Thermodynamics of Coaxially Stacked Helixes with GA and CC Mismatches†

James Kim, Amy E. Walter, and Douglas H. Turner*

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

Received April 15, 1996; Revised Manuscript Received August 12, 1996⊗

ABSTRACT: The thermodynamics of RNA helix−helix interfaces with intervening single- and tandem-GA or single-CC mismatches were studied by UV melting experiments. The model system consists of a hairpin with a four- or five-nucleotide 5′ overhang which is bound by a short oligomer, creating the helical interface. Single GA interfaces are found to have favorable free energy increments of about 2 kcal/mol. This is similar to those reported for coaxially stacked flush interfaces of AU base pairs [Walter A. E., & Turner, D. H. (1994) *Biochemistry 33,* 12715−12719]. The free energy increment of the GA mismatches depends little on the sequence of the closing base pairs of the helixes, whether the break in the phosphate backbone is 5′ or 3′ with respect to the mismatch or whether the chains are extended beyond the helix−helix interface. Surprisingly, interfaces with single-CC mismatches have favorable free energy increments similar to those of GA interfaces, even though CC mismatches in coaxial stacks occur much less frequently in known RNA secondary structures. The results provide experimental support for the assumption that a bonus free energy is required for coaxially stacked helixes with intervening GA mismatches when free energy minimization is used to predict RNA secondary structures [Walter, A. E., Turner, D. H., Kim, J., Lyttle, M. H., Mueller, P., Mathews, D. H., & Zuker, M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91,* 9218−9222].

Much of the diversity of RNA functions (Watson et al., 1987) stems from the variety of structures that RNA can adopt. A detailed understanding of RNA structure and energetics is crucial for understanding the functions and mechanisms of RNA. Unfortunately, obtaining structural information has been difficult. Knowledge of RNA energetics, however, can facilitate acquisition of structural information in many ways. For example, it can help in deriving secondary structure from sequence.

Although RNA secondary structure prediction has vastly improved (Turner et al., 1988; Jaeger et al., 1989; Walter et al., 1994; Gultyaev et al., 1995; Schmitz & Steger, 1996), one secondary structural motif which continues to be poorly predicted is the multibranch loop (Zuker et al., 1991). This is understandable since the multibranch loop is the most poorly understood among the various structural motifs. One interaction that contributes to the stabilities of many multibranch loops is coaxial stacking (Walter et al., 1994; Walter & Turner, 1994). Presumably, coaxial stacking has a favorable free energy due to both stacking and hydrogen bonding interactions even though there is a break in the sugar-phosphate backbone. An example of the strength of such interactions in the complete absence of a backbone is the self-assembly of adenosine on poly U (Huang & Ts'o, 1966).

Coaxial stacking of helixes has been found in crystal structures of tRNA (Kim et al., 1974; Robertus et al., 1974; Westhof et al., 1985; Biou et al., 1994) and hammerhead ribozymes (Pley et al., 1994; Scott et al., 1995). It has also been seen in NMR model systems of group I introns (Chastain & Tinoco, 1992, 1993) and natural antisense regulators (Marino et al., 1995). Coaxial stacking has also been proposed in modeling studies of group I introns (Kim & Cech, 1987; Michel & Westhof, 1990) and rRNA (Stern et al., 1988; Gutell et al., 1994; Laing & Draper, 1994). In yeast phenylalanine tRNA, the D and anticodon arms are separated by a GA mismatch but coaxially stack with each other. Such interactions have also been proposed in models for rRNA (Stern et al., 1988; Laing & Draper, 1994). GA mismatches between helixes are also seen to occur in self-alkylating (Wilson & Szostak, 1995) and self-ligating (Ekland et al., 1995) ribozymes.

Model systems with a four- or five-nucleotide overhang on a hairpin stem have been shown to provide thermodynamic parameters for coaxially stacked helixes with Watson−Crick interfaces (Walter et al., 1994; Walter & Turner, 1994). This model system is extended here to study the thermodynamic effects of coaxial stacking of helixes with intervening GA or CC mismatches.

MATERIALS AND METHODS

Oligoribonucleotides. Oligonucleotides were synthesized in 1 μmol quantities with the β-cyanoethyl (CE) phosphoramidite method (Usman et al., 1987; Ogilvie et al., 1988; Wu et al., 1989; Wang et al., 1990) on an Applied Biosystems DNA/RNA synthesizer model 392. Sequences were cleaved from the CPG support and base blocking groups were removed by treatment with ethanolic ammonium hydroxide (3:1, v/v, 30% ammonium hydroxide/95% ethanol) for at least 16 h at 55 °C. Samples were incubated in 1 M triethylamino−hydrogen fluoride in pyridine for at least 48 h at 55 °C to remove the 2′-*tert*-butyldimethylsilyl protecting groups. Salts were then removed by a Sep Pak C-18 cartridge (Waters).

