A Model for the Stabilities of RNA Hairpins Based on a Study of the Sequence Dependence of Stability for Hairpins of Six Nucleotides†

Martin J. Serra,*,‡ Theresa J. Axenson,§ and Douglas H. Turner§

Department of Chemistry, Allegheny College, Meadville, Pennsylvania 16335, and Department of Chemistry, University of Rochester, Rochester, New York 14627-0216

Received June 15, 1994; Revised Manuscript Received September 15, 1994⊗

ABSTRACT: Thermodynamic parameters are reported for hairpin formation in 1 M NaCl by RNA sequences of the type GGCXUAAUYGCC, where XY is the set of 10 possible mismatch base pairs. A nearest neighbor analysis of the data indicates the free energy for loop formation at 37 °C varies from 2.9 to 4.5 kcal/mol. Thermodynamic parameters are also reported for hairpin formation by RNA sequences of the type GGXGUAAUAYCC (where XY are CG, GC, AU, UA, GU, and UG), with the common naturally occurring GA first mismatch (45% of small and large subunit rRNA loops of six). These results allow the development of a model to predict the stability of RNA hairpin loops. The model includes the size of the loop, the identity of the closing base pair, the free energy increment ($\Delta G°_{37MM}$) for interaction of the closing base pair with the first mismatch, and an additional stabilization term for GA and UU first mismatches. $\Delta G°_{37L}(n) = \Delta G°_{37i}(n) + \Delta G°_{37MM} + 0.4$ (if closed by AU or UA) -0.7 (if first mismatch is GA or UU). Here $\Delta G°_{37i}(n)$ is the free energy for initiating a loop of n nucleotides. $\Delta G°_{37i}(n)$ for $n = 4-9$ is 4.9, 4.4, 5.0, 5.0, 5.1, and 5.2 kcal/mol, respectively. The $\Delta G°_{37MM}$ is derived from measurements of model duplexes with terminal mismatches. The model gives good agreement when tested against four naturally occurring hairpin sequences.

RNAs are intimately associated with a wide variety of biological functions (Watson et al., 1987). An understanding of the molecular basis of these functions requires knowledge of the three-dimensional structure of RNA. The three-dimensional structures of RNAs have been difficult to determine. The only naturally occurring RNAs or fragments thereof whose three-dimensional structures are known are tRNAs (Kim et al., 1974; Robertus et al., 1974; Westhof et al., 1985), GNRA and UUCG hairpins (Cheong et al., 1990; Heus & Pardi, 1991), and the E loop from 5S rRNA (Wimberly et al., 1993). Thus, methods are being developed to model the structures of RNAs (Michel & Westhof, 1990; Woese & Pace, 1993; Major et al., 1991; Malhotra et al., 1990; Brimacombe et al., 1988; Stern et al., 1988).

An important step in modeling RNA structure is determining the secondary structure. Several approaches are commonly used to predict RNA secondary structure. Phylogenetic studies can reveal areas of similar secondary structure by comparing RNAs with varying sequence but similar function (Noller, 1984; Woese & Pace, 1993; Gutell et al., 1994). Structure mapping has also been used to probe RNA secondary structure (Parker, 1989). Thermodynamic stability can be used to predict the most stable secondary structure (Tinoco et al., 1971; DeLisi & Crothers, 1971; Turner et al., 1988). All three methods can be used cooperatively. The main drawback to thermodynamic structural prediction has been the lack of experimental data for structural motifs other than the double helix (Turner et al., 1987).

Hairpins are an important RNA structural motif and have received increased investigation recently (Groebe & Uhlenbeck, 1988; Antao et al., 1991; Antao & Tinoco, 1992; SantaLucia et al., 1992; Serra et al., 1993). We have previously shown that the closing base pair of an RNA hairpin affects its stability (Serra et al., 1993). Here we show that the first mismatch in the loop is also an important determinant of stability. Combined with previous studies, the results permit the development of a model to predict the stabilities of RNA hairpins.

MATERIALS AND METHODS

RNA Synthesis and Purification. Oligomers were synthesized on solid support using the phosphoramidite approach (Wu et al., 1989; Usman et al., 1987). After ammonia and fluoride deprotection, the crude oligomer was purified by SEP-PAK C18 (Waters) chromatography. The oligomer was further purified by preparative TLC (*n*-propanol:ammonium hydroxide:water, 55:35:10) (Chou et al., 1989) and desalted by SEP-PAK C18 chromatography. Purities were checked by analytical C-8 HPLC (Beckman) or 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, unless indicated otherwise. These conditions provide thermodynamic parameters that are reasonable approximations for solutions containing $10^{-3}-10^{-1}$ M Mg^{2+} in the presence of 0.15 M Na^+ (Williams et al., 1989). Single-strand extinction coefficients were calculated from the extinction coefficients for dinucleoside monophosphates and nucleosides, as described previously

† This work is supported by NIH Grant GM22939 (D.H.T.), NIH Grant GM49429 (M.J.S.), Research Corporation (M.J.S.), and the DANA Foundation. D.H.T. is a Guggenheim Fellow and an American Cancer Society Scholar.

* Author to whom correspondence should be addressed.
‡ Allegheny College.
§ University of Rochester.

