting the thermodynamic values for hairpin loops ($n > 3$) in kilocalories per mole:

$$\Delta G°_{37L}(n) = \Delta G°_{37iL}(n) + \Delta G°_{37mm} + 0.6 \text{ (if closed by AU, UA, or UG)} + [-0.7 \text{ (if closed by GU)}] + [-0.7 \text{ (if first mismatch is GA or UU except for hairpins closed by GU)}] \quad (2)$$

$$\Delta H°_{L}(n) = \Delta H°_{mm} + 0.6 \text{ (if closed by AU, UA, or UG)} + [-0.7 \text{ (if closed by GU)}] + [-0.7 \text{ (if first mismatch is GA or UU except for hairpins closed by GU)}] \quad (3)$$

$$\Delta S°_{L}(n) = \Delta S°_{iL}(n) + \Delta S°_{mm} \quad (4)$$

This model assumes that $\Delta H°_{iL} = 0$ and therefore $\Delta S°_{iL} = -\Delta G°_{37iL}/310.15$.

Hairpin loops of three closed by GU have the same stability as hairpin loops closed by GC/CG base pairs while loops closed with UG have similar stability to loops closed with AU/UA base pairs (Table 5). The model to predict the stability for hairpin loops of three (in kilocalories per mole) is "loop" sequence-independent:

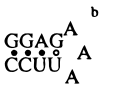
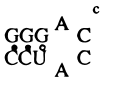
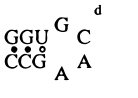
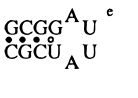
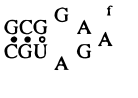
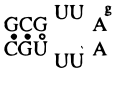
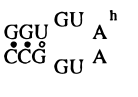
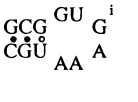
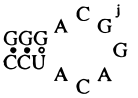
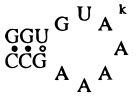
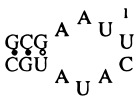
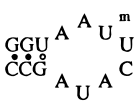
$$\Delta G°_{37L}(3) = +4.8 + 0.6 \text{ (if closed by AU, UA, or UG)} \quad (5)$$

The model was used to predict the thermodynamics for the hairpins in Tables 1 and 4. In most cases, there is good agreement between the measured and predicted T_M and $\Delta G°_{37}$. The stability of hairpin loops closed by wobble base pairs is less predictable than hairpins closed with Watson−Crick base pairs. This is most evident with the loops of four closed by GU where $\Delta G°_{37iL}$ varies between +2.4 (ACCA) and +5.4 (AUUA). These results are surprising since the AUUA loop would be able to fold, forming a "U-turn" (*26*). Two other hairpins closed by GU are also not predicted well, AUAAUA $\Delta G°_{37L}$ (measured +2.9 kcal/mol; predicted +4.0 kcal/mol) and AUAAAAUA $\Delta G°_{37L}$ (measured +2.6 kcal/mol; predicted +3.9 kcal/mol) (Table 1). Hairpin loops of eight closed by AU base pairs were found to be the most variable in our study of hairpins closed by Watson−Crick base pairs (*15*). This variability may be due to the structural flexibility of large loops.

All of the hairpin loops closed by UG base pairs in this study are predicted very well. Only one hairpin, studied previously (*20*), is greater than 1 kcal/mol more stable than predicted. The hairpin loop GUAAUA $\Delta G°_{37L}$ (measured +3.1 kcal/mol; predicted +4.4 kcal/mol) represents an "unusually stable" first mismatch.

Four hairpins closed by GU and having an "unusually stable" (GA or UU) first mismatch have been studied, three in Table 4 and one previously (*20*). In all cases, inclusion of the additional stabilization term for the GA or UU first mismatch produces poorer predictions of the actual hairpin stability. Therefore, for hairpins closed by GU, the additional (GA, UU) stabilization is not included in the predictions. Four hairpins closed by UG and having GA or UU as the first mismatch have also been studied, three in Table 4 and one previously (*20*). In all cases, the inclusion of the additional stabilization term predicts the actual stability better.