Short (tetramer, pentamer, and hexamer) oligonucleotides were then purified by thin-layer chromatography (Baker Si500F plates) in 55:35:10, v/v/v, 1-propanol/ammonia/water

† This work is supported by NIH Grants GM 22939 and GM 07102.
* Author to whom correspondence should be addressed.
⊗ Abstract published in *Advance ACS Abstracts,* October 15, 1996.

S0006-2960(96)00913-0 CCC: $12.00

Table 1: Phylogenetically Determined Number of Possible Coaxially Stacked Helixes with Only Watson−Crick or GU Base Pairs at the Interfaces or with One or Two Mismatches at the Interfaces

RNA[a]	no. of Watson−Crick and GU coaxial stacks[b]	no. of coaxial stacks with single mismatches[c]	no. of coaxial stacks with double mismatches[c]
small-subunit rRNA	179	222	46
large-subunit rRNA	312	316	394
tRNA	97	56	0
RNase MRP RNA	7	14	0
RNase P RNA	5	8	0
V4 of 18s	21	18	9
group I introns	10	6	0
sum	631	640	440

[a] 21 small-subunit rRNA (Gutell et al., 1985); 21 large-subunit rRNA (Gutell et al.,1992); 77 tRNA (Hall, R., 1971; Sprinzl et al., 1991); 6 RNase MRP RNA (Schmitt et al., 1993); 2 RNase P RNA (Pace et al., 1989); 9 V4 of 18s rRNA (Nickrent & Sargent, 1991); 3 group I intron RNA (Michel & Westhof, 1990). [b] Watson−Crick and GU helical interfaces. [c] Helical interfaces with single or double mismatches. GU wobbles were treated in the same manner as Watson−Crick pairs due to their stability.

Table 2: Percentage of Occurrences of Single Mismatches and the Double Mismatch $\begin{smallmatrix}5'GA\\3'AG\end{smallmatrix}$ between Helices[a]

mismatch[b]	% of occurrence	mismatch[b]	% of occurrence
GA	14	GG	4
AG	12	CU	2
AA	10	CC	1
UU	5	UC	1
CA	4		
AC	4	$\begin{smallmatrix}GA\\AG\end{smallmatrix}$	3

[a] For the data set given in Table 1. Only single and double mismatches other than GU and UG were searched. Presumably, anything larger would not allow coaxial stacking. [b] Mismatches, XY, are defined such that $\begin{smallmatrix}5'X\\3'Y\end{smallmatrix}$ occurs between helices. Therefore, a GA mismatch is $\begin{smallmatrix}5'G\\3'A\end{smallmatrix}$ and an AG mismatch is $\begin{smallmatrix}5'A\\3'G\end{smallmatrix}$.

Table 3: Percentage of Occurrences of Base Pairs 5′ and 3′ to the GA Mismatch between Helices[a]

base pair[b]	% of occurrence 5′ to GA	% of occurrence 3′ to GA
AU	12	15
UA	26	9
GC	14	15
CG	27	11
GU	3	2
UG	15	12

[a] For the data set given in Table 1. [b]Base pairs, XY, are defined such that $\begin{smallmatrix}X\\Y\end{smallmatrix}$ occur $5'\left(\begin{smallmatrix}5'XG\\3'YA\end{smallmatrix}\right)$ or $3'\left(\begin{smallmatrix}5'GX\\3'AY\end{smallmatrix}\right)$ to the GA mismatch.

(Chou et al., 1989). After elution of samples from the silica gel with sterile water, remaining salts were removed by a Sep Pak C-18 cartridge (Waters). Purity of these short oligomers was checked by HPLC on a C-8 RP column and found to be greater than 95% pure, except for 5′GGA-CAAA3′ (>90%), 5′GGACA3′ (>92%), and 5′GGUCC3′ (>92%).

Long hairpins were purified on 8 M urea/20% polyacrylamide gels. Bands were visualized under ultraviolet light, and the least mobile band was cut out. Hairpins were eluted from gel by a crush and soak procedure using 0.5 M