⊗ Abstract published in *Advance ACS Abstracts,* November 1, 1994.

0006-2960/94/0433-14289$04.50/0

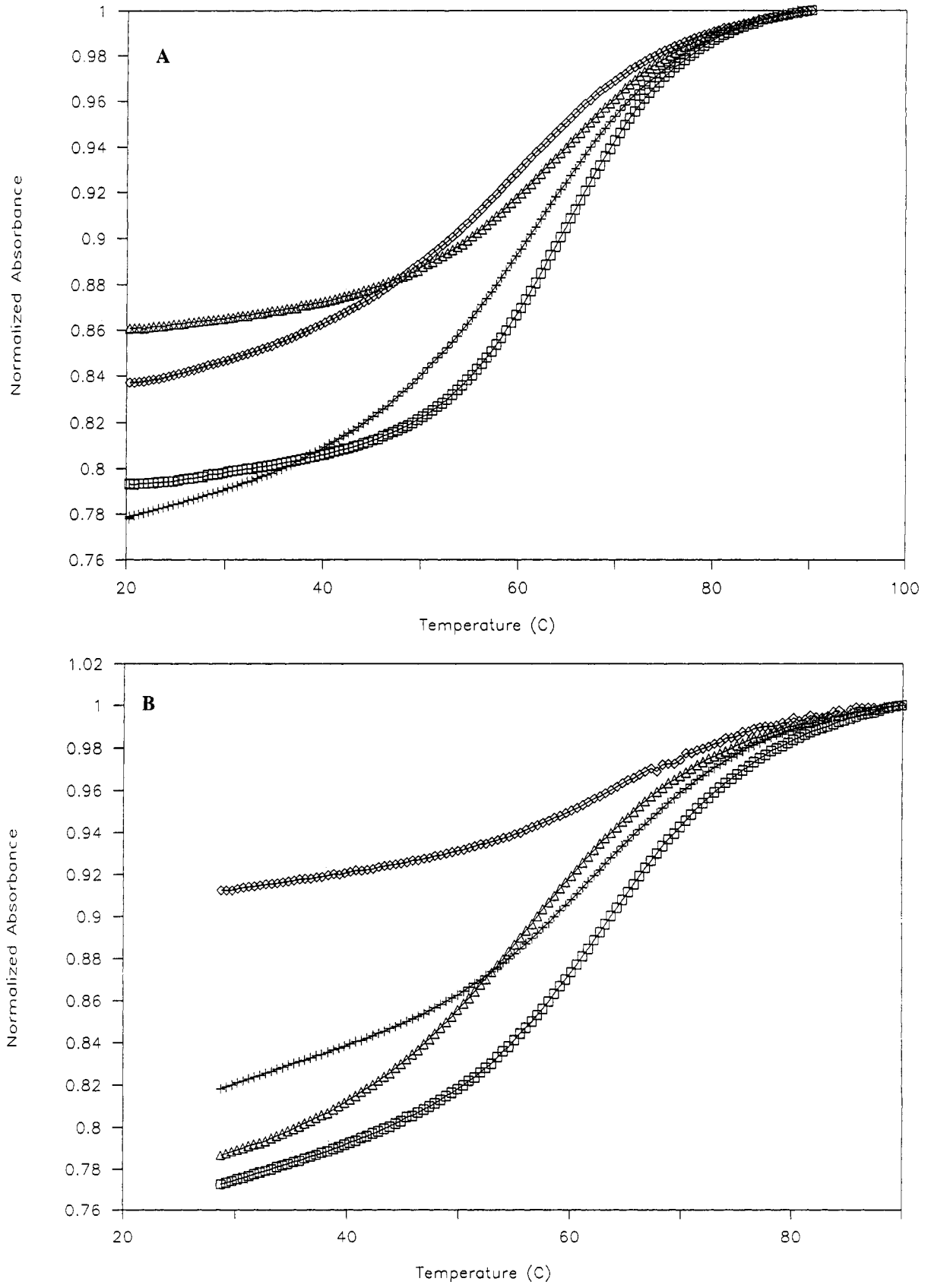

FIGURE 1: Normalized melting curves. Sequences are (A) ΔΔΔΔΔΔΔΔΔΔΔΔΔ, GGCAUAAUCGCC: ◇◇◇◇◇◇◇◇◇◇◇◇, GGCCUAAUAGCC: ++++++++++++, GGCCUAAUCGCC: □□□□□□□□□□□□, GGCUUAAUGCC: and (B) ΔΔΔΔΔΔΔΔΔΔΔ, GGCCUAAUUGCC: ◇◇◇◇◇◇◇◇◇◇◇, GGCGUAAUGGCC: +++++++++, GGCAUAAUGGCC: □□□□□□□□□□, GCUUAAUCGCC. Buffer is 1 M NaCl, 10 mM sodium cacodylate, and 0.5 mM EDTA, pH 7.0.

