Thermodynamics of Nonsymmetric Tandem Mismatches Adjacent to G·C Base Pairs in RNA†

Tianbing Xia, Jeffrey A. McDowell, and Douglas H. Turner*

Department of Chemistry, University of Rochester, Rochester, New York 14627-0216

Received May 7, 1997; Revised Manuscript Received July 31, 1997⊗

ABSTRACT: The thermodynamic stabilities and structures of a series of RNA duplexes containing nonsymmetric tandem mismatches in the context of $^{5'GAGX\,YGAG3'}_{3'CUCWZCUC5'}$, where $^{5'X\,Y3'}_{3'WZ5'}$ are tandem mismatches, were studied by UV melting and imino proton NMR. The contribution of one mismatch to the free energy increment for tandem mismatch formation depends on the identity of the other mismatch. Imino proton NMR indicates that this is partly because the structure of a mismatch is dependent on the adjacent mismatch. The results suggest that differences in size, shape, and hydrogen bonding of the adjacent mismatches play important roles in determining loop stability. A model for predicting stabilities of all possible tandem mismatches is proposed based on these and previous results.

RNA molecules have important biological functions (Watson et al., 1987), including catalysis, regulation, splicing, and coding for proteins. The revolutions in molecular biology have accelerated the discovery of RNAs with particular functions. Understanding structure−function relationships for these molecules depends on knowing three dimensional structures, so RNA structure determination is increasingly interesting. The major structure determination approaches, X-ray diffraction and NMR, however, can't keep pace with the rate of discovery and sequencing of interesting new RNA molecules. Therefore, it is necessary to develop methods to reasonably predict RNA secondary and three dimensional structures from primary sequences. Studies of RNA unfolding indicate secondary structure interactions dominate over tertiary interactions (Crothers et al., 1974; Hilbers et al., 1976; Banerjee et al., 1993; Jaeger et al., 1993; Gluick & Draper, 1994). Thus, thermodynamic parameters measured for secondary structure motifs in small model systems are used to predict secondary structures by free energy minimization (Tinoco et al., 1971; Zuker & Stiegler, 1981; Turner et al., 1988; Zuker, 1989; Jaeger et al., 1989; Walter et al., 1994; Lück et al., 1996; McCaskill, 1990). This method requires a large database of thermodynamic parameters (Tinoco et al., 1973; Freier et al., 1986; Turner et al., 1988; Serra & Turner, 1995).

Internal loops are one motif with a wide variety of possible sequences for which relatively little experimental data are available. Our group has previously studied symmetric tandem mismatches (Wu et al., 1995), e.g., $^{5'GGAC3'}_{3'CAGG5'}$, $^{5'GUUC3'}_{3'CUUG5'}$, etc. The results show that stabilities of tandem mismatches are quite sequence dependent, indicating that the rules for internal loops are complex. Thus, more data are necessary to define and understand the factors determining internal loop stability.

This work is aimed at providing a thermodynamic database for nonsymmetric tandem mismatches, e.g., $^{5'GGAG3'}_{3'CAAC5'}$, $^{5'GGUG3}_{3'CAUC5'}$, etc., and relating loop stability to structure by NMR. A widely used RNA folding algorithm (Walter et al., 1994) assumes that the contributions to loop stability of the first mismatch on each side of an internal loop are independent of each other. Here it is found that, in nonsymmetric tandem mismatches, the contribution of each mismatch depends on the identity of the other mismatch. The results suggest a structural basis for this dependence, thus providing new insight into the interactions determining stability and structure of internal loops. From these results and previous data, a preliminary model is proposed for predicting the free energy increments for all tandem mismatches so that they can be included more realistically in algorithms for predicting RNA folding.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Purification. RNA oligomers were synthesized on an Applied Biosystems 392 DNA/RNA synthesizer from monomers with 2′-hydroxyls protected as the *tert*-butyl-dimethylsilyl ether and 5′-hydroxyls protected as the dimethoxytrityl group (Usman et al., 1987). Upon completion of coupling on the synthesizer, the oligomer was removed from solid support and deprotected by treatment with concentrated ammonia in ethanol (3:1 by volume) at 55 °C overnight. The 2′-silyl protection was removed by treatment with freshly made 1.0 M triethylammonium hydrogen fluoride (50 equiv) in pyridine at 55 °C for 48 h. The crude samples were then dried and partitioned between water and diethyl ether. Then, the samples were desalted through a Sep-pak C-18 cartridge (Waters) and purified on a Si500F thin-layer chromatography plate (Baker) developed with 1-propanol−ammonia−water (55:35:10 by volume). The least mobile product band was visualized with ultraviolet light, cut out, and eluted with water. Then the samples were desalted again with a Sep-pak C-18 cartridge. The purity of each oligomer was checked by HPLC on a C-8 analytical reverse phase column (Hamilton) and was greater than 95%.

UV Melting and Thermodynamic Parameters. Thermodynamic parameters were measured in 1.0 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7.0. Two molecules were also measured in 1.0 M NaCl, 10 mM

† This work is supported by NIH Grant GM 22939.

* Author to whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* October 1, 1997.

S0006-2960(97)01069-6 CCC: $14.00

sodium phosphate, and 0.5 mM Na_2EDTA, pH 7.0. Oligoribonucleotide single strand concentrations, C_T, were calculated from high-temperature absorbances and single-strand extinction coefficients calculated as described previously (Borer, 1975; Richards, 1975). Oligomers were mixed in 1:1 concentration ratios for non-self-complementary duplexes. Small errors in mixing ratios are not expected to affect the results since the effect of mixing ratio not being 1:1 is small up to 50% of excess of one strand (Peritz et al., 1991). Absorbance vs temperature melting curves were measured at 280 nm with a heating rate of 1 °C/min from 0 to 90 °C on a Gilford 250 spectrometer controlled by a Gilford 2527 thermoprogrammer. Thermodynamic parameters for duplex formation were derived by fitting the shape of each curve to the two-state model with sloping base lines using a nonlinear least-squares program (Petersheim & Turner, 1983; McDowell & Turner, 1996), and by plotting the reciprocal of the melting temperature T_M^{-1} vs $\ln(C_T/4)$ for non-self-complementary sequences (Borer, 1974):

$$T_M^{-1} = \frac{R}{\Delta H^\circ}\ln(C_T/4) + \frac{\Delta S^\circ}{\Delta H^\circ} \quad (1)$$

Here, R is the gas constant, 1.987 cal K^{-1} mol^{-1}. Except for $^{5'GAGGCGAG3'}_{3'CUCAACUC5'}$, the ΔH°s derived from the two methods agree within 15%, consistent with the two-state transition model (Turner et al., 1988; Freier et al., 1986; Petersheim & Turner, 1983; Albergo et al., 1981).

Error limits reported for ΔG°_{37}, ΔH°, and ΔS° derived from fitted parameters are standard deviations. The sample covariance of ΔH° and ΔS°, Cov(ΔH°,ΔS°), can be calculated from the standard deviations in ΔG°_{37}, ΔH°, and ΔS° by the following equations (Snedecor & Cochran, 1982):

$$(\sigma_{\Delta G^\circ 37})^2 = (\sigma_{\Delta H^\circ})^2 + T^2(\sigma_{\Delta S^\circ})^2 - 2T\mathrm{Cov}(\Delta H^\circ,\Delta S^\circ) \quad (2)$$

or

$$\mathrm{Cov}(\Delta H^\circ,\Delta S^\circ) = \frac{1}{N-1}\sum_i[(\Delta H_i - \overline{\Delta H})(\Delta S_i - \overline{\Delta S})] \quad (3)$$

where T is the temperature in Kelvin, and $\mathrm{Cov}(\Delta H^\circ,\Delta S^\circ) = r\sigma_{\Delta H^\circ}\sigma_{\Delta S^\circ}$, where r is the correlation coefficient between ΔH° and ΔS°. The covariance derived by eqs 2 and 3 indicates that the correlation coefficient between ΔH° and ΔS° is greater than 0.99; therefore, ΔG°_{37} is a more accurate parameter than either ΔH° or ΔS° individually (Petersheim & Turner, 1983; SantaLucia et al., 1991a).

Error limits reported for ΔH° and ΔS° derived from T_M^{-1} vs $\ln(C_T/4)$ data reflect the standard deviation in the slopes and intercepts (Meyer, 1975) with all data points weighted equally. Error in ΔH° is proportional to the standard deviation in the slope. Error in ΔS° is slightly larger than the error in ΔH° because ΔS° depends on the slope, intercept, and the covariance of the slope and intercept. The error in ΔG°_{37} can be calculated as follows (Bevington, 1969; SantaLucia et al., 1991a):

$$(\sigma_{\Delta G^\circ 37})^2 = (2.303R)^2\left[\sigma_m^2\left(\frac{Tb-1}{m^2}\right)^2 + \sigma_b^2\frac{T^2}{m^2} - 2\mathrm{Cov}(m,b)\frac{(T^2b-T)}{m^3}\right] \quad (4)$$

where m, b, σ_m^2, σ_b^2, and Cov(m,b) are the slope, intercept, variance of slope, variance of intercept, and covariance of slope and intercept, respectively, for eq 1.

NMR Spectroscopy. After purification, RNA oligomers were dialyzed against 0.1 mM Na_2EDTA and double-distilled water for 24 h each, dried down and dissolved in 80 mM NaCl, 0.5 mM Na_2EDTA, and 10 mM sodium phosphates in 90% H_2O and 10% D_2O, pH 7.0. One dimensional imino proton NMR spectra were recorded by a Varian VXR 500S or Varian INOVA 500 spectrometer at 500 MHz. A binomial $1:\bar{3}:3:\bar{1}$ pulse sequence was used to suppress the solvent peak (Hore, 1983). The frequency offset was set to maximize the signal-to-noise ratio at about 13 ppm with the first nodes at 21 and 5 ppm. Spectra were collected with 12 000 points over a sweep width of 12 kHz, multiplied by a 2−4 Hz line-broadening exponential function, and Fourier-transformed on a Sun 4/260 computer running Varian VNMR software.