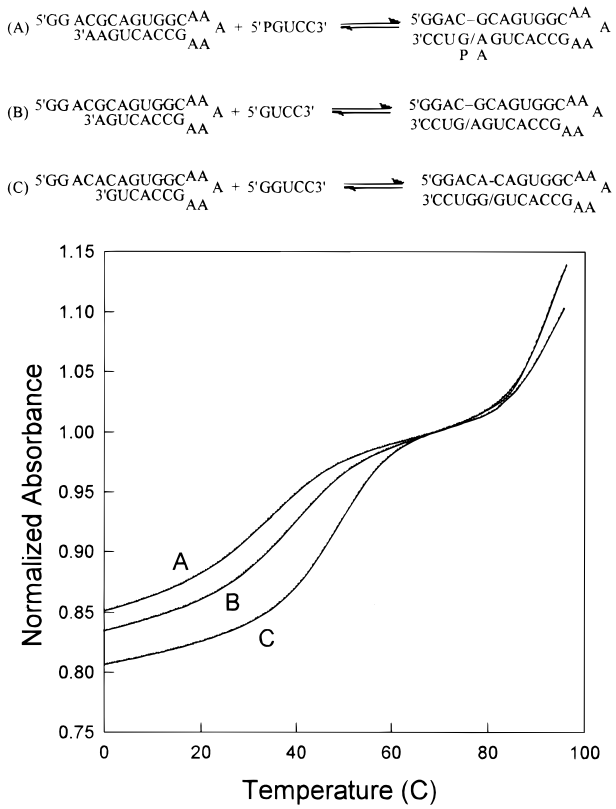

FIGURE 1: Typical melting curves. Absorbances were normalized by division with the absorbance at 69 °C. Sequences corresponding to the curves are shown above the plot. The lowest temperature transition of each curve represents the melting of the short oligomer from the hairpin. This transition is fit to obtain the thermodynamic parameters of coaxial stacking. The highest temperature transition is that of the hairpin. The strand concentrations and T_M values are (A) 1.97×10^{-4} M, 33.1 °C; (B) 1.87×10^{-4} M, 38.1 °C; (C) 1.91×10^{-4} M, 47.6 °C.

ammonium acetate, 0.1 mM EDTA. Hairpins were ethanol precipitated and further desalted with a G-25 Sephadex column or a Sep Pak C-18 cartridge (Waters). Purity was checked by 5′ ^{32}P-labeling with T4 polynucleotide kinase, and the hairpins were found to be at least 95% pure.

Measurement of Thermodynamic Data. Thermodynamic data were obtained by measuring absorbance at 280 nm vs temperature from 0 to 95 °C. The apparatus consisted of a Gilford 250 spectrometer and a Gilford 2527 thermoprogrammer interfaced to a Zenith Z-248 computer. A Gateway 4DX-33 computer was used to process the data. Oligonucleotides were dissolved in a melting buffer of 1.0 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na_2EDTA at pH 7.0. Sequences containing $\begin{smallmatrix}5'GCACC-C\\3'CGUGC/G\end{smallmatrix}$ and $\begin{smallmatrix}5'GCACC-C\\3'CGUGC/G\\A\ A\end{smallmatrix}$ were also melted in the same buffer adjusted to pH 5.5. The total strand concentrations were calculated from the absorbance measured at 95 °C and 280 nm (Borer, 1975; Richards, 1975).

Melting curves were analyzed by fitting to a two-state model with sloping base lines through the use of a nonlinear least-squares program (Petersheim & Turner, 1983). The thermodynamic parameters were obtained with two methods. One method used the average enthalpy and entropy changes from fitting the curves. The second method used plots in which the inverse melting temperatures, T_M^{-1}, vs the log of total strand concentrations, C_T, were fit to a straight line

Table 4: List of Sequences Studied[a]

sequence no.	sequence	sequence no.	sequence
(1)	5'GGACG-CAGUGGC$^{AA}_{\ A}$ 3'CCUGA /GUCACCG$_{AA}$	(11)	5'GGACG-CAGUGGC$^{AA}_{\ A}$ 3'CCUGA/GUCACCG^{AA} A A
(2)	5'GGAC-GCAGUGGC$^{AA}_{\ A}$ 3'CCUG/AGUCACCG$_{AA}$	(12)	5'GGACG-CAGUGGC$^{AA}_{\ A}$ 3'CCUGA/GUCACCG$_{AA}$ P A
(3)	5'GGACG-AGUGGC^{G}C 3'CCUGA /UCACCG$_{A}$A	(13)	5'GGAC-GCAGUGGC$^{AA}_{\ A}$ 3'CCUG/AGUCACCG$_{AA}$ P A
(4)	5'GGAC-GAGUGGC^{G}C 3'CCUG/ AUCACCG$_{A}$A	(14)	5'GGAC-GCAGUGGC$^{AA}_{\ A}$ 3'CCUG/AGUCACCG$_{AA}$ G G
(5)	5'GGAC-GAGUGGC^{G}C 3'CCUG /AGCACCG$_{A}$A	(15)	5'GGACG-AGUGGC^{G}C 3'CCUGA/UCACCG$_{AA}$ A A
(6)	5'GGACG-AGUGGC^{G}C 3'CCUGA /GCACCG$_{A}$A	(16)	5'GGACG-AGUGGC^{G}C 3'CCUGA /GCACCG$_{AA}$ A P
(7)	5'GGACA-CAGUGGC$^{AA}_{\ A}$ 3'CCUGG/GUCACCG$_{AA}$	(17)	5'GGAGC-CAGUGGC$^{AA}_{\ A}$ 3'CCUCC/GUCACCG$_{AA}$ A A
(8)	5'GGAGC-CAGUGGC$^{AA}_{\ A}$ 3'CCUCC/GUCACCG$_{AA}$	(18)	5'GGACC-CAGUGGC$^{AA}_{\ A}$ 3'CCUGC/GUCACCG$_{AA}$ A A
(9)	5'GGACC-CAGUGGC$^{AA}_{\ A}$ 3'CCUGC/GUCACCG$_{AA}$	(19)	5'GGAC AAA3' 3'CCUG5'
(10)	5'GGACG-CAGUGGC$^{AA}_{\ A}$ 3'CCUGA/GUCACCG$_{AA}$ 5'A P3'	(20)	5'GGACAAAGAGUGGC^{G}C 3'CCUG CUCACCG$_{A}$A