(Borer, 1975; Richards, 1975). Strand concentrations were determined from high-temperature absorbance at 280 nm. Absorbance versus temperature melting curves were measured at 280 nm with a heating rate of 1.0 °C min^{-1}, on a Gilford 250 or Perkin Elmer Lambda 2S spectrophotometer as described previously (Freier et al., 1983; Petersheim &

Table 1: Thermodynamic Parameters for Hairpin Formation at 1 M NaCl

```
  A A
U   U
X   Y
 C•G
 G•C
 G•C
```

XY	T_M (°C)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (eu)	$\Delta G°_{37}$ (kcal/mol)	$\Delta H°_L$[a] (kcal/mol)	$\Delta S°_L$[a] (eu)	$\Delta G°_{37L}$[a] measd (pred) (kcal/mol)	percent occurrence[b]
GA[c]	68.7	−36.9 ± 2.4	−107.8 ± 7.2	−3.42 ± 0.24	−10.5	43.1	2.9 (2.9)	45
UU	64.9	−37.4 ± 2.3	−100.5 ± 6.6	−3.19 ± 0.08	−11.0	35.8	3.1 (3.1)	8
GG	66.2	−36.2 ± 2.1	−104.3 ± 6.0	−3.12 ± 0.19	−9.8	39.6	3.2 (3.4)	17
AG	65.9	−35.3 ± 2.8	−103.8 ± 8.8	−3.07 ± 0.19	−8.9	39.1	3.3 (3.6)	6
UC	64.9	−34.9 ± 0.1	−103.2 ± 2.9	−2.99 ± 0.03	−8.5	38.5	3.3 (3.6)	4
AA[c]	64.0	−34.3 ± 1.3	−101.7 ± 3.2	−2.74 ± 0.13	−7.9	37.0	3.6 (3.5)	6
AC	65.3	−30.5 ± 1.3	−90.1 ± 3.7	−2.56 ± 0.12	−4.1	25.4	3.8 (3.5)	1
CC	64.4	−31.1 ± 1.1	−92.6 ± 5.2	−2.30 ± 0.07	−4.7	27.9	4.1 (3.9)	4
CA	62.9	−28.2 ± 1.4	−84.0 ± 4.0	−2.17 ± 0.20	−1.8	19.3	4.2 (4.0)	5
CU	58.0	−29.8 ± 1.7	−90.2 ± 4.6	−1.87 ± 0.06	−3.4	25.5	4.5 (4.2)	4

[a] Calculated from equations equivalent to $\Delta G°_{L37}$ = $\Delta G°_{37}$(hairpin formation) − $\Delta G°_{37}$(pred stem). [b] Percent of phylogenetically determined hairpins of six in small and large subunit rRNAs that have the same first mismatch (Gutell et al., 1985, 1993); total number of loops of six is 478. [c] Serra et al. (1993).

Turner, 1983). Oligomer concentrations were varied over at least a 20-fold range.

Absorbance versus temperature profiles were fit to a two-state model with sloping base lines by using a nonlinear least-squares program (Petersheim & Turner, 1983; Freier et al., 1983). For hairpin melts, this program was adapted for a unimolecular transition. Thermodynamic parameters for hairpin formation were obtained from averages of the fits of the individual melting curves. Melting temperatures and thermodynamic parameters were concentration independent as expected for unimolecular hairpin formation. 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 reciprocal melting temperatures, T_m^{-1}, versus log C_t gave enthalpy and entropy changes (Borer et al., 1974):

$$T_m^{-1} = (2.3R/\Delta H°)\log C_t + \Delta S°/\Delta H° \quad (1)$$

Here C_t is the total concentration of oligomers. Parameters derived from the two methods agreed within 10%, consistent with the two-state model.

RESULTS

Phylogenetic studies of rRNA molecules have revealed a marked preference for certain loop sequences. In particular, the first mismatch in loops of four (Woese et al., 1990) and six is frequently GA (approximately 45% for loops of six) (Gutell et al., 1985, 1993). Hairpin loops with a GA as the first mismatch are also about 0.7 kcal/mol more stable than identical loops with an AA mismatch (SantaLucia et al., 1992; Serra et al., 1993). To determine the role of the first mismatch on the stability of hairpin loops, a complete set of hairpins closed with CG pairs but with differing first mismatches was prepared and the thermodynamics of hairpin formation were measured by optical melting.

Table 2: Thermodynamic Parameters for Hairpin Formation at 0.1 M NaCl

XY	T_M (°C)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$\Delta G°_{37}$ (kcal/mol)	$\Delta\Delta G°_{37,0.1-1.0}$[a]
GA	64.9	31.2 ± 1.8	92.4 ± 5.3	−2.70 ± 0.15	0.7
UU	61.0	37.9 ± 1.5	113.5 ± 4.3	−2.72 ± 0.15	0.5
AG	61.4	31.0 ± 0.6	92.3 ± 1.9	−2.24 ± 0.04	0.9
UC	59.5	34.9 ± 0.2	104.9 ± 0.8	−2.34 ± 0.04	0.8
CC	56.7	31.7 ± 2.6	96.0 ± 7.7	−1.90 ± 0.23	0.5
CU	55.1	28.5 ± 1.8	86.7 ± 5.3	−1.57 ± 0.17	0.3

[a] Difference in free energy change at 37 °C measured at 0.1 and 1 M NaCl.