Table 4: Thermodynamic Parameters for Hairpin Formation of Natural Sequences in 1.0 M NaCl

RNA Hairpin	T_M °C Measured (Predicted)	ΔH° kcal/mol Measured (Predicted)	ΔS° eu Measured (Predicted)	$\Delta G°_{37}$ kcal/mol Measured (Predicted)	$\Delta G°_{37L}$[a] kcal/mol Measured (Predicted)
	46.1 (48.4)	-31.2±3.6 (-28.2)	-97.6±11.9 (-87.7)	-0.9±0.4 (-0.9)	+4.8 (+4.8)
	53.7 (43.8)	-41.7±3.5 (-22.6)	-128.6±11.5 (-71.3)	-1.9±0.3 (-0.4)	+2.4 (+3.9)
	52.5 (48.2)	-30.0±2.4 (-18.7)	-92.1±7.6 (-58.2)	-1.4±0.2 (-0.7)	+3.6 (+4.3)
	51.8 (66.8)	-14.2±1.9 (-32.6)	-93.8±10.7 (-95.9)	-1.4±0.2 (-2.9)	+5.4 (+3.9)
	41 (52.0)	-21.4±4.4 (-18.6)	-68.1±14.0 (-57.2)	-0.3±0.3 (-0.9)	+4.3 (+3.7)
	33.5 (55.7)	-29.8±2.6 (-19.7)	-97.2±8.5 (-59.9)	-0.3±0.1 (-1.1)	+4.3 (+3.5)
	48.9 (46.4)	-35.1±1.8 (-30.2)	-109.0±5.0 (-94.5)	-1.3±0.1 (-0.8)	+3.7 (+4.2)
	41.1 (52.0)	-34.5±3.2 (-18.6)	-109.0±10.0 (-57.2)	-0.4±0.3 (-0.9)	+4.2 (+3.7)
	45.0 (42.5)	-39.8±3.5 (-22.6)	-125.0±11.2 (-71.6)	-1.0±0.2 (-0.3)	+3.3 (+4.0)
	43.5 (46.5)	-31.3±3.5 (-18.7)	-98.8±10.6 (-58.5)	-0.6±0.2 (-0.6)	+4.4 (+4.4)
	42.9 (46.7)	-32.6±3.8 (-21.4)	-103.0±10.4 (-66.9)	-0.6±0.2 (-0.7)	+4.0 (+3.9)
	37.6 (42.4)	-26.6±1.7 (-24.8)	-85.5±5.9 (-78.6)	-0.0±0.2 (-0.3)	+5.0 (+4.7)
	50.2 (48.7)	-32.7±2.7 (-22.4)	-101.3±8.8 (-69.6)	-1.3±0.2 (-1.0)	+3.7 (+4.3)

[a] Calculated from $\Delta G°_{37iL} = \Delta G°_{37}$(hairpin formation) − $\Delta G°_{37}$(pred. stem). Sequences were modeled on the following (numbers refer to *E. coli* equivalent positions): [b]*Thermoplasma acidophilum* large subunit rRNA 780; [c]*Pseudomonas aeroginosa* large subunit rRNA 410; [d]*Halobacterium marismortui* (HC10) small subuint rRNA 895; [e]Random sequence, with loop similar to model sequences; [f]*Chlorella ellipsoidea* large subunit rRNA 2028; [g]random sequence with unusually stable first mismatch; [h]*Borrelia burgdorferi* small subunit rRNA 1130; [i]*E. coli* small subunit rRNA position 690; [j]*Oryza sativa* large subunit rRNA 2306; [k]*Methanobacterium formicicum* small subunit rRNA 450; [l]*Chlamydomonas eugameto* large subunit rRNA 158; [m]*Chlamydomonas reinhardtii* large subunit rRNA 158; and [n]random sequence with unusually stable GA first mismatch (*30*, *31*).

The large difference in average stability (1.3 kcal/mol) between hairpin loops closed by GU and UG is surprising, since hairpin loops closed by GC and AU pairs have similar stability when the base pair is reversed. Several lines of

Table 5: ΔG°_{37iL} for Hairpin Loop Formation in 1 M NaCl

loop size	ΔG°_{37iL} (GU)[a]	ΔG°_{37iL} (GC or CG)	ΔG°_{37iL}(GC) − ΔG°_{37iL}(GU)
3[b]	4.7 (2)	4.8	−0.1
4	4.1 (3)	4.9	−0.8
5	4.3 (3)	5.0	−0.8
6	4.6 (5)	5.0	−0.4
7	4.0 (3)	5.0	−1.0
8	3.5 (2)	4.9	−1.4
avg			−0.7
loop size	ΔG^0_{37iL} (UG)[a]	ΔG°_{37iL} (AU or UA)	ΔG°_{37iL} (AU) − ΔG°_{37iL} UG
3[b]	5.9 (1)	5.7	+0.2
4	5.4 (2)	5.4	0.0
5	5.2 (1)	5.8	−0.6
6	5.1 (5)	5.3	−0.2
7	5.8 (2)	5.8	0.0
8	5.5 (3)	5.4	+0.1
avg			−0.1

[a] Average calculated ΔG°_{37iL} for measured hairpin loops (*15*, *22*). Number in parentheses represents number of hairpins used in average. [b] For hairpin loops of three, ΔG°_{37iL} is loop sequence-independent.

nonthermodynamic evidence suggested that hairpin loops closed by UG might in fact be more stable than loops closed by GU.