[a] "–" denotes a continuation of the phosphate backbone; "/" denotes a break in the phosphate backbone; "P" denotes a purine riboside.

(Borer et al., 1974):

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

RESULTS

The choice of sequences for study was guided by searching many known RNA secondary structures for possible coaxially stacked helixes. Table 1 shows that the potential for coaxial stacking with intervening mismatches (e.g., the coaxial stacking of the D arm and the anticodon arm in yeast phenylalanine tRNA) is equal to or greater than that of coaxial stacking with some combination of AU, GC, or GU pairs at the interface. Table 2 shows that the most common intervening mismatch is a single GA. The double-GA mismatch sequence, $\begin{matrix}5'GA\\3'AG\end{matrix}$, also occurs rather frequently (Table 2), and a CC mismatch is rare. Thus these three motifs were chosen for study as representative of different classes. Table 3 shows that for the single-GA mismatch, the most common base pairs 5′ and 3′ of the G are CG or UA and AU or GC, respectively. Thus the $\begin{matrix}5'CGA\\3'GAU\end{matrix}$ motif was studied. The motif $\begin{matrix}5'CGC\\3'GAG\end{matrix}$ also occurs naturally and was studied to minimize potential effects due to fraying of the AU base pair.

As shown in Figure 1, coaxial stacking of helixes is modeled with a short oligomer bound to either a four- or five-nucleotide overhang at the 5′ end of the hairpin. The hairpin was designed to be stable below 80 °C so that the transition associated with disruption of the duplex is distinct from melting of the hairpin. Table 4 lists the sequences used. The initial hairpin loop was the GCAA tetraloop. Some of these sequences, however, formed self-complementary duplexes with an internal loop rather than hairpins. Thus, the loop was changed to AAAAA to ensure formation of the hairpin. Sequences of the 5′ overhang for the CC mismatch models (sequences 8, 9, 17, and 18 in Table 4) are also slightly different from the 5′ overhang of the GA mismatch models in order to avoid association of the overhangs. As a control, sequence 20 in Table 4 was studied where three adenosines were inserted between the hairpin stem and duplex helix in order to disrupt any coaxial stacking of helixes. For sequences 10, 12, 13, and 16 in Table 4, purine (P) was used as a dangling nucleotide in order to prevent any base pairing.

Figure 1 shows typical melting curves in which the lowest temperature transition reflects the binding of the short oligomer to the 5′ overhang of the hairpin stem. The T_M^{-1} vs $\ln(C_T/4)$ plots for all the duplex transitions are shown in Figure 2, and the results are summarized in Table 5. With two exceptions, all parameters derived from T_M^{-1} plots and fits of melting curves agree within 15%, which is consistent with the two-state model (Freier et al., 1983; Petersheim & Turner, 1983). The exceptions are $\begin{matrix}5'GGAGC\text{-}CA\\3'CCUCC/GU\end{matrix}$ (8) and $\begin{matrix}5'GGAGC\text{-}CA\\3'CCUCC/GU\\A\ \ A\end{matrix}$ (17) (numbers in parentheses that follow the sequences correspond to the sequence numbers in Table 4; this convention will be followed throughout the text). The extrapolation to 37 °C from their respective melting temperatures, however, is short, and therefore the measured ΔG°_{37} is reasonably reliable (Freier et al., 1984).