Typical melting curves are shown in Figure 1, and the measured thermodynamic parameters in 1 M NaCl are listed in Table 1. The free energy change for folding, $\Delta G°$, of the RNA hairpins varies from −3.4 to −1.9 kcal/mol at 37 °C. The stability of an RNA hairpin can be dissected into its two structural motifs, the double-helical stem and the loop. The free energy for loop formation at 37 °C, $\Delta G°_{37L}$, can be calculated from $\Delta G°_{37L}$ = $\Delta G°_{37}$(measured for folding of hairpin) − $\Delta G°_{37}$(stem). Since all the hairpins contained a common stem, ggg [$\Delta G°_{37}$ = −6.3 kcal/mol calculated according to Freier et al. (1986a)], the free energy for loop formation varied by an amount equivalent to the difference in hairpin stability. The results of this analysis are presented in Table 1. Also listed in Table 1 are the number of occurrences of the various first mismatches in hairpin loops of six in phylogenetically determined secondary structures

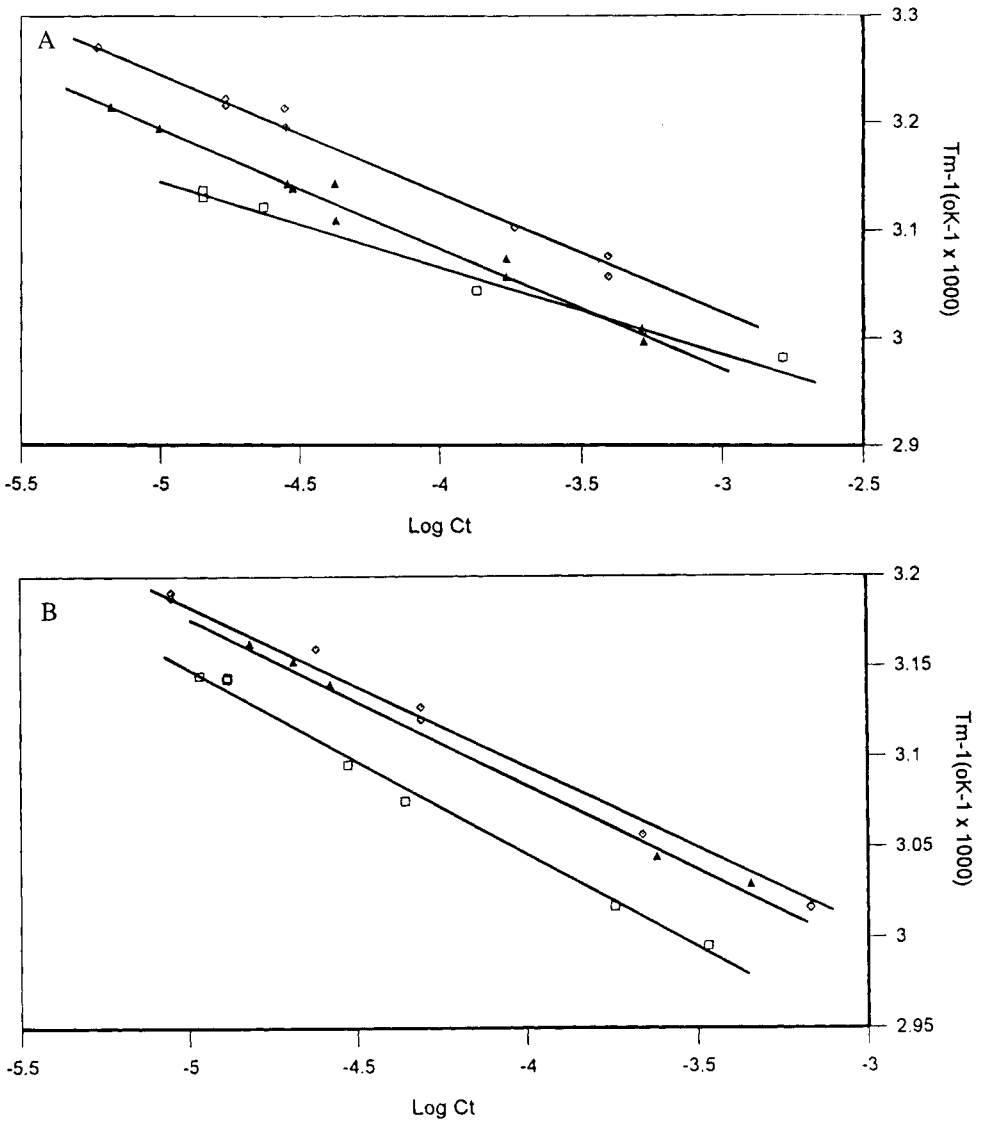

FIGURE 2: Reciprocal melting temperature versus log(concentration) plots. (A) ◇, UGGCCC; ▲, GGCGCA: □, CGGCCU: and (B) ◇, AGCGCA: ▲, GGCGCG: □, CGGCCA in 1 M NaCl, 10 mM sodium cacodylate, and 0.5 mM EDTA, pH 7.0.