RESULTS

Phylogenetic Analysis. A survey of phylogenetic structures of 79 group I introns, 101 small subunit rRNAs, and 218 large subunit rRNAs (Damberger & Gutell, 1994; Gutell et al., 1993; Gutell, 1994) revealed 1890 tandem mismatches (Table 1). Most of these are tandem G·Us (1124 occurrences). About 41% of the 766 other tandem mismatches are symmetric $^{AA}_{AA}$, $^{GA}_{AG}$, and $^{UU}_{UU}$ mismatches (see diagonal of Table 1) if closing base pairs are not taken into account. Thermodynamic parameters were measured for tandem mismatches that occur most often, as well as for some that occur rarely.

Thermodynamic Parameters. For all the tandem mismatches, $^{5'X\ Y3'}_{3'WZ5'}$, stability was determined in the same context, $^{5'GAGX\ YGAG3'}_{3'CUCWZ\ CUC5'}$. The G·G mismatch was not studied because of aggregation problems encountered with symmetric tandem G·G mismatches (SantaLucia et al., 1991b). Typical plots of T_M^{-1} vs $\ln(C_T/4)$ are shown in Figure 1 and Supporting Information (see paragraph at the end of paper regarding Supporting Information). Thermodynamic parameters of duplex formation derived from these plots and from averaging parameters obtained by fitting the shape of each individual melting curve to the two-state model are listed in Table 2. Two mismatches with high natural occurrence, $^{5'GAGAAGAG3'}_{3'CUCGGCUC5'}$ and $^{5'GAGAAGAG3'}_{3'CUCAGCUC5'}$, were measured in both cacodylate and phosphate buffer systems; the thermodynamic parameters are essentially the same.

In Table 2, duplexes with mismatches are listed in order of decreasing stability at 37 °C. At the top of Table 2 are thermodynamic parameters for the reference duplex, $^{5'GAG}_{3'CUC}$ $^{GAG3'}_{CUC5'}$. As shown in Table 2, parameters for the reference duplex and for $^{5'GAGUGGAG3'}_{3'CUCGUCUC5'}$ are predicted reasonably well by the nearest-neighbor parameters of Freier et al. (1986) and He et al. (1991). The stability of $^{5'GAGGUGAG3'}_{3'CUCUGCUC5'}$ (ΔG° = −7.63 kcal/mol), however, could not be unambiguously predicted because there are two sets of nearest-neighbor parameters for $^{5'GU3'}_{3'UG5'}$ depending on the flanking base pairs (He et al., 1991). Set I, derived by omitting results for oligomers containing the motif $^{GGUC}_{CUGG}$, predicts a ΔG°_{37} of −5.7 kcal/mol. Set II, derived by omitting results for oligomers containing the motifs $^{CGUG}_{GUGC}$, $^{AGUU}_{UUGA}$, and $^{UGUA}_{AUGU}$, predicts a ΔG°_{37} of −8.2 kcal/mol. Evidently, the average

Table 1: Frequency of Occurrence of Tandem Mismatches[a] in Phylogenetic Structures of 79 Group I Introns, 101 Small Subunit rRNAs, and 218 Large Subunit rRNAs

5′XY3′ / 3′WZ5′	U / U	G / U	C / U	U / G	G / G	A / G	U / C	C / C	A / C	G / A	C / A	A / A
A	9	2	2	0	5	160	0	0	0	6	4	26
A	12	1	3	1	2	21	0	0	3	7	1	
C	14	1	30	0	2	52	0	1	3	3		
A	4	1	1	1	7	3	4	0	1			
G	18	0	0	0	6	2	0	0				
A	7	1	10	0	3	2	1					
C	7	0	1	5	1	188						
U	5	0	1	0	0							
G	1	200	0	150								
G	13	0	0									
U	2	774										
C	99											
G												
U												
U												

[a] The left column represents the $^{5'X}_{3'W}$ mismatch, and the top row represents the $^{Y3'}_{Z5'}$ mismatch, respectively, regardless of loop closing base pairs. The table is symmetric along the diagonal, so the lower-right part is omitted. The total number of occurrences of tandem mismatches is 1890.

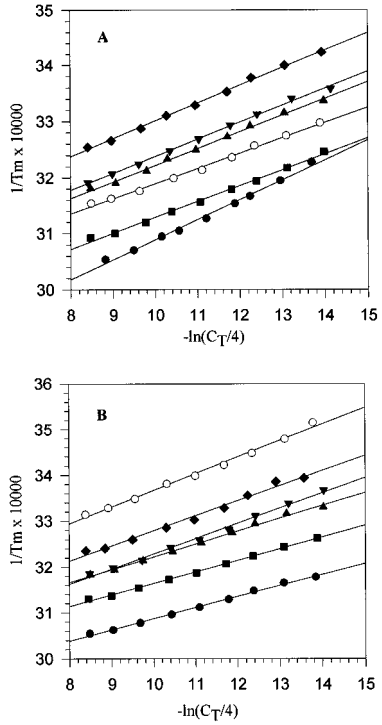


FIGURE 1: Reciprocal of melting temperature versus $-\ln(C_T/4)$ plots for (A) $^{GAGGAG}_{CUCCUC}$ (●), $^{GAGUUGAG}_{CUCGGCUC}$ (■), $^{GAGUUGAG}_{CUCGUCUC}$ (▲), $^{GAGUUGAG}_{CUCUUCUC}$ (▼), $^{GAGUUGAG}_{CUCAGCUC}$ (◆), $^{GAGGAGAG}_{CUCUGCUC}$ (○) and for (B) $^{GAGUGGAG}_{CUCGUCUC}$ (●), $^{GAGGUG}_{AG}$ (■), $^{GAGUGGAG}_{UC}$ (▲), $^{GAGUGGAG}_{CUCUUCUC}$ (▼), $^{GAGUGGAG}_{CUCGACUC}$ (◆), $^{GAGUAGAG}_{CUCUACUC}$ (○). Buffer is 1.0 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7. Slopes ($-R/\Delta H°$) and intercepts ($\Delta S°/\Delta H°$) for each plot give the $1/T_M$ vs $\ln(C_T/4)$ parameters in Table 2 (see eq 1).

of the two sets of parameters (−6.95 kcal/mol) makes a reasonable prediction.

Thermodynamic parameters for internal loop formation are listed in Table 3. These parameters were calculated as follows (Gralla & Crothers, 1973). For any tandem mismatch motif $^{5'GX\,YG3'}_{3'CWZC5'}$, the nearest-neighbor free energy parameter at 37 °C is

$$\Delta G°_{37,\mathrm{loop}}\left(^{5'GX\,YG\,3'}_{3'CWZC\,5'}\right) = \Delta G°_{37}\left(^{5'GAGX\,YGAG3'}_{3'CUCWZCUC5'}\right) - \Delta G°_{37}\left(^{5'GAGGAG3'}_{3'CUCCUC5'}\right) + \Delta G°_{37}\left(^{5'GG3'}_{3'CC5'}\right) \quad (5)$$

In this equation, $\Delta G°_{37}\left(^{5'GAGX\,YGAG3'}_{3'CUCWZCUC5'}\right)$ and $\Delta G°_{37}\left(^{5'GAGGAG3'}_{3'CUCCUC5'}\right)$ are the free energy changes for duplex formation with and without mismatches, respectively, derived from T_M^{-1} vs $\ln(C_T/4)$ plots (Table 2). $\Delta G°_{37}\left(^{5'GG3'}_{3'CC5'}\right)$ is the free energy increment of −2.9 kcal/mol for the nearest-neighbor $\left(^{5'GG3'}_{3'CC5'}\right)$ interaction (Freier et al., 1986) that is interrupted by the tandem mismatches. The free energy increments for loop formation cover a wide range, from about −4.1 to 2.6 kcal/mol. Enthalpy and entropy increments are calculated similarly and also cover a wide range (Table 3).

Imino Proton NMR. One-dimensional imino proton NMR spectra of 12 different duplexes in H_2O are shown in Figure 2. Since these duplexes are non-self-complementary, six resonances are expected for the six Watson−Crick base pairs in each duplex. This is observed for the reference duplex $^{5'GAGGAG3'}_{3'CUCCUC5'}$ (Figure 2a). Resonances were assigned based on chemical shift and NMR melting. In Figure 2a, the two resonances above 14 ppm are typical of A·U base pairs (Hilbers, 1979; Pardi et al., 1981; Ulrich et al., 1983; Chou et al., 1983). The resonances at 12.39 and 13.53 ppm are assigned to the two terminal G·C base pairs because they broaden first when temperature is raised, due to terminal base pair fraying. The other two resonances at 12.30 and 12.78 ppm are assigned to the internal G·C base pairs. The spectra for duplexes with tandem mismatches have features similar to those in the spectrum for the reference duplex, $^{5'GAGGA}_{3'CUCCU}$