In tRNAs, the ends of helices closed by UG are found in a "stacked" configuration while GU pairs are unstacked (*27*). Helices ending in $\frac{XU}{YG}$ are thermodynamically more stable than the corresponding helix ending in $\frac{XG}{YU}$ (*28*). In small subunit rRNA, there is only a slight preference for the more stable stacked wobble pair at the end of helical regions (*29*). The presence of conserved wobble pairs in both rRNA and tRNAs suggest that they may play structural/functional roles.

In *Escherichia coli* rRNA, 10 of the 14 hairpin loops closed by wobble base pairs are closed with UG. Nearly twice as many hairpin loops in phylogenetic rRNA structures are closed with UG (10%) as with GU (5%) (*2*, *30*, *31*). It is interesting that small loop sizes have a preponderance of UG base closures while larger loops tend to be closed more frequently by GU. This trend is not related to the thermodynamic stability of the loop and may be related to either the structural flexibility of the unstacked UG base pair or tertiary interactions of the RNA.

Nearly 40% of hairpins, with loop sizes between three and nine in small and large rRNA, have unusually stable first mismatches (*2*, *30*, *31*). Almost half (198 of 428) of the hairpins closed by UG have unusually stable first mismatches. GA and UU first mismatches almost never occur (less than 5%) in hairpins closed by GU. There are two notable exceptions where highly conserved hairpin loops occur. These are at position 690 (*G*GUGAAA*U*) of the small ribosomal RNA and at position 1950 (*G*UAAGU*U*) of the large ribosomal RNA. A hairpin modeled on the sequence *G*GUGAAA*U* was included in this study (Table 4). The conserved nature of these two hairpin loop sequences in the absence of additional stabilization by the first mismatches suggests that these loops have a functional role in the ribosome.

While GA mismatches can hydrogen-bond in a variety of geometries, only two conformations have been observed in RNA. The guanine can hydrogen-bond using its imino proton as observed in tRNA (*26*, *32*) or can hydrogen-bond in a "sheared" base pair conformation as seen at loop−helix junctions (*7*, *33*, *34*). Structural investigations of hairpin loops of four and six closed by CG base pairs with GA first mismatches, where the GA provides additional stabilization, have shown the GA mismatch to adopt the "sheared" conformation (*7*, *35*, *36*). The additional stabilization of these hairpin loops is achieved by the stacking interactions of the first mismatch (GA) on the closing base pair and a network of hydrogen bonds involving the first mismatch and the rest of the loop (*7*, *35*−*37*). The absence of additional stabilization by GA or UU first mismatches on hairpins closed by GU base pairs suggest that these hairpin loops may adopt a different structure. For example, the different stacking interactions with a GU pair may distort a sheared GA conformation. The structural plasticity of GA mismatches in RNA has been observed in internal loops where these bases have been shown to adopt different conformations depending upon the closing base pair (*38*). The structural variability leads to marked differences in the stability of the internal loops. The lack of additional stabilization of the two conserved hairpin loops described above suggests that they may have been selected for their unique structure and ability to form tertiary interactions. Structural investigations of these two hairpin loops and elucidation of their tertiary interactions will be necessary to fully comprehend the role of these hairpins in ribosome structure and function.

The results presented here provide a simple model for predicting the stability of hairpins closed by wobble base pairs. These results should improve the prediction of RNA secondary structure from sequence.

ACKNOWLEDGMENT

We thank Dr. Douglas H. Turner for many stimulating discussions and careful reading of this manuscript.

SUPPORTING INFORMATION AVAILABLE

Figure 1, showing representative melting curves for the hairpins presented in Table 1, and Figure 2, showing $1/T_M$ vs log (C_t) for the terminal mismatches presented in Table 2 (3 pages). Ordering information is given on any current masthead page.