Table 3: 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)
GGCGCG	48.8	132.5	7.76	50.8	48.2	130.3	7.74	50.9
GGCGCA	41.7	110.2	7.48	51.0	46.8	126.1	7.68	50.9
CGGCCU	46.3	123.2	8.13	54.4	49.1	131.7	8.28	54.4
AGCGCA	48.7	132.4	7.64	50.1	45.2	121.5	7.53	50.2
UGGCCC	41.2	110.6	6.87	46.3	38.0	100.3	6.86	47.0
CGGCCA	47.1	125.1	8.30	55.3	47.7	127.0	8.33	55.3
reference duplexes								
GCGC[c,d]	30.5	83.4	4.61	26.5				
GGCC[e]	35.8	98.1	5.37	27.1				

[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] Solutions are 1 M NaCl, 10 mM sodium phosphate, and 0.5 mM EDTA, pH 7. [d] Freier et al. (1985). [e] Freier et al. (1983).

of small and large subunit rRNAs (Gutell et al., 1985, 1993). Interestingly, the relative occurrence of the first mismatch trends with the stability of the loop.

Thermodynamic parameters for some hairpins were also measured in 0.1 M NaCl. These are listed in Table 2. The trends are essentially the same as at 1 M NaCl, but the hairpins are less stable by an average of 0.6 kcal/mol at 37 °C.

Jaeger et al. (1989) suggested that the stability of hairpin loops depends on the stacking interactions of the first mismatch with the closing base pair. To test this hypothesis, thermodynamic parameters for several mismatches, XY, on a CG base pair were measured using self-complementary oligomers of the type YGCGCX or YGGCCX. Plots of T_m^{-1} vs log C_t are shown in Figure 2, and thermodynamic parameters are presented in Table 3. The nearest neighbor

Table 4: Thermodynamic Parameters for Terminal Mismatches on a CG Pair (1 M NaCl)

→
CX
GY
←

X	A	C	G	U
		$-\Delta H°$ (kcal/mol)		
A	9.1	5.6	5.6	
C	(5.7)	(3.4)		2.7
G	8.2[b]		9.2	
U		5.3		8.6[c]
		$-\Delta S°$ (eu)		
A	24.5	13.5	13.4	
C	(15.2)	(7.6)		6.3
G	21.8[b]		24.6	
U		12.6		23.9[c]
		$-\Delta G°_{37}$[a] (kcal/mol)		
A	−1.5	−1.5	−1.4	
C	(−1.0)	(−1.1)		−0.8
G	−1.4[b]		−1.6	
U		−1.4		−1.2[c]

[a] $\Delta G°_{37}$'s calculated as illustrated in text. Values in parentheses are estimated. [b] SantaLucia et al. (1991b). [c] Freier et al. (1986b).

thermodynamic parameters for a terminal mismatch on a CG base pair are derived from equations equivalent to (Hickey & Turner, 1985): $\Delta G°\binom{CX}{GY} = 0.5[\Delta G°(YGGCCX) - \Delta G°(GGCC)]$. Table 4 summarizes these parameters.

Since a large fraction of natural hairpin loops starts with a GA mismatch, the dependence of stability on the closing base pair was investigated for the hairpin loop GUAAUA. These results are listed in Table 5. When compared with previous results for the hairpin loop AUAAUA (Serra et al., 1993), the GA mismatch stabilizes the hairpin by 1.0, 1.1, 0.7, and 0.7 kcal/mol, respectively, for AU, UA, GC, and CG closure.

To test the generality of conclusions from this work, thermodynamic parameters were also measured for four hairpins of six nucleotides that occur in small and large subunit rRNAs. These results are listed in Table 6.

DISCUSSION

Due to lack of experimental data, previous models for predicting the stability for RNA hairpins have assumed that stability depends only on the number of nucleotides in the loop and the closing base pairs (Tinoco et al., 1973) or on the size of the loop plus interactions of the first mismatch with the closing base pair (Jaeger et al., 1989). The free energy differences for the hairpins listed in Table 1 allow the development of a more accurate model for predicting hairpin stability.

For hairpins closed by CG or GC, Figure 3 shows a plot of the free energy change for hairpin loop formation, $\Delta G°_{37L}$, versus the free energy increment for the first mismatch in the loop. There is a trend in which loops with more favorable mismatches are more stable. Loops with first mismatches of GA and UU appear to be somewhat extra stable, as has also been observed for internal loops with tandem mismatches (SantaLucia et al., 1991a). When these GA and

Table 5: Thermodynamic Parameters for Hairpin Formation with GA First Mismatch[a]

RNA Hairpin	T_M (°C)	$\Delta H°$ Meas kcal/mol	$\Delta S°$ Meas (eu)	$\Delta G°_{37}$ Meas (kcal/mol)	$\Delta G°_{L37}$[b] Meas (pred.) (kcal/mol)
GUG GU A[d] / CAU AU A	19.0	-21.3±5.8	-71.7±19.2	+1.08±0.30	4.0
GGU GU A / CCA AU A	50.2	-36.9±2.3	-114.0±7.4	-1.50±0.05	3.6 (3.2)
GGA GU A / CCU AU A	51.8	-37.2±2.2	-114.4±7.1	-1.73±0.12	3.5 (3.5)
GGU GU A / CCG AU A	53.3	-37.7±0.6	-115.6±2.2	-1.87±0.05	3.1
GUG GU A / CAC AU A	48.5	-29.4±1.7	-91.3±6.6	-1.05±0.18	2.9 (2.7)
GGC GU A[e] / CCG AU A	68.7	-36.9±2.4	-107.8±7.2	-3.42±0.24	2.9 (2.9)