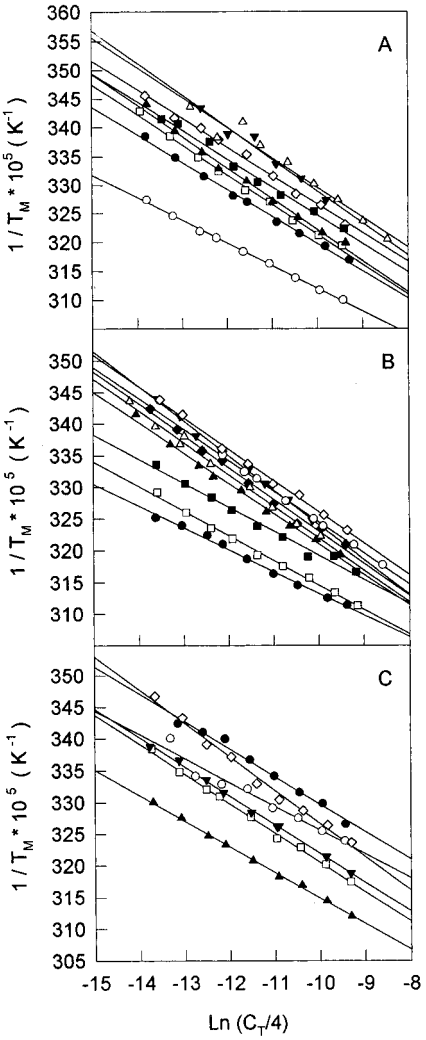

FIGURE 2: Plots of T_M^{-1} vs ln(C_T/4). The sequence numbers correspond to those listed in Table 4. For (A), (●) 6, (□) 7, (▲) 5, (◇) 8, (▼) 9, (○) 16, (■) 12, (△) $\frac{\text{GGACG}}{\text{CCUGA}}$. For (B), (●) 3, (□) 10, (▲) 4, (◆) 11, (▼) 2, (○) 1, (■) 14, (△) 15, (◇) 13. For (C), (○) 17, (□) 19, (▲) 19 pH = 5.5, (◇) 20, (▼) 20 pH = 5.5, (●) 18.

The thermodynamic parameters for coaxial stacking are calculated from equations equivalent to (Walter et al., 1994) $\Delta\Delta G^\circ_{37} = \Delta G^\circ_{37}$(hairpin + short oligomer) − ΔG°_{37}(short oligomer duplex). For cases where nucleotides at the 3′ end of the hairpin stem must be displaced to allow complete binding of the short oligonucleotide, corrected $\Delta\Delta G^\circ_{37}$ values are calculated from $\Delta\Delta G^\circ_{37}$(corr.) = ΔG°_{37}(hairpin + short oligomer) − ΔG°_{37}(short oligomer duplex) + ΔG°_{37}-(predicted for displacing the 3′ dangling end). Here ΔG°_{37}-(predicted for displacing the 3′ dangling end) is the nearest neighbor ΔG°_{37} for the 3′ dangling end. For $\begin{matrix}5'\text{GGACG-C}\\3'\text{CCUGA/G}\\\quad\quad(A)_2\end{matrix}$ (11) and $\begin{matrix}5'\text{GGACG-A}\\3'\text{CCUGA/U}\\\quad\quad(A)_2\end{matrix}$ (15), the ΔG°_{37}(predicted for displacing the 3′ dangling end) = $\Delta G^\circ_{37}\binom{5'\text{AA}}{3'\text{G}}$ + $\Delta G^\circ_{37}\binom{5'\text{GA}}{3'\text{C}}$ or $\binom{5'\text{UA}}{3'\text{A}}$ as appropriate) (Turner et al., 1988). Since there is no measured $\Delta G^\circ_{37}\binom{5'\text{AA}}{3'\text{G}}$, it was approximated with $\Delta G^\circ_{37}\binom{5'\text{AG}}{3'\text{G}}$ as measured by SantaLucia et al. (1991b). The parameters for $\Delta\Delta H^\circ$ and $\Delta\Delta S^\circ$ were calculated in an analogous manner. These results are listed in Table 6. For the $\begin{matrix}5'\text{GGACG}\\3'\text{CCUGA}\end{matrix}$ duplex, measured thermodynamic values were used. The measured ΔG°_{37} of this duplex is identical to that predicted by nearest neighbor parameters (Freier et al., 1986; SantaLucia et al., 1991b). The other pentamer duplexes were measured but showed non-two-state behavior. The T_M values for the tetramer duplexes are too low to allow accurate measurements (Walter et al., 1994; Walter & Turner, 1994). Thus, the thermodynamic parameters for the other short duplexes were calculated from nearest neighbor parameters.