Table 2: Thermodynamic Parameters of Duplex Formation

	$1/T_M$ vs $\ln(C_T/4)$ parameters				curve fit parameters			
RNA duplexes[a]	$-\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m[b] (°C)	$-\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m[b] (°C)
GAGGAG	8.50(0.05)	55.7(1.7)	152.2(5.2)	48.4	8.66(0.2)	58.8(4.9)	161.8(15.2)	48.6
CUCCUC	7.50[c]	54.0[c]	149.9[c]	42.7[c]				
GAGUGGAG	9.66(0.06)	82.3(1.8)	234.1(5.6)	49.3	9.59(0.2)	80.4(4.0)	228.2(12.3)	49.3
CUCGUCUC	8.80[d]	73.6[d]	208.9[d]	46.9[d]				
GAGUUGAG	8.24(0.04)	70.0(1.8)	199.2(5.6)	44.7	8.55(0.3)	80.1(8.0)	230.8(25.0)	45.0
CUCGGCUC	8.22(0.15)[e]	73.9(1.9)[e]	211.8(5.9)[e]	44.2[e]	8.30(0.1)[e]	76.8(3.6)[e]	220.9(11.2)[e]	44.3[e]
GAGGUGAG	7.63(0.02)	78.4(1.5)	228.2(4.7)	41.4	7.63(0.1)	76.1(2.7)	220.8(8.7)	41.6
CUCUGCUC								
GAGGAGAG	6.95(0.01)	73.0(1.9)	213.1(6.2)	38.8	7.00(0.2)	62.0(5.0)	177.4(16.6)	39.4
CUCAGCUC								
GAGGAGAG	6.50(0.04)	55.9(2.3)	159.1(7.7)	36.8	6.43(0.1)	59.6(4.1)	171.3(13.2)	36.5
CUCUGCUC								
GAGUUGAG	6.21(0.02)	67.0(1.5)	195.9(4.8)	35.5	6.27(0.1)	68.6(5.4)	201.1(17.5)	35.8
CUCGUCUC								
GAGUGGAG	6.21(0.02)	70.9(1.7)	208.4(5.6)	35.6	6.29(0.1)	67.9(4.0)	198.6(13.1)	35.8
CUCUUCUC								
GAGAGGAG	6.21(0.02)	42.2(1.2)	116.1(3.9)	34.7	6.14(0.2)	43.2(4.6)	119.4(15.1)	34.2
CUCCACUC								
GAGUAGAG	6.09(0.02)	63.8(1.8)	186.0(5.8)	34.9	6.14(0.1)	68.4(7.2)	200.8(23.1)	35.3
CUCGGCUC								
GAGUGGAG	6.07(0.03)	59.9(1.8)	173.5(5.7)	34.7	6.23(0.2)	55.8(6.0)	159.9(20.0)	35.3
CUCGACUC								
GAGUCGAG	5.97(0.03)	60.8(1.4)	176.7(4.6)	34.2	5.97(0.1)	60.7(2.0)	176.5(6.6)	34.2
CUCGACUC								
GAGAGGAG	5.95(0.02)	66.8(1.8)	196.3(6.0)	34.4	6.11(0.3)	55.3(3.4)	158.6(11.6)	34.7
CUCGACUC								
GAGCGGAG	5.93(0.04)	43.3(1.6)	120.3(5.2)	32.7	5.79(0.2)	49.6(2.3)	141.2(7.5)	32.5
CUCAACUC								
GAGUUGAG	5.88(0.02)	65.8(1.2)	193.2(3.8)	34.0	5.98(0.1)	67.6(8.3)	198.6(26.8)	34.5
CUCUUCUC								
GAGCGGAG	5.88(0.06)	36.9(1.7)	99.9(5.5)	31.6	6.00(0.2)	39.7(7.3)	108.9(23.7)	32.9
CUCCACUC								
GAGAGGAG	5.87(0.03)	57.9(1.3)	167.9(4.3)	33.5	5.97(0.1)	52.4(1.1)	149.6(3.2)	33.7
CUCGUCUC								
GAGGCGAG	5.82(0.05)	64.6(2.7)	189.5(8.8)	33.6	5.86(0.1)	60.3(2.7)	175.6(8.5)	33.6
CUCUACUC								
GAGAAGAG	5.78(0.02)	61.1(1.2)	178.5(4.1)	33.3	5.89(0.1)	58.6(5.2)	170.0(17.3)	33.7
CUCGGCUC	5.49(0.05)[f]	66.0(2.1)[f]	195.1(7.1)[f]	32.2[f]	5.66(0.2)[f]	58.5(1.9)[f]	170.3(6.5)[f]	32.5[f]
GAGGAGAG	5.74(0.06)	48.6(2.3)	138.2(7.5)	32.1	5.69(0.2)	55.7(6.5)	161.4(20.5)	32.4
CUCUACUC								
GAGCAGAG	5.60(0.03)	45.8(1.0)	129.6(3.3)	30.8	5.48(0.1)	50.9(1.6)	146.3(5.1)	30.7
CUCAGCUC								
GAGUCGAG	5.57(0.04)	62.5(1.8)	183.6(6.0)	32.3	5.65(0.1)	58.3(1.7)	169.6(5.7)	32.4
CUCGCCUC								
GAGGAGAG	5.56(0.06)	62.4(3.4)	183.3(11)	32.2	5.68(0.2)	53.5(1.2)	154.1(3.8)	32.1
CUCAACUC								
GAGAAGAG	5.42(0.03)	58.3(1.4)	170.6(4.7)	31.2	5.56(0.2)	54.0(4.0)	156.2(13.3)	31.5
CUCAGCUC	5.33(0.04)[f]	69.1(1.7)[f]	205.5(5.5)[f]	31.7[f]	5.56(0.2)[f]	59.5(3.2)[f]	173.9(10.8)[f]	32.0[f]
GAGUAGAG	5.42(0.16)	50.6(4.2)	145.7(14)	30.3	5.38(0.2)	51.7(5.0)	149.3(15.9)	30.3
CUCUCCUC								
GAGGCGAG	5.37(0.05)	63.3(1.7)	186.7(5.6)	31.4	5.42(0.1)	61.3(1.9)	180.2(6.3)	31.5
CUCUUCUC								
GAGAGGAG	5.30(0.27)	52.2(6.0)	151.3(20)	29.9	5.32(0.3)	55.7(8.1)	162.6(25.9)	30.4
CUCAUCUC								
GAGUCGAG	5.30(0.06)	59.4(2.2)	174.4(7.4)	30.7	5.34(0.1)	56.1(2.5)	163.7(8.2)	30.6
CUCUCCUC								
GAGUCGAG	5.22(0.08)	43.2(2.1)	122.4(6.9)	27.9	5.06(0.2)	46.6(6.1)	133.8(19.7)	27.5
CUCUACUC								
GAGAGGAG	5.20(0.1)	52.0(2.8)	150.8(9.2)	29.3	5.29(0.1)	49.1(3.1)	141.3(9.8)	29.3
CUCAACUC								
GAGGUGAG	5.19(0.04)	63.7(1.5)	188.6(5.0)	30.6	5.41(0.3)	58.1(5.1)	169.7(17.3)	31.1
CUCUUCUC								
GAGUUGAG	5.17(0.03)	60.5(1.0)	178.5(3.2)	30.2	5.18(0.1)	61.1(1.4)	180.2(4.7)	30.3
CUCCUCUC								
GAGGCGAG	5.14(0.04)	58.8(1.4)	173.0(4.6)	29.8	5.40(0.1)	47.3(5.1)	135.0(16.0)	29.7
CUCAACUC								
GAGUAGAG	5.12(0.08)	60.7(2.5)	179.1(8.3)	30.0	5.13(0.4)	59.7(8.9)	175.9(30.1)	29.9
CUCGACUC								
GAGUUGAG	5.06(0.04)	56.9(1.4)	167.0(4.5)	29.2	5.04(0.1)	56.7(2.8)	166.4(9.1)	29.1
CUCUCCUC								
GAGGUGAG	4.99(0.04)	60.3(1.2)	178.4(4.1)	29.3	5.23(0.1)	52.1(2.8)	151.3(9.0)	29.5
CUCAGCUC								
GAGAUGAG	4.93(0.08)	46.3(2.2)	133.3(7.3)	26.6	4.97(0.2)	47.6(4.7)	137.5(15.4)	27.2
CUCCUCUC								
GAGAAGAG	4.89(0.06)	55.7(1.7)	163.8(5.6)	28.1	5.03(0.1)	50.7(2.3)	147.3(7.3)	28.1
CUCGACUC								

Table 2 (continued)

	$1/T_M$ vs $\ln(C_T/4)$ parameters				curve fit parameters			
RNA duplexes[a]	$-\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m[b] (°C)	$-\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m[b] (°C)
GAGACGAG CUCGACUC	4.79(0.08)	57.1(2.1)	168.6(7.1)	27.8	5.03(0.2)	50.6(1.6)	146.9(5.1)	28.1
GAGGUGAG CUCAUCUC	4.75(0.06)	57.5(1.3)	170.0(4.3)	27.7	4.79(0.1)	56.3(1.4)	166.0(4.8)	27.7
GAGAAGAG CUCAACUC	4.75(0.05)	47.6(1.0)	138.3(3.4)	25.8	4.76(0.1)	46.9(2.7)	135.8(9.2)	25.7
GAGACGAG CUCGUCUC	4.71(0.08)	59.8(2.2)	177.7(7.5)	27.9	4.74(0.2)	61.8(6.1)	183.9(20.3)	28.3
GAGUUGAG CUCUGCUC	4.68(0.04)	62.5(1.1)	186.4(3.6)	28.1	4.86(0.2)	58.4(4.4)	172.7(14.7)	28.4
GAGAUGAG CUCGGCUC	4.67(0.09)	55.6(1.9)	164.2(6.2)	26.9	4.94(0.2)	49.0(3.8)	142.2(12.7)	27.3
GAGACGAG CUCAACUC	4.62(0.16)	44.5(2.6)	128.7(8.9)	24.2	4.66(0.2)	43.7(4.3)	125.7(14.3)	24.2
GAGAAGAG CUCCGCUC	4.58(0.10)	53.4(2.5)	157.3(8.3)	26.1	4.82(0.2)	46.2(4.1)	133.4(13.8)	26.0
GAGAAGAG CUCGCCUC	4.55(0.12)	53.0(2.8)	156.4(9.5)	25.8	4.59(0.1)	52.1(2.6)	153.2(8.6)	25.9
GAGGCGAG CUCACCUC	4.52(0.08)	48.6(1.5)	142.3(5.0)	24.7	4.82(0.4)	41.1(7.7)	116.9(25.7)	24.6
GAGUCGAG CUCCUCUC	4.52(0.09)	52.9(2.0)	155.9(6.7)	25.6	4.84(0.2)	44.9(3.1)	129.0(10.1)	25.8
GAGACGAG CUCGCCUC	4.50(0.05)	47.3(1.0)	137.9(3.5)	24.2	4.61(0.1)	45.4(3.4)	131.5(11.2)	24.4
GAGUGGAG CUCUACUC	4.34(0.08)	50.4(1.4)	148.4(4.9)	24.1	4.39(0.2)	48.3(3.8)	141.5(12.8)	23.9
GAGAUGAG CUCAGCUC	4.32(0.11)	54.6(2.3)	162.0(7.6)	24.9	4.59(0.1)	48.2(2.0)	140.5(6.2)	25.0
GAGUAGAG CUCUGCUC	4.31(0.05)	55.9(1.2)	166.5(3.9)	25.1	4.53(0.3)	52.9(6.6)	155.8(22.2)	25.7
GAGACGAG CUCCACUC	4.16(0.07)	53.8(1.3)	160.1(4.3)	23.9	4.44(0.3)	47.4(5.2)	138.6(17.7)	23.9
GAGAUGAG CUCGUCUC	3.84(0.10)	55.3(1.4)	165.8(5.0)	22.6	3.75(0.4)	56.4(4.6)	169.7(16.1)	22.4
GAGAUGAG CUCAUCUC	3.82(0.01)	57.2(1.4)	172.0(4.9)	23.0	4.05(0.2)	53.1(2.5)	158.0(8.7)	23.1
GAGUAGAG CUCUACUC	3.74(0.10)	54.7(1.7)	164.2(5.7)	22.0	3.82(0.2)	52.8(3.6)	157.8(12.1)	21.8
GAGUUGAG CUCCCCUC	3.56(0.07)	47.8(1.1)	142.7(3.6)	18.9	3.2(0.1)	54.2(3.3)	164.3(10.8)	19.0
GAGACGAG CUCCCCUC	3.33(0.10)	49.8(1.6)	149.8(5.4)	18.2	3.53(0.1)	46.2(2.6)	137.5(8.7)	18.1
GAGACGAG CUCACCUC	3.03(0.09)	53.0(1.2)	161.1(4.3)	17.8	3.37(0.8)	48.1(11)	144.3(39.6)	17.9