[a] Measurements were made in 1.0 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7. Errors in $\Delta H°$, and $\Delta G°$ are standard deviations. $\Delta G°_{37}$ was calculated from $\Delta H°$ and $\Delta S°$ before rounding off. [b] $\Delta G°_{L37} = \Delta G°_{37}$(hairpin formation) − $\Delta G°_{37}$(pred stem). [c] $\Delta G°_{L37}$ predicted as described in text. [d] Low-melting hairpin prevented accurate determination of lower base lines and leads to unusually large standard deviations. [e] Serra et al. (1993).

Table 6: Thermodynamic Parameters for Hairpin Formation of Natural Sequences in 1 M NaCl

Hairpin		T_M meas (°C)	$\Delta H°$ meas (kcal/mol)	$\Delta S°$ meas (eu)	$\Delta G°$ meas (kcal/mol)	$\Delta G°_{37L}$ meas (pred) (kcal/mol)
A A U U G G U•A G•C G•C	a	46.6	-35.4±1.7	-110.8±5.2	-1.0±0.2	4.0 (4.2)
A A U U U U G•C G•C C•G	b	54.7	-36.1±2.8	-110.1±8.9	-1.9±0.2	3.0 (3.6)
A C C C A A U•A C•G U•A C•G	c	53.1	36.7±3.9	-112.6±12.2	-1.8±0.2	3.9 (4.4)
C A A C G A C•G C•G A•U	d	52.9	34.2±4.0	-105.1±10.2	-1.6±0.2	3.4 (2.9)

[a] Sequence modeled on *Zea mays* SRP-RNA U(GUAGCG)A position 10 (Zwieb, 1989). [b] *Homo sapien* small subunit rRNA position 1090 (Gutell et al., 1993). [c] Bovine mitochondrial small subunit rRNA position 1265 (Gutell et al., 1993). [d] *Escherichia coli* large subunit rRNA position 1612 (Gutell et al., 1993).

UU mismatches are omitted, a linear fit to the data in Figure 3 gives $\Delta G°_{37L}(6) = 5.4 + 1.3\Delta G°_{37MM}$ ($r = 0.89$; average deviation = 0.17 kcal/mol), where $\Delta G°_{37MM}$ is the free energy increment for the first mismatch in the loop as measured at duplex ends (Table 4). The coefficient of 1.3 suggests interactions of mismatches at the base of hairpin loops may be somewhat more favorable than at duplex ends. A similar context dependence has been observed for hydrogen bonding between GA mismatches in internal loops, hairpin loops, and duplex ends (SantaLucia et al., 1992). The data in Figure 3, however, are fit almost as well by $\Delta G°_{37L}(6) = 5.0 + \Delta G°_{37MM}$ (average deviation = 0.20 kcal/mol). Given experimental error and the likelihood that any context dependence of mismatch interactions gets smaller as the loop gets larger, this simple equation is probably a more reasonable form for hairpin loops in general.

A previous study of AUAAUA hairpin loops indicated loops closed by AU or UA are less stable than loops closed by CG or GC (Serra et al., 1993). The less favorable stacking of mismatches on AU and UA pairs compared with CG and GC pairs (Turner et al., 1988) probably accounts for part of the difference. Comparison of the data of Serra et al. (1993) and the U(GUAAUG)A hairpin loop in Table 6 with the plot in Figure 3, however, suggests hairpin loops closed by AU or UA are an additional 0.4 kcal/mol less stable on average, when the first mismatch is not GA or UU.

Hairpin loops with the UAAU motif, GA or UU as the first mismatch, and Watson–Crick closing base pairs are more stable by 0.7 kcal/mol on average when compared with the simple model presented above (Tables 1 and 5, calculation done before rounding off). The additional stability of loops with first mismatches of GA and UU is likely due to hydrogen bonding. GA and UU mismatches form stable hydrogen bonds when present in tandem mismatches in RNA internal loops (SantaLucia et al., 1991a,b; SantaLucia & Turner, 1993). NMR-derived structures of GCAA and GAAA tetraloops include hydrogen bonding between the G and A in the first mismatches (Heus & Pardi, 1991), and optical melting studies indicate this makes a small contribution to stability (SantaLucia et al., 1992).