For $\begin{matrix}5'\text{GGACG-C}\\3'\text{CCUGA/G}\\\quad\quad(A)_2\end{matrix}$ (11), an alternate conformation, $\begin{matrix}5'\text{GGAC-GC}\\3'\text{CCUG/AG}\\\quad\quad\text{A A}\end{matrix}$, is possible. Comparison of the average of fits and T_M^{-1} vs ln(C_T/4) plots suggests that the transition for this system is two-state (Table 5). This would imply either that one of these two conformations predominates or that the two conformations have similar thermodynamics. In order to understand which conformation might dominate, purine riboside, P, was substituted such that only $\begin{matrix}5'\text{GGAC-GC}\\3'\text{CCUG/AG}\\\quad\quad\text{P A}\end{matrix}$ (13) or $\begin{matrix}5'\text{GGACG-C}\\3'\text{CCUGA/G}\\\quad\quad\text{P}\\\quad\quad\text{A}\end{matrix}$ (12) could be formed. As shown in Table 5, the ΔG°_{37} for $\begin{matrix}5'\text{GGACG-C}\\3'\text{CCUGA/G}\\\quad\quad(A)_2\end{matrix}$ (11) is 0.3 kcal/mol less favorable than $\begin{matrix}5'\text{GGACG-C}\\3'\text{CCUGA/G}\\\quad\quad\text{P}\\\quad\quad\text{A}\end{matrix}$ (12) and 0.6 kcal/mol more favorable than $\begin{matrix}5'\text{GGAC-GC}\\3'\text{CCUG/AG}\\\quad\quad\text{P A}\end{matrix}$ (13). These differences of 0.3 and 0.6 kcal/mol are close to experimental error and are therefore ambiguous. It appears that $\begin{matrix}5'\text{GGAC-GC}\\3'\text{CCUG/AG}\\\quad\quad\text{A A}\end{matrix}$ and $\begin{matrix}5'\text{GGACG-C}\\3'\text{CCUGA/G}\\\quad\quad(A)_2\end{matrix}$ (11) have similar thermodynamics, and both conformations may exist in equilibrium with each other. Thus the thermodynamics of the model system probably provides a reasonable approximation for both potential structures. Similar arguments can be applied to $\begin{matrix}5'\text{GGACG-A}\\3'\text{CCUGA/U}\\\quad\quad(A)_2\end{matrix}$ (15).

DISCUSSION

Coaxial stacking of helixes is an important determinant of RNA secondary and three-dimensional structure. Thus knowledge of the energetics of coaxial stacking is required for successful prediction of RNA structure. It has been shown previously that coaxial stacking of helixes with Watson−Crick interfaces allows a short oligomer to bind more tightly to an overhang adjacent to a hairpin stem than predicted from the nearest neighbor parameters of the base pairs (Walter et al., 1994; Walter & Turner, 1994). Here we show that a GA or CC mismatch between helixes also leads to tighter binding.

When Coaxial Stacking Is Not Possible, Enhanced Binding Is Not Observed. The enhanced binding of oligomers adjacent to a pre-existing helix provides an empirical measure of the energetics of coaxial stacking. A control experiment

Table 5: Thermodynamics of Duplex Formation[a]