[a] Listed in order of decreasing loop stability. Top strand is in a 5′ to 3′ direction, bottom stand is in 3′ to 5′ direction. [b] Calculated for 10^{-4} M oligomer concentration. [c] Predicted from nearest-neighbor parameters (Freier et al., 1986). [d] Predicted from nearest-neighbor parameters (Freier et al., 1986; He et al., 1991). [e] Measured by He et al., 1991. [f] Measured in sodium phosphate buffer system.

$^{G3'}_{C5'}$. In these spectra, extra resonances are presumably from the mismatches.

The spectrum for $^{5'GAGAAGAG3'}_{3'CUCGGCUC5'}$ (Figure 2b) has eight resonances. The two extra imino proton resonances downfield of 11.5 ppm suggest that both guanines in A·G mismatches are in the imino hydrogen-bonded A·G conformation shown in Figure 3 (Cheng et al., 1992; Walter et al., 1994; Wu & Turner, 1996). The spectrum for $^{5'GAGAGGAG3'}_{3'CUCGACUC5'}$ (Figure 2c) has seven resonances, but integration of the peak at 11.95 ppm suggests that it arises from two protons. This spectrum is very similar to that for $^{5'GAGAAGAG3'}_{3'CUCGGCUC5'}$. Evidently, both A·G mismatches in $^{5'GAGAGGAG3'}_{3'CUCGACUC5'}$ are also imino hydrogen bonded.

Four extra resonances between 10.3 and 11.2 ppm are observed in the spectrum for $^{5'GAGUUGAG3'}_{3'CUCUUCUC5'}$ (Figure 2d). They are characteristic of U·U mismatches (SantaLucia et al., 1991b; Wu et al., 1995; Nikonowicz & Pardi, 1992). These sharp resonances indicate that all four U imino protons are protected from exchange with water and presumably these NMR features reflect tight imino proton hydrogen bonding in both U·U mismatches.

There are nine resonances in the spectrum for $^{5'GAGUAGA}_{3'CUCUGCU\ G3'\ C5'}$ (Figure 2e). The three resonances between 10.1 and 11.4 ppm are presumably from the U·U and G·A mismatches. Note that these resonances are much broader than those in $^{5'GAGUUGAG3'}_{3'CUCUUCUC5'}$. This suggests that the U·U mismatch in $^{5'GAGUAGAG3'}_{3'CUCUGCUC5'}$ is not as tightly hydrogen bonded. A broad resonance observed upfield of 11.5 ppm for an imino proton in a G·A mismatch is considered to indicate the sheared conformation shown in Figure 3 (SantaLucia & Turner, 1993; Walter et al., 1994; Cheng et al., 1992; Li et al., 1991; Green et al., 1994). This is consistent with the fact that there is no NOE from these resonances to an AH2 (data not shown). The spectrum for $^{5'GAGGUGAG3'}_{3'CUCAUCUC5'}$ (Figure 2f) is similar to that for $^{5'GAGUAGAG3'}_{3'CUCUGCUC5'}$; thus, the G·A mismatch in this motif also appears to have a sheared conformation. The resonances at 10.10 and 10.98 ppm in the spectrum for $^{5'GAGAUGAG3'}_{3'CUCGUCUC5'}$ (Figure 2g) are from either the two mismatched Us or one

Table 3: Thermodynamic Parameters for Loop Formation by Nonsymmetric Tandem Mismatches[a]

GXYG / CWZC	G / U	U / G	A / G	G / A	U / U	A / C	U / C	C / U	C / C	C / A	A / A
(a) Free Energy Increments ($\Delta G^\circ_{37,loop}$) for Loop Formation in Kilocalories per Mol[b]											
U	−4.06	−2.64	−0.49	−0.47	−0.61				0.03	−0.37	0.48
G	(0.08)	(0.06)	(0.05)	(0.06)	(0.05)				(0.06)	(0.06)	(0.09)
G		−2.03	−0.90		0.41			0.23		−0.22	−0.14
U		(0.05)	(0.06)		(0.06)			(0.07)		(0.07)	(0.08)
G		0.61	−1.35		0.85				1.08	0.46	0.04
A		(0.06)	(0.05)		(0.08)				(0.09)	(0.06)	(0.08)
A	−0.27	0.93	−0.18	−0.35	1.76	1.05		0.89	1.10	0.81	0.71
G	(0.06)	(0.10)	(0.05)	(0.05)	(0.11)	(0.13)		(0.09)	(0.07)	(0.09)	(0.08)
U	−0.61	0.92	1.29	1.26	−0.28	0.18	0.54		0.30	0.38	1.86
U	(0.05)	(0.06)	(0.07)	(0.09)	(0.05)	(0.17)	(0.06)		(0.08)	(0.09)	(0.11)
C			0.0	−0.33							
A			(0.06)	(0.06)							
U					0.43		2.04				
C					(0.06)		(0.09)				
C				−0.28							
C				(0.08)							
A			1.02	−0.61	0.67				2.27	1.44	
C			(0.11)	(0.05)	(0.09)				(0.11)	(0.09)	
A	0.30	1.28	0.18	0.40	1.78				2.57	0.98	0.85
A	(0.27)	(0.12)	(0.06)	(0.11)	(0.05)				(0.10)	(0.17)	(0.07)
(b) Enthalpy Change Increments ($\Delta H^\circ_{37,loop}$) for Loop Formation in Kilocalories per Mol											
U	−38.8	−26.5	−20.3	−16.4	−23.5				−19.0	−17.3	−17.2
G	(2.5)	(2.5)	(2.5)	(2.5)	(2.3)				(2.5)	(2.2)	(3.0)
G		−34.9	−12.4		−20.2			−19.8		−21.1	−5.1
U		(2.3)	(2.9)		(2.3)			(2.4)		(3.2)	(2.9)
G		−16.8	−29.5		−14.0				−5.1	−15.3	−18.0
A		(2.1)	(2.5)		(2.1)				(2.3)	(2.2)	(3.8)
A	−14.4	−12.1	−17.6	−23.3	−11.8	−9.5		−16.3	−3.8	−13.6	−12.2
G	(2.1)	(2.5)	(2.1)	(2.5)	(2.2)	(3.3)		(2.8)	(2.0)	(2.7)	(2.4)
U	−27.4	−19.0	−12.4	−6.9	−22.3	−7.1	−13.4		−15.9	0.3	−11.2
U	(2.4)	(2.0)	(2.1)	(2.2)	(2.1)	(4.5)	(2.2)		(2.8)	(2.7)	(2.4)
C			−2.3	0.2							
A			(2.0)	(2.3)							
U					−17.0		−4.3				
C					(2.0)		(2.0)				
C				6.6							
C				(2.4)							
A			−9.9	1.3	−2.8				−6.3	−10.3	
C			(3.0)	(2.1)	(2.8)				(2.3)	(2.1)	
A	−8.7	−11.1	−14.8	−8.5	−13.7				−9.5	−1.0	−4.1
A	(6.2)	(2.9)	(2.2)	(3.3)	(2.2)				(2.1)	(3.1)	(2.0)
(c) Entropy Change Increments ($\Delta S^\circ_{37,loop}$) for Loop Formation in Entropy Units											
U	−111.6	−76.7	−63.5	−51.0	−73.4				−61.1	−54.2	−56.6
G	(7.6)		(7.8)	(7.7)	(7.1)				(7.9)	(6.9)	(9.8)
G		−105.7	−36.6		−66.1			−64.2		−67.0	−15.7
U		(7.0)	(9.3)		(7.2)			(7.6)		(10.2)	(9.1)
G		−55.9	−90.6		−47.5				−19.8	−50.5	−60.8
A		(6.6)	(8.1)		(6.7)				(7.2)	(6.9)	(12.2)
A	−45.4	−41.7	−56.0	−73.8	−43.3	−33.9		−55.2	−15.4	−46.1	−41.3
G	(6.7)	(8.1)	(6.6)	(7.9)	(7.2)	(10.8)		(9.1)	(6.3)	(8.8)	(7.6)
U	−85.9	−63.9	−44.0	−25.9	−70.7	−23.2	−44.5		−51.9	0.10	−41.7
U	(7.6)	(6.3)	(6.5)	(7.1)	(6.4)	(14.9)	(6.9)		(9.0)	(8.6)	(7.7)
C			−7.1	2.2							
A			(6.2)	(7.4)							
U					−56.0		−20.2				
C					(6.1)		(6.3)				
C				22.6							
C				(7.6)							
A			−34.8	6.4	−10.8				−27.3	−37.6	
C			(9.8)	(6.5)	(9.0)				(7.5)	(6.7)	
A	−28.8	−39.5	−48.1	−28.3	−49.5				−38.6	−6.2	−15.8
A	(20.7)	(9.2)	(7.0)	(10.6)	(7.1)				(6.7)	(10.3)	(6.2)

[a] The left column represents the $^{5'X}_{3'W}$ mismatches, while the top row represents the $^{Y3'}_{Z5'}$ mismatches in the context of $^{5'GX\,YG3'}_{3'CWZ\,C5'}$. Values in parentheses are estimated errors. [b] Calculated by eq 5.