Collectively, the results presented here and in Serra et al. (1993) suggest the following simple model for the free energy increment of a hairpin loop of $n \geq 4$ nucleotides:

$$\Delta G°_{37L}(n) = \Delta G°_{37i}(n) + \Delta G°_{37MM} + 0.4 \text{ (if closed by AU or UA)} \quad -0.7 \text{ (if first mismatch is GA or UU)} \quad (2)$$

Here $\Delta G°_{37i}(n)$ is the free energy for initiating a loop of n nucleotides. Based on the data in Table 1 and the literature

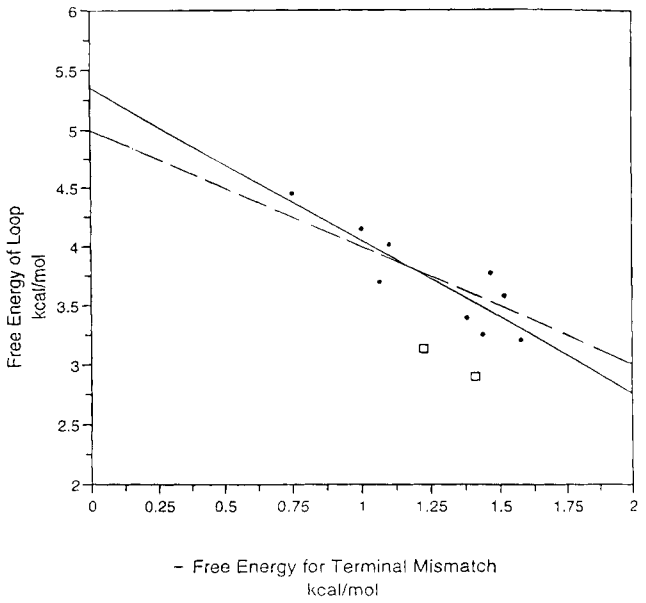

FIGURE 3: Plot of the free energy change for hairpin loop formation, $\Delta G°_{37L}$, versus the free energy increment, $\Delta G°_{37MM}$, for the first mismatch of the loop. Loops are closed by GC or CG. First mismatches GA and UU, □: all others, ●. Solid line is least-squares fit to filled in circles. Dashed line is least-squares fit to filled in circles with slope forced to be 1.

(Groebe & Uhlenbeck, 1988; Gralla & Crothers, 1973; Riesner et al., 1973; Coutts et al., 1974; Serra et al., 1993), $\Delta G°_{37i}(n)$ for $n = 4-9$ is 4.9, 4.4, 5.0, 5.0, 5.1, and 5.2 kcal/mol, respectively.

The above model was used to predict $\Delta G°_{37L}$ for the sequences listed in Tables 1, 5, and 6 and for those studied previously (Serra et al., 1993). Predicted values are listed in parentheses in Tables 1, 5, and 6. For sequences used to develop the model, $\Delta G°_{37L}$ is predicted within 0.4 kcal/mol (Tables 1 and 5). For several naturally occurring hairpins not included in development of the model, $\Delta G°_{37L}$ is predicted within 0.6 kcal/mol (Table 6). This is good agreement considering the simplicity and uncertainty of the nearest neighbor model and suggests the model contains the major determinants of hairpin stability.

The model of hairpin loop stability is consistent with most previous observations for the sequence dependence of hairpin loops between four and seven nucleotides (supplementary material; see paragraph at end of paper regarding supplementary material). SantaLucia et al. (1992) found that the C(GCAA)G hairpin loop is 0.8 kcal/mol more stable than C(ACAA)G at 37 °C, consistent with eq 2. Tetraloops with first mismatches of GA or UU were found to be 0.5−0.6 kcal/mol more stable relative to loops with AA first mismatches (Antao & Tinoco, 1992). Groebe and Uhlenbeck (1988) determined the thermodynamic stabilities for hairpin loops of sequence $C(A_n)G$ and $C(U_n)G$. For $n = 4$, 5, and 7, the U_n loops are more stable by an average of 0.4 kcal/mol, as predicted by eq 2. The U_9 loop, however, is 2.4 kcal/mol more stable than A_9, suggesting new terms may be important for large loops. Loops of three nucleotides are also not predicted well. This may reflect the fact that as the loop becomes small it is more constrained so that interactions of the first mismatch differ from the interactions of terminal mismatches on a duplex. For the majority of hairpin loops found in natural RNA, however, eq 2 should provide a better approximation of stability than those used previously. This should help improve predictions of RNA secondary structure from sequence.

SUPPLEMENTARY MATERIAL AVAILABLE

One table listing measured and predicted $\Delta G°_{L,37}$ for hairpins studied by Groebe and Uhlenbeck (1988), Antao and Tinoco (1992), and Puglisi et al. (1988) (1 page). Ordering information is given on any current masthead page.

REFERENCES

Antao, V. P, & Tinoco, I., Jr. (1992) *Nucleic Acids Res. 20*, 819−824.

Antao, V. P, Lai, S. Y., & Tinoco, I., Jr. (1991) *Nucleic Acids Res. 19*, 5901−5905.

Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology*: Nucleic Acids (Fasman, C. D., Ed.) 3rd ed., Vol. I, p 589, CRC, Cleveland, OH.

Borer, P. N., Dengler, B., Tinoco, I., Jr., & Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843−853.

Brimacombe, R., Atmadja, J., Stiege, W., & Schuler, D. (1988) *J. Mol. Biol. 199*, 115−136.

Cheong, C., Varani, G., & Tinoco, I., Jr. (1990) *Nature* 346, 680−682.

Chou, S. H., Flynn, P., & Reid, B. (1989) *Biochemistry 28*, 2422−2435.

Coutts, S. M., Gangloff, J., & Dirheimer, G. (1974) *Biochemistry 13*, 3938−3948.