		$1/T_M$ vs $\ln(C_T/4)$ analysis				average of curve fits			
sequence no.	interface	$-\Delta H°$ (kcal mol^{-1})	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal mol^{-1})	T_M[b] (°C)	$-\Delta H°$ (kcal mol^{-1})	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal mol^{-1})	T_M[b] (°C)
(1)	GGACG-C- CCUGA/G-	57.49 ± 1.7	160.23 ± 5.4	7.80 ±0.03	44.0	50.80 ± 1.2	138.92 ± 3.7	7.72 ± 0.12	44.2
(2)	GGAC - GC- CCUG / AG-	38.11 ± 1.0	102.62 ± 3.3	6.28 ± 0.03	35.0	39.53 ± 3.0	107.12 ± 9.7	6.31 ± 0.07	35.3
(3)	GGACG-A- CCUGA/U-	51.09 ± 1.2	140.85 ± 3.9	7.41 ± 0.02	42.4	50.45 ± 3.4	138.64 ± 11.1	7.45 ± 0.04	42.7
(4)	GGAC-GA- CCUG/AU-	40.29 ± 1.6	110.99 ± 5.3	5.86 ± 0.06	31.9	38.44 ± 3.6	104.65 ± 11.9	5.99 ± 0.15	32.7
(5)	GGAC-GA- CCUG/AG-	40.13 ± 1.0	111.03 ± 3.3	5.70 ± 0.04	30.6	40.77 ± 3.1	112.95 ± 10.2	5.74 ± 0.08	31.1
(6)	GGACG-A- CCUGA/G-	52.45 ± 2.4	147.59 ± 7.7	6.67 ± 0.04	37.8	47.87 ± 1.8	132.65 ± 5.9	6.73 ± 0.13	38.3
(7)	GGACA - C- CCUGG / G-	50.28 ± 0.7	137.01 ± 2.4	7.79 ± 0.01	44.9	50.81 ± 1.9	138.6 ± 6.0	7.82 ± 0.03	45.1
(8)	GGAGC-CA- CCUCC/GU-	52.94 ± 4.7	152.52 ± 15.6	5.64 ± 0.14	31.8	38.15 ± 3.4	103.68 ± 10.5	6.00 ± 0.20	32.7
(9)	GCACC-CA- CGUGC/GU-	42.80 ± 0.8	117.34 ± 2.7	6.41 ± 0.02	36.1	41.36 ± 1.4	112.65 ± 4.4	6.42 ± 0.07	36.2
	pH = 5.5	49.08 ± 0.8	134.66 ± 2.4	7.31 ± 0.01	42.0	48.13 ± 1.5	131.58 ± 5.0	7.32 ± 0.05	42.2
(10)	GGACG - C- CCUGA / G- A P	41.45 ± 1.2	113.21 ± 3.9	6.34 ± 0.03	35.6	41.20 ± 2.2	112.25 ± 7.3	6.38 ± 0.07	35.9
(11)	GGACG-C- CCUGA/G- A A	36.70 ± 1.3	98.41 ± 4.2	6.18 ± 0.04	34.0	38.07 ± 3.6	102.71 ± 11.8	6.22 ± 0.08	34.5
(12)	GGACG-C- CCUGA/G- P A	41.97 ± 0.9	114.38 ± 3.0	6.50 ± 0.02	36.8	44.96 ± 3.6	123.93 ± 11.5	6.52 ± 0.04	36.9
(13)	GGAC - GC- CCUG / AG- P A	39.86 ± 1.1	110.34 ± 3.8	5.63 ± 0.05	30.2	36.87 ± 2.7	100.20 ± 9.2	5.79 ± 0.15	30.9
(14)	GGAC - GC- CCUG /AG- G G	35.72 ± 3.1	97.66 ± 10.4	5.43 ± 0.16	27.7	33.27 ± 6.4	89.29 ± 21.7	5.58 ± 0.37	28.4
(15)	GGACG-A- CCUGA/U- A A	39.39 ± 1.4	107.41 ± 4.7	6.08 ± 0.05	33.5	39.24 ± 5.2	106.56 ± 17.3	6.19 ± 0.16	34.3
(16)	GGACG - A- CCUGA / G- A P	39.44 ± 1.2	107.12 ± 4.1	6.22 ± 0.04	34.5	41.24 ± 4.1	112.86 ± 13.2	6.23 ± 0.07	34.8
(17)	GGAGC-CA- CCUCC/GU- A A	45.87 ± 2.2	131.38 ± 7.2	5.12 ± 0.09	27.8	36.64 ± 2.2	100.38 ± 7.5	5.51 ± 0.20	28.6
(18)	GCACC-CA- CGUGC/GU- A A	37.81 ± 1.4	103.65 ± 4.5	5.66 ± 0.06	30.0	40.84 ± 2.0	113.79 ± 6.6	5.58 ± 0.10	29.9
	pH = 5.5	43.71 ± 0.7	120.85 ± 2.4	6.23 ± 0.02	34.9	42.91 ± 1.4	118.21 ± 4.6	6.25 ± 0.04	35.0
	GGACG CCUGA	38.15 ± 2.9	105.89 ± 9.6	5.31 ± 0.13	27.4	34.74 ± 5.3	94.15 ± 17.7	5.53 ± 0.27	28.3
(19)	GGACAAA CCUG	40.60 ± 0.5	111.89 ± 1.8	5.90 ± 0.01	32.2	39.49 ± 3.1	108.09 ± 10.2	5.96 ± 0.07	32.6
(20)	GGACAAAG- CCUG C-	36.31 ± 0.9	97.82 ± 3.2	5.97 ± 0.04	32.3	41.45 ± 3.6	114.81 ± 11.6	5.84 ± 0.08	31.9
	GGAC CCUG	(35.7)	(102.2)	(3.9)					
	GGACA CCUGG	(41.30)	(116.0)	(5.3)					
	GCACC CGUGC	(38.3)	(107.3)	(5.0)					
	GGAGC CCUCC	(32.6)	(91.3)	(4.2)					

[a] T_M calculated for 10^{-4} M total strand concentration. [b] Numbers in parentheses are predicted values. The errors for the predicted thermodynamics are estimated as 10%, 10%, and 5% for $\Delta H°$, $\Delta S°$, and $\Delta G°$, respectively. For experimental measurements, significant figures are given beyond error estimates to allow accuate calculation of T_M and other parameters.

demonstrates that a single-stranded AAA linker inserted between helixes to prevent coaxial stacking eliminates tighter binding. The measured $\Delta G°_{37}$ of this complex is within 0.1 kcal/mol of the measured $\Delta G°_{37}$ of the duplex without the adjoining hairpin (Table 5, sequences 19 and 20). Thus, a system that is not expected to have coaxial stacking does not exhibit a favorable free energy increment. This contrasts with the previous result for the same sequence without the AAA linker where binding was enhanced by 3 kcal/mol due to coaxial stacking (Walter et al., 1994). Free energies for