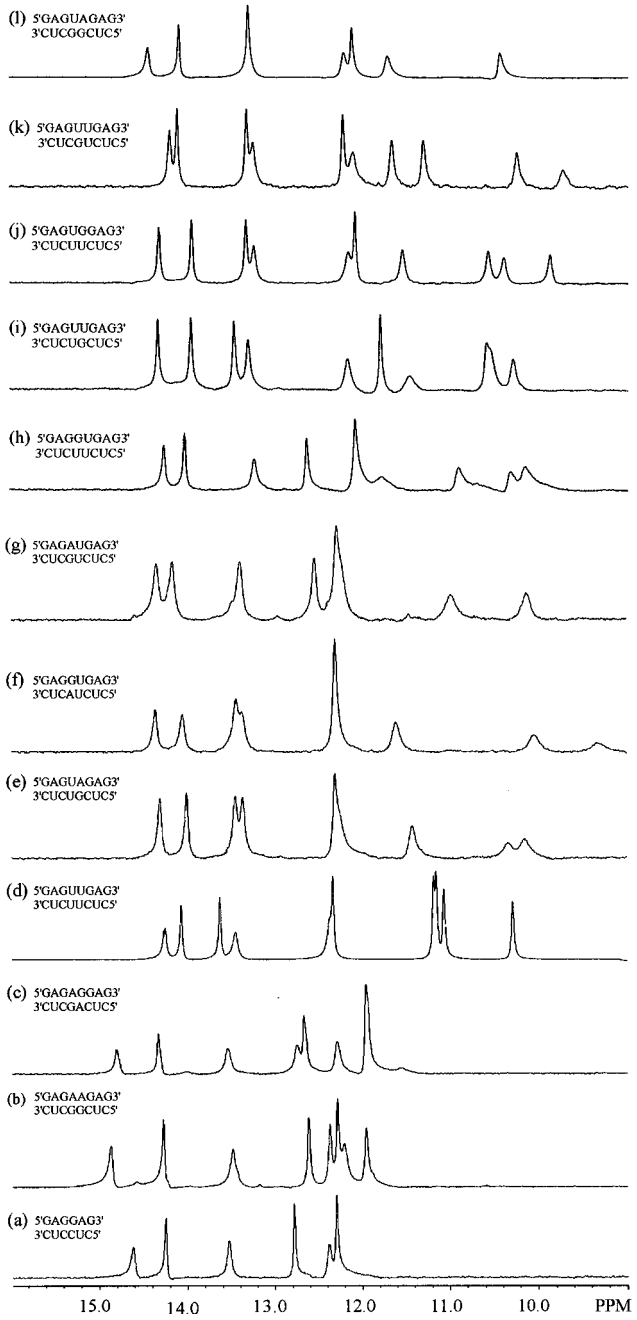
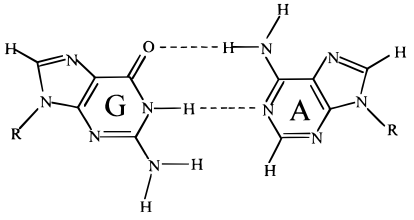


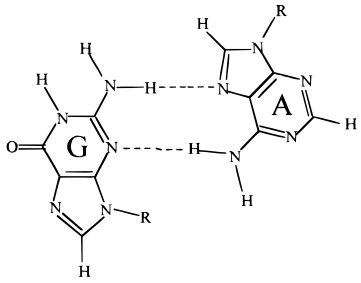
FIGURE 2: Imino proton NMR spectra (500 MHz) in 80 mM NaCl, 0.5 mM Na_2EDTA, and 10 mM sodium phosphates in 90% H_2O and 10% D_2O at pH 7 for (bottom to top) (a) 0.5 mM $^{GAGGAG}_{CUCCUC}$ at 10 °C, (b) 1.4 mM $^{GAGAAGAG}_{CUCGGCUC}$ at 10 °C, (c) 0.75 mM $^{GAGAGGAG}_{CUCGACUC}$ at 15 °C, (d) 2.6 mM $^{GAGUUGAG}_{CUCUUCUC}$ at 14 °C, (e) 0.8 mM $^{GAGUAGAG}_{CUCUGCUC}$ at 1 °C, (f) 2.2 mM $^{GAGGUGAG}_{CUCAUCUC}$ at 10 °C, (g) 1.3 mM $^{GAGAUGAG}_{CUCGUCUC}$ at 8 °C, (h) 0.9 mM $^{GAGGUGAG}_{CUCUUCUC}$ at 0 °C, (i) 0.9 mM $^{GAGUUGAG}_{CUCUGCUC}$ at 3 °C, (j) 1.2 mM $^{GAGUGGAG}_{CUCUUCUC}$ at 6 °C, (k) 1.7 mM $^{GAGUUGAG}_{CUCGUCUC}$ at 6 °C, and (l) 0.9 mM $^{GAGUAGAG}_{CUCGGCUC}$ at 10 °C. Concentrations are total strand concentration.

from G in the A·G mismatch and the other from one of the two Us. From geometric considerations (see Discussion), the A·G mismatch cannot form a sheared conformation if there is a Watson−Crick base pair 5′ to the adenine, so the two resonances are probably from the two Us. Evidently, the imino proton of the A·G mismatch is rapidly exchanging with water. Thus, it does not give rise to a resonance.

Four spectra for combinations of U·G and U·U mismatches have similar features (Figure 2, spectra h−k). They



(I) Imino Hydrogen-bonded Conformation



(II) Sheared Conformation

FIGURE 3: Two known hydrogen-bonding patterns of G·A mismatches in a G·A tandem. (I) Imino hydrogen bonded conformation, (II) Sheared conformation.

all have four extra resonances between 11.8 and 9.6 ppm (with two resonances overlapping at about 10.4 ppm in spectrum i). Since both U·G and U·U mismatches have two imino protons, these four resonances are presumably from the mismatches. Imino protons of G in U·G wobble pairs typically give sharp resonances around 11 ppm (McDowell & Turner, 1996; McDowell et al., 1997; White et al., 1992; Allain & Varani, 1995; Geerdes & Hilbers, 1979). Although it is difficult to assign the four resonances individually, the broadness suggests that neither of the two mismatches forms strong hydrogen bonds such as those observed in symmetric tandem U·G and U·U mismatches (SantaLucia et al., 1991b; He et al., 1991; McDowell & Turner, 1996; McDowell et al., 1997; Nikonowicz & Pardi, 1992).

The spectrum for $^{5'GAGUAGAG3'}_{3'CUCGGCUC5'}$ (Figure 2l) has only two extra resonances, at 11.71 and 10.43 ppm. Neither of them gives an NOE to an AH2 (data not shown); thus, the A·G mismatch is not in an imino hydrogen-bonded conformation. Also for geometric reason, it cannot form a sheared conformation, either. Thus, these two resonances are assigned to the U·G mismatch. As in the motif $^{5'GAGAUGAG3'}_{3'CUCGUCUC5'}$ (Figure 2g), the lack of an imino proton resonance from G in the A·G mismatch indicates that the G imino proton is exchanging rapidly with water.

DISCUSSION

Studies on symmetric tandem mismatches showed that some mismatches are stabilizing, e.g., U·G, G·U, G·A, A·G, and U·U, partly due to hydrogen bonding (Wu et al., 1995). Other mismatches are destabilizing, e.g., A·A, A·C, C·A, C·C, C·U, and U·C. A periodicity of stability as a function of mismatches and closing base pairs has been summarized in a "periodic table" of symmetric tandem mismatches (Wu et al., 1995). For predicting RNA secondary structure from sequence, it is important to know if mismatches that are stabilizing in symmetric contexts are still stabilizing in

Table 4: Table of Δ Values for Nonsymmetric Tandem Mismatches[a]

GXYG / CWZC	G / U	U / G	A / G	G / A	U / U	A / C	U / C	C / U	C / C	C / A	A / A
U / G	0.3	0.1	2.5	2.3	1.8	*(0.0)*	*(1.0)*	*(1.0)*	1.4	0.9	2.0
G / U	*(0.0)*	0.3	1.4	*(2.0)*	2.5	*(0.0)*	*(1.0)*	1.4	*(1.0)*	0.7	1.1
G / A	*(2.0)*	2.3	0.2	*(0.0)*	2.3	*(0.0)*	*(1.0)*	*(1.0)*	1.4	0.7	0.5
A / G	2.3	1.9	0.6	0.6	2.4	0.9	*(1.0)*	0.6	0.7	0.3	0.5
U / U	1.6	1.5	1.7	1.8	0.0	−0.3	−0.1	*(0.0)*	−0.5	−0.5	1.2
C / A	*(0.0)*	*(0.0)*	−0.3	−0.5	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*
C / U	*(1.0)*	*(1.0)*	*(1.0)*	*(1.0)*	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*
U / C	*(1.0)*	*(1.0)*	*(1.0)*	*(1.0)*	0.0	*(0.0)*	0.5	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*
C / C	*(1.0)*	*(1.0)*	*(1.0)*	−0.5	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*
C / A	*(0.0)*	*(0.0)*	0.7	−0.8	0.2	*(0.0)*	*(0.0)*	*(0.0)*	0.8	−0.2	*(0.0)*
A / A	1.5	0.8	−0.4	−0.1	1.0	*(0.0)*	*(0.0)*	*(0.0)*	0.8	−0.9	−0.8

[a] The left column represents the $^{5'X}_{3'W}$ mismatches, while the top row represents the $^{Y3'}_{Z5'}$ mismatches in the context of $^{5'GX\,YG3'}_{3'CWZC5'}$. Values in italics and parentheses are predicted. For tandem mismatches of similar size or of two destabilizing mismatches or at least one A·C mismatch, penalty is predicted to be zero (see Discussion); penalty is 2.0 kcal/mol for two different stabilizing mismatches, and 1.0 kcal/mol for one stabilizing and one non-A·C destabilizing mismatch of different size.

nonsymmetric contexts and if the rules revealed by the periodic table of symmetric tandem mismatches can be used to predict stabilities of nonsymmetric tandem mismatches.