DeLisi, C., & Crothers, D. M. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 2682−2685.

Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., & Turner, D. H. (1983) *Biochemistry 22*, 6198−6206.

Freier, S. M., Sinclair, A., Neilson, T., & Turner, D. H. (1985) *J. Mol. Biol. 185*, 645−647.

Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986a) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.

Freier, S. M., Kierzek, R., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986b) *Biochemistry 25*, 3209−3213.

Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 73*, 497−511.

Groebe, D. R., & Uhlenbeck, O. C. (1988) *Nucleic Acids Res. 16*, 11725−11735.

Gutell, R. R., Weiser, B., Woese, C. R., & Noller, H. F. (1985) *Proc. Nucleic Acid Res. Mol. Biol. 32*, 155−216.

Gutell, R. R., Gray, M. W., & Schnare, M. N. (1993) *Nucleic Acids Res. 21*, 3055−3074.

Gutell, R. R., Larsen, N., & Woese, C. R. (1994) *Microbiol. Rev. 58*, 10−26.

Heus, H. A., & Pardi, A. (1991) *Science 253*, 191−194.

Hickey, D. R., & Turner, D. H. (1985) *Biochemistry 24*, 2086−2094.

Jaeger, J. A., Turner, D. H., & Zuker, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7706−7710.

Kim, S. H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H. J., Seeman, N. C., & Rich, A. (1974) *Science 185*, 435−440.

Major, F., Turcotte, M., Gautheret, D., LaPalme, G., Fillion, E., & Cedergren, R. (1991) *Science 253*, 1255−1260.

Malhotra, A., Tan, R. K. Z., & Harvey, S. C. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1950−1954.

Michel, F., & Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.

Noller, H. F. (1984) *Annu. Rev. Biochem. 53*, 119−162.

Parker, R. (1989) in *Methods in Enzymology* (Dahlberg, J. E., & Abelson, J. N., Eds.) Vol. 180, p 510, Academic Press, San Diego, CA.

Petersheim, M., & Turner, D. H. (1983) *Biochemistry 22*, 256−263.

Richards, E. G. (1975) in *Handbook of Biochemistry and Molecular Biology*: *Nucleic Acids* (Fasman, C. D., Ed.) 3rd ed., Vol. I, p 197, CRC, Cleveland, OH.

Riesner, D., Maass, G., Thiebe, R., Philippsen, P., & Zachau, H. G. (1973) *Eur. J. Biochem. 36*, 76−88.

Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. D., Clark, B. F. C., & Klug, A. (1974) *Nature 250*, 546−551.

SantaLucia, J., Jr., & Turner, D. H. (1993) *Biochemistry 32*, 12612−12623.

SantaLucia, J., Jr., Kierzek, R., & Turner, D. H. (1991a) *Biochemistry 30*, 8242−8251.

SantaLucia, J., Jr., Kierzek, R., & Turner, D. H. (1991b) *J. Am. Chem. Soc. 113*, 4313−4322.

SantaLucia, J., Jr., Kierzek, R., & Turner, D. H. (1992) *Science 256*, 217−219.

Serra, M. J., Lyttle, M. H., Axenson, T. J., Schadt, C. A., & Turner, D. H. (1993) *Nucleic Acids Res. 21*, 3845−3849.

Stern, S., Bryn, W., & Noller, H. F. (1988) *J. Mol. Biol. 204*, 447−481.

Tinoco, I., Jr., Uhlenbeck, O. C., & Levine, M. D. (1971) *Nature (London) 230*, 362−367.

Tinoco, I., Jr., Borer, P. N., Dengler, B., Levin, M. D., Uhlenbeck, O. C., Crothers, D. M., & Gralla, J. (1973) *Nature (London), New Biol. 246*, 40−41.

Turner, D. H., Sugimoto, N., Jaeger, J. A., Longfellow, C. E., Freier, S. M., & Kierzek, R. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 123−133.

Turner, D. H., Sugimoto, N., & Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.

Usman, N., Ogilvie, K. K., Jiang, M. Y., & Cedergren, R. L. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.

Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A., & Weiner, A. M. (1987) *Molecular Biology of the Gene*, The Benjamin/Cummings Publishing Co., Menlo Park, CA.

Westhof, E., Dumas, P., & Moras, D. (1985) *J. Mol. Biol. 184*, 119−145.

Williams, A. P., Longfellow, C. E., Freier, S. M., Kierzek, R., & Turner, D. H. (1989) *Biochemistry 28*, 4283−4291.

Wimberly, B., Varani, G., & Tinoco, I., Jr. (1993) *Biochemistry 32*, 1078−1087.

Woese, C. R., & Pace, N. R. (1993) in *The RNA World* (Gesteland, R. F., & Atkins, J. F., Eds.) pp 91−117, Cold Spring Harbor Press, Cold Spring Harbor, NY.

Woese, C. R., Winker, S., & Gutell, R. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8467−8471.

Wu, T., Ogilvie, K. K., & Pon, R. T. (1989) *Nucleic Acids Res. 17*, 3501−3517.

Zwieb, C. (1989) *Prog. Nucleic Acid Res. Biol. 37*, 207−234.