REFERENCES

1. Woese, C. R., and Pace, N. R. (1993) in *The RNA World* (Gesteland, R. F., and Atkins, J. F., Eds.) pp 91−117, Cold Spring Harbor Press, Cold Spring Harbor, NY.
2. Gutell, R. R., Power, A., Hertz, G. Z., Putz, E. J., and Stromo, G. D. (1992) *Nucleic Acids Res. 20*, 5785−5795.
3. Parker, R. (1989) *Methods Enzymol. 180*, 510−517.
4. Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J., and Ehresmann, B. (1987) *Nucleic Acids Res. 15*, 9109−9127.
5. Woo, N. H., Roe, B. A., and Rich, A. (1980) *Nature 286*, 346−351.
6. Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994) *Nature 372*, 68−74.
7. Heus, H. A., and Pardi, A. (1991) *Science 253*, 191−194.
8. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1678−1686.
9. Gautheret, D., Major, F., and Cedergren, R. (1993) *J. Mol. Biol. 229*, 1049−1064.
10. Michel, F., and Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.
11. Martinez, H. M. (1990) *Methods Enzymol. 183*, 306−317.

12. Zuker, M. (1989) *Science 244*, 48−52.
13. Gultyaev, A. P., van Batenburg, F. H., and Pleij, C. W. (1995) *J. Mol. Biol. 250*, 37−51.
14. Walter, A. E., Turner, D. H., Kim, J., Little, M. H., Muller, P., Mathews, D. H., and Zuker, M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9218−9222.
15. Serra, M. J., Barnes, T. W., Betschart, K., Gutierrez, M. J., Sprouse, K. J., Riley, C. K., Stewart, L., and Temel, R. E. (1997) *Biochemistry 36*, 4844−4851.
16. Wu, T., Ogilvie, K. K., and Pon, R. T. (1989) *Nucleic Acids Res. 17*, 3501−3517.
17. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. L. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.
18. Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 589, CRC Press, Cleveland, OH.
19. Richards, E. G. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 197, CRC Press, Cleveland, OH.
20. Serra, M. J., Axenson, T. J., and Turner, D. H. (1994) *Biochemistry 33*, 14289−14296.
21. McDowell, J. A., and Turner, D. H. (1996) *Biochemistry 35*, 14077−14089.
22. Serra, M. J., Lyttle, M. H., Axenson, T. J., Schadt, C. A., and Turner, D. H. (1993) *Nucleic Acids Res. 21*, 3845−3849.
23. Serra, M. J., and Turner, D. H. (1995) *Methods Enzymol. 259*, 242−261.
24. Jaeger, J. A., Turner, D. H., and Zuker, M. (1990) *Methods Enzymol. 183*, 281−306.
25. Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., and Turner, D. H. (1983) *Biochemistry 22*, 6198−6206.
26. Quigley, G. J., and Rich, A. (1976) *Science 194*, 796−806.
27. Mizuno, H., and Sundaralingam, M. (1978) *Nucleic Acids Res. 5*, 4451−4461.
28. Freier, S. M., Kierzek, R., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) *Biochemistry 25*, 3209−3213.
29. Gautheret, D. , Konings, D., and Gutell, R. R. (1995) *RNA 1*, 807−814.
30. Gutell, R. R., Schnare, M. N., and Gray, M. W. (1992) *Nucleic Acids Res. 20* (Suppl.), 2095−2109.
31. Gutell, R. R. (1993) *Nucleic Acids Res. 21*, 3051−3054
32. Kim, S. H., Suddath, F. L., Quigley, G. J., McPhereson, A., Sussman, J. L., Wang, A., Seeman, N. C., and Rich, A. (1974) *Science 185*, 435−440.
33. Varani, G., Wimbley, B., and Tinoco, I., Jr. (1989) *Biochemistry 28*, 7760−7772.
34. SantaLucia, J., and Turner, D. H. (1993) *Biochemistry 32*, 12612−12623.
35. Fountain, M. A., Serra, M. J., Krugh, T. R., and Turner, D. H. (1996) *Biochemistry 35*, 6539−6548.
36. Huang, S., Wang, Y.-X., and Draper, D. E. (1996) *J. Mol. Biol. 258*, 308−321.
37. Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H. M., and Pardi, A. (1996) *J. Mol. Biol. 264*, 968−980.
38. Walter, A. E., Wu, M., and Turner, D. H. (1994) *Biochemistry 33*. 11349−11354.

BI972050V