The simplest approximation for predicting stabilities of nonsymmetric tandem mismatches is to average the appropriate values for symmetric tandem mismatches. This is shown below:

$$\Delta G^\circ_{\text{approx}}\left(^{5'GX\,YG3'}_{3'CWZC5'}\right) = \frac{1}{2}[\Delta G^\circ\left(^{5'GX\,WC3'}_{3'CWX\,G5'}\right) + \Delta G^\circ\left(^{5'CZ\,YG3'}_{3'GYZC5'}\right)] \quad (6)$$

This is a nearest-neighbor approximation. It assumes that for each motif of tandem mismatches, there are three components corresponding to three nearest-neighbor interactions (Turner et al., 1988), and that the nearest-neighbor interactions between two mismatches are the average of the symmetric counterparts. The difference (Δ) between experimental values and approximations is defined as:

$$\Delta = \Delta G^\circ_{\text{exp}} - \Delta G^\circ_{\text{approx}} \quad (7)$$

For example, to calculate Δ for the motif $^{5'GGUG3'}_{3'CAUC5'}$, we take the average of $\Delta G^\circ_{37}\left(^{5'GGAC3'}_{3'CAGG5'}\right)$ (−2.7 kcal/mol) and $\Delta G^\circ_{37}\left(^{5'CUUG3'}_{3'GUUC5'}\right)$ (−0.1 kcal/mol) by eq 6, giving −1.4 kcal/mol; the experimental value of $\Delta G^\circ_{37}\left(^{5'GGUG3'}_{3'CAUC5'}\right)$ is 0.85 kcal/mol (Table 3). Thus, the Δ value by eq 7 is 2.3 kcal/mol for this motif. Values of Δ for each motif are listed in Table 4. Positive Δ values indicate that the experimentally measured stability increments are less favorable than predicted by eq 6. Negative Δ values indicate that the measured stability increments are more favorable than predicted. Numbers in italics and parentheses in Table 4 are predicted by the rules specified at the end of this paper.

The mismatches in Table 4 are ordered such that they are less favorable in a symmetric context (Wu et al., 1995) as you go from left to right or top to bottom. The upper left corner, from U·G to U·U, contains stabilizing mismatches. Here, Δ values range from 0.0 to 2.5 kcal/mol. The Δ values less than 0.6 kcal/mol are associated with two identical mismatches, e.g., two U·G mismatches, $^{5'GUUG3'}_{3'CGGC5'}$ (Δ is 0.1 kcal/mol). Thus, Δ is close to zero if both mismatches are stabilizing and have the same nucleotide composition. The Δ values greater than 1.4 kcal/mol are associated with combinations of stabilizing mismatches of different size, e.g., $^{5'GUAG3'}_{3'CUGC5'}$ (Δ is 1.7 kcal/mol). G·A, U·U, and G·U mismatches in symmetric contexts are known to have hydrogen bonds between the mismatched bases (SantaLucia et al., 1991b; SantaLucia & Turner, 1993; Nikonowicz & Pardi, 1993; Walter et al., 1994; Wu et al., 1995; McDowell & Turner, 1996; McDowell et al., 1997). Presumably, hydrogen-bonded mismatches are relatively rigid with certain sizes and shapes. This suggests that backbone distortions required to accommodate different sizes of hydrogen-bonded mismatches may be an important factor in determining stability.

The lower right corner of Table 4 contains combinations of two destabilizing mismatches. Since these mismatches are not hydrogen bonded in symmetric contexts, they are likely flexible. Therefore, it is reasonable to see small Δ values for these motifs, e.g., $^{5'GACG3'}_{3'CCAC5'}$ (Δ is −0.2 kcal/mol), $^{5'GACG3'}_{3'CCCC5'}$ (Δ is 0.8 kcal/mol). The Δ values in Table 4 are also near zero for combinations of one stabilizing mismatch and one destabilizing mismatch of the same size, e.g., $^{5'GA\ 3'CG}_{AG3'\ AC5'}$ (Δ is 0.5 kcal/mol), and for combinations of A·C with G·A or U·U. The average Δ for such sequences is less than 0.1 kcal/mol. For combinations of one stabilizing mismatch and one non-A·C destabilizing mismatch of different size, however, the average of the Δ values is 1.0 kcal/mol, e.g., $^{5'GGCG3'}_{3'CACC5'}$ (Δ is 1.4 kcal/mol). It is clear from Table 4 that the difference in size of the two neighboring mismatches and the type of mismatch both play a role in determining stability.

Following the nearest-neighbor model, the free energy increment of tandem mismatches can be broken down into

three components. These three terms for any motif of $^{5'GX\,YG3'}_{3'CWZC5'}$ are $^{5'GX\,3'}_{3'CW5'}$, $^{5'X\,Y3'}_{3'WZ5'}$, and $^{5'YG3'}_{3'ZC5'}$. Thus, the Δ values of Table 4 represent the error in approximating these nearest neighbors. There are many potential sources of this error. First of all, the interactions between closing base pair and mismatch, e.g., $^{5'GX\,3'}_{3'CW5'}$, in a nonsymmetric motif are not necessarily the same as that in a symmetric $^{5'GX\,WC3'}_{3'CWX\,G5'}$ motif. Secondly, interactions between two mismatches, e.g., $^{5'X\,Y3'}_{3'WZ5'}$, will not in general be the average of $^{5'X\,W3'}_{3'WX\,5'}$ and $^{5'Z\,Y\,3'}_{3'YZ\,5'}$. In addition, there can be non-nearest-neighbor effects which will be discussed later. All these are reflected in Δ values.

One factor that affects Δ is mismatch structure. That is, the structure of the mismatch may be the same or different in the symmetric and nonsymmetric contexts. NMR spectroscopy provides information on structures. The stability of $^{5'GUUG3'}_{3'CUUC5'}$ is predicted well by eq 6 (Δ = 0.0 kcal/mol). The NMR spectrum for this motif (Figure 2d) shows that both U·U mismatches are imino hydrogen bonded. Chemical shifts of the four U imino protons are very close to those in the symmetric motifs $^{5'GUUC3'}_{3'CUUG5'}$ (Wu et al., 1995) and $^{5'CUUG3'}_{3'GUUC5'}$ (SantaLucia et al., 1991b). This suggests that each U·U mismatch has the same structure as in the symmetric context. Equation 6 also predicts the stability of $^{5'GAGG3'}_{3'CGAC5'}$ well (Δ = 0.6 kcal/mol). The imino-proton spectrum for $^{5'GAGG3'}_{3'CGAC5'}$ shows that both the A·G mismatches are imino hydrogen bonded (Figure 2c). The symmetric tandem A·G mismatches in $^{5'GAGC3'}_{3'CGAG5'}$ and $^{5'CAGG3'}_{3'GGAC5'}$ are also imino hydrogen bonded (SantaLucia et al., 1991b; Walter et al., 1994; Wu et al., 1997). Thus, there is little structural change for the two A·G mismatches going into the nonsymmetric context. For the above two motifs, the interaction between mismatch and closing base pair is the same as in the symmetric motifs. Since both U·U and A·G mismatches have the same structure as they do in the symmetric motifs and the sizes of the two mismatches in both motifs are the same, the interactions between the mismatches in both motifs should be predicted well by the average of those in the symmetric cases. Thus, Δ is close to 0 kcal/mol.

In $^{5'GAAG3'}_{3'CGGC5'}$, both mismatches are also imino hydrogen bonded. Although the A·G mismatches in the symmetric $^{5'GAGC3'}_{3'CGAG5'}$ motif are imino hydrogen bonded (Walter et al., 1994), the G·A mismatches in the symmetric $^{5'CGAG3'}_{3'GAGC5'}$ motif have the sheared conformation shown in Figure 3 (SantaLucia & Turner, 1993). So the right G·A mismatch gives up the sheared conformation that is favorable for the corresponding symmetric context. This is because a sheared conformation and an imino hydrogen-bonded conformation cannot be incorporated next to each other in a helix due to their remarkably different backbone distortions (Gautheret et al., 1994; SantaLucia & Turner, 1993; Wu & Turner, 1996). Furthermore, due to the short P−P distance of a sheared A·G mismatch (12.7 Å), a Watson−Crick base pair (P−P distance is 17.5 Å in a regular A-RNA) cannot be connected to the 5′ side of the adenine in a sheared A·G mismatch (Gautheret et al., 1994). As a result, the A·G mismatches of the $^{AA}_{GG}$ motif cannot both adopt sheared conformations. Interestingly, the Δ value for this motif is still only 0.6 kcal/mol, possibly because better stacking between mismatches partially compensates the unfavorable electrostatic interactions between an imino hydrogen-bonded G·A mismatch and C·G closure (Wu & Turner, 1996).

The mismatches in the $^{5'GGUG3'}_{3'CAUC5'}$ motif also differ in structures from their symmetric counterparts. In this motif, the G·A mismatch switches from the imino hydrogen-bonded conformation that is favorable in the symmetric context to the sheared conformation. This is because the backbone of an imino hydrogen-bonded G·A mismatch (P-P distance 20.4 Å) is wider than the standard Watson−Crick backbone (Wu & Turner, 1996). This makes the adjacent smaller U·U mismatch unable to hydrogen bond. When the G·A mismatch is sheared, the U·U mismatch is on the narrower side of the G·A mismatch, i.e., 5′ to the adenine. Even though the G·A mismatch switches to a narrower sheared conformation, the U·U mismatch is still not able to tightly hydrogen bond, as evidenced by broad resonances for the U imino protons (Figure 2f). This is also seen in $^{5'GUAG3'}_{3'CUGC5'}$, where the G·A mismatch keeps the sheared conformation which is favorable in a C·G closing context. In the related motif, $^{5'GAUG3'}_{3'CGUC5'}$, the 1-D imino proton NMR spectrum shows that the A·G mismatch is not imino hydrogen bonded. But since there is a Watson−Crick G·C base pair 5′ to the adenine, this A·G mismatch cannot have a sheared conformation, either. The NMR spectrum also shows no evidence of a sheared conformation. Thus the A·G mismatch in this motif has a conformation other than the two structures shown in Figure 3. Due to severe spectral overlap, it is difficult to determine the structures of mismatches in another interesting and related motif, $^{5'GUGG3'}_{3'CUAC5'}$ (data not shown). In all of the above three motifs, resonances of U imino protons are broad compared to those in $^{5'GUUG3'}_{3'CUUC5'}$, indicating that they exchange faster with water. This is also observed for combinations of U·G and U·U mismatches where neither U·G nor U·U form strong hydrogen bonds. These are typical examples of size and shape complementarity of the two adjacent mismatches being important. Large unfavorable Δ values, from 1.5 to 2.5 kcal/mol, result when G·U, G·A, and U·U mismatches of different size are combined.

The phenomenon of mismatches affecting each other's structure produces a non-nearest-neighbor sequence effect. Studies on symmetric tandem G·A mismatches in RNA showed that the conformation of the mismatches is determined by the closing Watson−Crick base pairs (SantaLucia & Turner, 1993; Walter et al., 1994; Wu & Turner, 1996). The results presented here indicate that in nonsymmetric contexts, the structure of a mismatch is determined by both its closing base pair, and by the structure of its neighbor mismatch. Since the latter is partially determined by the other closing base pair, a Watson−Crick base pair can affect the structure of a mismatch which is not adjacent. Thus, the global minimum in energy depends on both mismatches and their closing base pairs.

Current RNA folding algorithms (Walter et al., 1994) assume that only the first mismatch on each side of an internal loop contributes to the stability, and this free energy increment is independent of the rest of the loop. This assumption is not reasonable for small loops such as tandem mismatches. As the size of an internal loop gets larger, however, the approximation of Walter et al. (1994) may become more reasonable since larger loops are likely to be more flexible and the first mismatch on each side would affect each other less.

The structural changes in base pair pattern are results of molecules seeking the global minimum in free energy. These changes in structure suggest that the nearest-neighbor interactions between hydrogen-bonded mismatches are significant so that eq 6 is not always a reasonable approximation. Presumably, hydrogen-bonded mismatches are rigid. If two adjacent rigid mismatches are different in size and shape, they can cause local backbone distortion or less favorable hydrogen bonding. Either distortion would make duplex formation less favorable. Thus, Δ values for these motifs are large positive numbers.

On the basis of the above ideas, it is reasonable to think that if one of the two mismatches is somehow flexible, it will give freedom to the other one. Presumably, mismatches that are destabilizing in a symmetric context are flexible. They are destabilizing because there is no particular interaction within the mismatch, such as hydrogen bonding. This is consistent with NMR imino proton spectra (SantaLucia et al., 1991b; Peritz et al., 1991; Wu et al., 1995). Table 5 shows how flexibility influences the Δ values. This table shows the effect of changing a U·G to C·A, a G·A to A·A, or a U·U to C·C or U·C in a tandem mismatch where the other mismatch is stabilizing (U·G, G·A, or U·U). That is, one of the two differently sized hydrogen-bonded mismatches is changed to a non-hydrogen-bonded mismatch of about the same size. The comparisons in Table 5 show that replacing a stabilizing mismatch by a destabilizing mismatch while keeping the other stabilizing mismatch unchanged reduces the penalty term by an average of 1.3 kcal/mol. The specific reduction depends on the particular mismatch replaced and the neighboring mismatch. For example, changing the A·G or U·U mismatch in $^{5'GAUG3'}_{3'CGUC5'}$ (Δ = 2.4 kcal/mol) to A·A ($^{5'GAUG3'}_{3'CAUC5'}$) or C·C ($^{5'GACG3'}_{3'CGCC5'}$) reduces Δ to 1.0 and 0.7 kcal/mol, respectively; changing the U·G in $^{5'GUGG3'}_{3'CUUC5'}$ (Δ = 1.6 kcal/mol) to a C·A ($^{5'GUAG3'}_{3'CUCC5'}$) reduces the Δ to −0.3 kcal/mol.

Presumably, nature selects particular sequences for functional reasons. Thermodynamic stability and structure are two correlated factors that influence function. Since RNA internal loops may play important roles in tertiary interactions (Michel & Westhof, 1990) and in protein recognition (Zwieb, 1992), this work represents a start on decoding the rules of molecular recognition for tandem mismatches in RNA. A survey of tandem mismatches in known RNA secondary structures suggests that tandem mismatches were not selected randomly during evolution. For all possible tandem mismatch patterns including adjacent G·C, A·U, or G·U base pairs, $^{5'RR3'}_{3'RR5'}$, $^{5'YY3'}_{3'YY5'}$, and $^{5'RY3'}_{3'RY5'}$ have sixteen possible combinations (R = purine, Y = pyrimidine), while $^{5'RR3'}_{3'RY5'}$, $^{5'RR3'}_{3'YR5'}$, and $^{5'RY3'}_{3'YY5'}$ have eight, and $^{5'RR3'}_{3'YY5'}$, $^{5'RY3'}_{3'YR5'}$, and $^{5'YR3'}_{3'RY5'}$ have four possible combinations, respectively. Table 1 lists the frequency of occurrences of tandem mismatches in 398 RNA secondary structures determined phylogenetically (Damberger & Gutell, 1994; Gutell et al., 1993; Gutell, 1994). Except for tandem G·U mismatches, all purine ($^{5'RR3'}_{3'RR5'}$) tandem mismatches dominate, occurring 443 times, or about 28 times per motif on average. All pyrimidine ($^{5'YY3'}_{3'YY5'}$) tandem mismatches are the second most prevalent with 148 occurrences (99 of them are tandem U·U mismatches), nine times per possible motif on average. Other tandem mismatches occur less frequently, a total of 175 occurrences, and most of them have at least one destabilizing mismatch such as A·C or A·A. Although it is clear that natural occurrences are not randomly distributed among different categories nor within each category, they do not necessarily follow the same trends as the thermodynamics.

Table 5: Influence of Flexibility of Mismatches on Δ Values

loop with rigid m.m.[a]	Δ[b] (kcal/mol)	loop with flexible m.m.[a]	Δ[b] (kcal/mol)	difference[c] (kcal/mol)
GAUG	2.4	GAUG	1.0	1.4
CGUC		CAUC		
GGUG	2.3	GAUG	1.0	1.3
CAUC		CAUC		
GAUG	1.9	GAUG	0.8	1.1
CGGC		CAGC		
GGUG	2.3	GAUG	0.8	1.5
CAGC		CAGC		
GAGG	2.3	GAGG	1.3	1.0
CGUC		CAUC		
GUAG	1.7	GUAG	1.2	0.5
CUGC		CUAC		
GUGG	1.8	GUAG	1.2	0.6
CUAC		CUAC		
GGAG	1.4	GGAG	1.1	0.3
CUGC		CUAC		
GUAG	2.5	GUAG	2.0	0.5
CGGC		CGAC		
GUGG	2.3	GUAG	2.0	0.3
CGAC		CGAC		
GUUG	1.8	GCUG	0.4	1.4
CGUC		CAUC		
GGUG	2.5	GAUG	0.2	2.3
CUUC		CCUC		
GUAG	2.1	GCAG	-0.3	2.4
CGGC		CAGC		
GGAG	1.4	GAAG	0.7	0.7
CUGC		CCGC		
GUGG	2.3	GCGG	-0.5	2.8
CGAC		CAAC		
GUUG	1.5	GUCG	-0.5	2.0
CUGC		CUAC		
GUGG	1.6	GUAG	-0.3	1.9
CUUC		CUCC		
GAUG	1.9	GACG	0.2	1.7
CGGC		CGAC		
GAGG	2.3	GAAG	0.9	1.4
CGUC		CGCC		
GGUG	2.3	GGCG	0.7	1.6
CAGC		CAAC		
GUGG	1.8	GCGG	-0.5	2.3
CUAC		CCAC		
GAUG	2.4	GACG	0.7	1.7
CGUC		CGCC		
GAUG	2.4	GAUG	0.6	1.8
CGUC		CGCC		
GGUG	2.3	GGCG	1.4	0.9
CAUC		CACC		
GUUG	1.8	GUCG	1.4	0.4
CGUC		CGCC		
GGUG	2.5	GGCG	1.4	1.1
CUUC		CUUC		

[a] m.m.= mismatch. [b] Data from Table 4. [c] Differences between the second and the fourth column.

Certain motifs occur more often than expected from thermodynamic stability. For example, the $^{5'AU3'}_{3'GU5'}$ motif occurs 14 times. This motif involves two extremely differently sized stabilizing mismatches and has an average unfavorable Δ of 2.1 kcal/mol and an average unfavorable loop free energy increment of 1.5 kcal/mol. The imino proton NMR spectrum (Figure 2g) indicates that, in the $^{5'GAUG3'}_{3'CGUC5'}$ motif, the U·U mismatch is loosely hydrogen bonded and the A·G mismatch is neither imino hydrogen bonded nor sheared. Thus, in the conformation of this motif,

several hydrogen-bonding groups of the A·G mismatch may be available for interactions. The U·U mismatch presumably has free carbonyl groups also available for hydrogen bonding. Thus, this thermodynamically unfavorable motif may have been selected during evolution because of a unique arrangement of hydrogen-bonding groups required for tertiary folding or molecular recognition. A phylogenetic search shows that this motif often occurs at the same position in large subunit ribosomal RNAs indicating a possible functional role.

From these and previous results, we propose an improved but still simple and preliminary model to predict stabilities of all 2628 possible tandem mismatches (including U·G as closing base pair). In this model, we add to eq 6 a penalty term, whose value depends on the type of both mismatches and closing base pairs. The general equation for predicting stability of the tandem mismatch motif $^{5'PXYS3'}_{3'QWZT5'}$ is

$$\Delta G^\circ_{pred}(^{5'PXYS3'}_{3'QWZT5'}) = \frac{1}{2}[\Delta G^\circ_{37}(^{5'PXWQ3'}_{3'QWXP5'}) + \Delta G^\circ_{37}(^{5'TZYS3'}_{3'SYZT5'})] + \Delta_P \quad (8)$$

where P·Q and S·T represent closing base pairs, X·W and Z·Y represent mismatches. The Δ_P is a penalty term preliminarily approximated as described below.

For identically sized tandem mismatches or combinations involving two destabilizing mismatches, or at least one A·C mismatch, e.g., $^{GG}_{AA}$, $^{AA}_{GC}$, $^{AA}_{AC}$, etc., Δ_P is set to zero regardless of closing base pairs. For any permutations of two G·C/C·G closing base pairs, with two differently sized stabilizing mismatches (i.e., G·U, G·A, or U·U), e.g., $^{5'UA3'}_{3'UG5'}$, Δ_P is set to 2 kcal/mol; with combinations of one stabilizing mismatch and one non-A·C destabilizing mismatch of different size, Δ_P is set to 1 kcal/mol. In Table 4, Δ values in italics and parentheses are calculated by this model. Motifs of two differently sized stabilizing mismatches are penalized by 1.5 kcal/mol for combinations of one G·C and one A·U or G·U closing base pair, and by 1.0 kcal/mol when both closing base pairs are either A·U or G·U. A·U closing base pairs are penalized less because A·U pairs are probably more flexible than G·C pairs. For combinations of one stabilizing mismatch and one non-A·C destabilizing mismatch of different size, penalties are set to 0.5 and 0.0 kcal/mol for G·C/A·U(G·U) and A·U/A·U(G·U) closing base pair patterns, respectively. This type of model has been added to the RNA folding algorithm, MFOLD (Zuker, 1989), and preliminary results indicate it improves structure prediction (Mathews et al., 1997).

NOTE ADDED IN PROOF

The thermodynamics of duplex formation have now been measured for (GGCAGGCC)$_2$, for which an NMR structure is available (Wu et al., 1997). With the 1 M NaCl, 10 mM sodium cacodylate buffer of this paper, values from T_M^{-1} plots and from averages of fits (in parentheses) are $\Delta G^\circ_{37} = -9.4 \pm 0.1$ (-9.6 ± 0.1) kcal/mol, $\Delta H^\circ = -69.8 \pm 1.9$ (-73.5 ± 1.5) kcal/mol, and $\Delta S^\circ = -194.6 \pm 5.7$ (-205.9 ± 4.7) eu. In a buffer of 150 mM KCl, 10 mM $MgCl_2$, and 10 mM sodium cacodylate, pH 7.0, the values are $\Delta G^\circ_{37} = -9.7 \pm 0.1$ (-9.5 ± 0.2) kcal/mol, $\Delta H^\circ = -78.6 \pm 2.8$ (-73.9 ± 3.5) kcal/mol, $\Delta S^\circ = -222.1 \pm 8.5$ (-207.6 ± 10.7) eu. For 1 M NaCl, the $\Delta G^\circ_{loop,37} = -0.1$ kcal/mol. This is similar to the $\Delta G^\circ_{loop,37}$ of -0.6 kcal/mol previously reported for (CGCAGGCG)$_2$ (SantaLucia et al., 1991b), supporting the assumptions in the model reported here.

ACKNOWLEDGMENT

We thank Dr. Xiaoqi Jiao for help with NMR experiments and Erik M. Roche for help with the phylogenetic analysis search.

SUPPORTING INFORMATION AVAILABLE

Eight figures showing plots of inverse melting temperature vs ln(C_T/4) for 47 sequences (8 pages). Ordering information is given on any current masthead page.

REFERENCES

Albergo, D. D., & Turner, D. H. (1981) *Biochemistry 20*, 1413−1418.
Allain, F. H.-T., & Varani, G. (1995) *Nucleic Acids Res. 23*, 341−350.
Banerjee, A. R., Jaeger, J. A., & Turner, D. H. (1993) *Biochemistry 32*, 153−163.
Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York, p 59−62.
Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 597, CRC Press, Cleveland OH.
Borer, P. N., Dengler, B., & Tinoco, I., Jr. (1974) *J. Mol. Biol. 86*, 843−853.
Cheng, J.-W., Chou, S.-H., & Reid, B. R. (1992) *J. Mol. Biol. 228*, 1037−1041.
Chou, S.-H., Hare, D. R., Wemmer, D. E., & Reid, B. R. (1983) *Biochemistry 22*, 3037−3041.
Crothers, D. M., Cole, P. E., Hilbers, C. W., & Shulman, R. G. (1974) *J. Mol. Biol. 87*, 63−88.
Damberger, S. H., & Gutell, R. (1994) *Nucleic Acids Res. 22*, 3508−3510.
Freier, S. M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M. H., & Turner, D. H. (1986) *Biochemistry 25*, 3214−3219.
Gautheret, D., Konings, D., & Gutell, R. R. (1994) *J. Mol. Biol. 242*, 1−8.
Geerdes, H. A. M., & Hilbers, C. W. (1979) *FEBS Lett. 107*, 125−128.
Gluick, T. C., & Draper, D. (1994) *J. Mol. Biol. 241*, 246−262.
Gralla, J., & Crothers, D. M. (1973a) *J. Mol. Biol. 78*, 301−319.
Gralla, J., & Crothers, D. M. (1973b) *J. Mol. Biol. 73*, 497−511.
Green, K. L., Jones, R. L., Li, Y., Robinson, H., Wang, H.-J., Zon, G., & Wilson, W. D. (1994) *Biochemistry 33*, 1053−1062.
Gregory, R. J., Cahill, P. B. F., Thurlow, O. L., & Zimmermann, R. A. (1988) *J. Mol. Biol. 204*, 295−307.
Gutell, R. R. (1994) *Nucleic Acids Res. 22*, 3502−3507.
Gutell, R. R., Gray, M. W., & Schnare, M. N. (1993) *Nucleic Acids Res. 21*, 3055−3074.
He, L., Kierzek, R., SantaLucia, J., Jr., Walter, A. E., & Turner, D. H. (1991) *Biochemistry 30*, 11124−11132.
Hilbers, C. W. (1979) in *Biological Applications of Magnetic Resonance* (Shulman, R. G., Ed.) Academic Press, New York.
Hilbers, C. W., Robillard, G. T., Shulman, R. G., Blake, R. D., Webb, P. K., Fresco, R., & Riesner, D. (1976) *Biochemistry 15*, 1874−1882.
Hore, P. J. (1983) *J. Magn. Reson. 55*, 283−300.
Jaeger, J. A., Turner, D. H., & Zuker, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7706−7710.
Jaeger, L., Westhof, E., & Michel, F. (1993) *J. Mol. Biol. 234*, 331−346.
Li, Y., Zon, G., & Wilson, W. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 26−30.
Lück, R., Steger, G., & Riesner, D. (1996) *J. Mol. Biol. 258*, 813−826.

Mathews, D. H., Andre, T. C., Kim, J., Turner, D. H., & Zuker, M. (1997) in *Molecular Modeling of Nucleic Acids* (Leontis, N. B., & Santa Lucia, J., Jr., Eds.) American Chemical Society (in press).
McCaskill, J. S. (1990) *Biopolymers 29*, 1105−1119.
McDowell, J. A., & Turner, D. H. (1996) *Biochemistry 35*, 14077−14089.
McDowell, J. A., He, L., Chen, X., & Turner, D. H. (1997) *Biochemistry 36*, 8030−8038.
Meyer, S. L. (1975) *Data Analysis for Scientists and Engineers*, Chapter 14, Wiley, New York.
Michel, F., & Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.
Nikonowicz, E. P., & Pardi, A. (1992) *Nature 335*, 184−186.
Nikonowicz, E. P., & Pardi, A. (1993) *J. Mol. Biol. 232*, 1141−1156.
Pardi, A. Martin, F. H., & Tinoco, I., Jr. (1981) *Biochemistry 20*, 3986−3996.
Peritz, A. E., Kierzek, R., Sugimoto, N., & Turner, D. H. (1991) *Biochemistry 30*, 6428−6436.
Petersheim, M., & Turner, D. H. (1983) *Biochemistry 22*, 256−263.
Richards, E. G. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 597, CRC Press, Cleveland OH.
SantaLucia, J., Jr., & Turner, D. H. (1993) *Biochemistry 32*, 12612−12623.
SantaLucia, J., Jr., Kierzek, R., & Turner, D. H. (1991a) *J. Am. Chem. Soc. 113*, 4313−4322.
SantaLucia, J., Jr., Kierzek, R., & Turner, D. H. (1991b) *Biochemistry 30*, 8242−8251.
Serra, M. J., & Turner, D. H. (1995) *Methods Enzymol. 259*, 242−261.
Snedecor, G. W., & Cochran, W. G. (1982) in *Statistical Methods*, 7th ed., p 189, The Iowa State University Press, Ames, IA.
Tinoco, I., Jr., Uhlenbeck, O. C., & Levine, M. D. (1971) *Nature 230*, 363−67.
Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., & Gralla, J. (1973) *Nat. New Biol. 246*, 40−41.
Turner, D. H., Freier, S. M., & Sugimoto, N. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−92.
Ulrich, E. L., John, E.-M. M., Gough, G. R., Brunden, M. J., Gilham, P. T., Westler, W. M., & Markley, J. L. (1983) *Biochemistry 22*, 4362−4365.
Usman, N., Ogilvie, K. K., Jiang, M.-V., & Cedergren, R. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.
Varani, G., Wimberly, B., & Tinoco, I., Jr. (1989) *Biochemistry 28*, 7760−7772.
Walter, A. E., Wu, M., & Turner, D. H. (1994) *Biochemistry 33*, 11349−11354.
Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A., & Weiner, A. M. (1987) *Molecular Biology of the Gene*, Benjamin Cummings, Inc., Menlo Park, CA.
White, S. A., Nilges, M., Huang, A., Brünger, A. T., & Moore, P. B. (1992) *Biochemsitry 31*, 1610−1621.
Wu, M., & Turner, D. H. (1996) *Biochemistry 35*, 9677−9689.
Wu, M., McDowell, J., & Turner, D. H. (1995) *Biochemistry 34*, 3204−3211.
Wu, M., SantaLucia, J., Jr., & Turner, D. H. (1997) *Biochemistry 36*, 4449−4460.
Zuker, M. (1989) *Science 244*, 48−52.
Zuker. M., & Stiegler, P. (1981) *Nucleic Acids Res. 9*, 133−148.
Zwieb, C. (1992) *J. Biol. Chem. 267*, 15650−15656.

BI971069V