Table 6: Thermodynamic Parameters of Coaxial Stacking[a]

sequence no.	interface	$-\Delta\Delta H°$ (kcal mol^{-1}) (corr.)	$-\Delta\Delta S°$ (eu) (corr.)	$-\Delta\Delta G°_{37}$ (kcal mol^{-1}) (corr.)
(1)	GGACG-C- CCUGA/G-	19.34 ± 3.4	54.35 ± 11.1	2.4 ± 0.1
(2)	GGAC - GC- CCUG / AG-	2.41 ± 3.7	0.42 ± 10.7	2.4 ± 0.2
(3)	GGACG-A- CCUGA/U-	12.94 ± 3.1	34.97 ± 10.4	2.1 ± 0.2
(4)	GGAC-GA- CCUG/AU-	4.59 ± 3.9	8.79 ± 14.5	2.00 ± 0.2
(5)	GGAC-GA- CCUG/AG-	4.43 ± 3.7	8.63 ± 10.7	1.8 ± 0.2
(6)	GGACG-A- CCUGA/G-	14.30 ± 3.8	41.71 ± 12.3	1.4 ± 0.2
(7)	GGACA - C- CCUGG / G-	8.98 ± 4.2	21.01 ± 11.8	2.5 ± 0.2
(8)	GGAGC-CA- CCUCC/GU-	20.34 ± 5.7	61.22 ± 18.1	1.4 ± 0.2
(9)	GCACC-CA- CGUGC/GU-	4.50 ± 3.9	10.04 ± 11.1	1.4 ± 0.2
	pH = 5.5	10.77 ± 3.9	27.36 ± 11.0	2.3 ± 0.2
(10)	GGACG - C- CCUGA / G- A P	3.30 ± 3.1 (10.70 ± 3.2)	7.33 ± 10.4 (27.33 ± 10.3)	0.9 ± 0.1 (2.0 ± 0.2)
(11)	GGACG-C- CCUGA/G- A A	-1.45 ± 3.2 (9.75 ± 3.4)	-7.47 ± 10.5 (23.12 ± 10.9)	0.9 ± 0.1 (2.7 ± 0.3)
(12)	GGACG-C- CCUGA/G- P A	3.82 ± 3.04 (11.22 ± 3.1)	8.50 ± 10.09 (28.50 ± 10.3)	1.2 ± 0.1 (2.3 ± 0.2)
(13)	GGAC - GC- CCUG / AG- P A	4.16 ± 3.75 (8.96 ± 3.9)	8.14 ± 10.9 (21.40 ± 11.2)	1.7 ±. 0.2 (2.4 ± 0.2)
(14)	GGAC - GC- CCUG /AG- G G	0.02 ± 4.72 (0.72 ± 4.8)	-4.54 ± 16.98 (8.76 ± 17.2)	1.5 ± 0.3 (2.2 ± 0.3)
(15)	GGACG-A- CCUGA/U- A A	1.24 ± 3.23 (−15.64 ± 3.5)	1.53 ± 10.72 (−43.42 ± 11.4)	0.8 ± 0.1 (2.2 ± 0.2)
(16)	GGACG - A- CCUGA / G- A P	1.29 ± 3.14 (3.39 ± 3.2)	1.24 ± 10.46 (6.04 ± 10.5)	0.8 ± 0.2 (1.4 ± 0.2)
(17)	GGAGC-CA- CCUCC/GU- A A	13.27 ± 3.91 (20.67 ± 4.0)	40.08 ± 11.63 (60.08 ± 11.8)	0.9 ± 0.2 (2.0 ± 0.3)
(18)	GCACC-CA- CGUGC/GU- A A	-0.49 ± 4.06 (6.91 ± 4.1)	-3.65 ± 11.65 (16.35 ± 11.8)	0.7 ± 0.2 (1.8 ± 0.3)
	pH = 5.5	5.41 ± 3.90 (12.81 ± 4.0)	13.55 ± 10.98 (33.55 ± 11.2)	1.2 ± 0.2 (2.3 ± 0.3)

[a] $\Delta\Delta H°$ and $\Delta\Delta S°$ were calculated with equations analogous to $\Delta\Delta G°_{37} = \Delta G°_{37}$(hairpin + short oligomer) − $\Delta G°_{37}$(short oligomer duplex). Corrected $\Delta\Delta H°$ and $\Delta\Delta S°$ were calculated with equations analogous to $\Delta\Delta G°_{37}$(corr.) = $\Delta\Delta G°_{37}$(uncorrected) − $\Delta G°_{37}$(predicted for displacing 3′ dangling end), where $\Delta G°_{37}$ (predicted for displacing 3′ dangling end) is the nearest neighbor $\Delta G°_{37}$ for the 3′ dangling nucleotide of the hairpin stem. Predicted $\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$ were used for short oligomer duplexes (Freier et al., 1986; SantaLucia et al., 1991b) except for $\frac{5'\text{GGACG}}{3'\text{CCUGA}}$, which was measured experimentally.

binding to receptors are expected to depend on the molecular weight of the complex (Williams et al., 1991). The difference in the translational component of binding a tetramer to a tetramer or 25-mer is predicted to be less than 0.6 kcal/mol, however. The results with the control containing an AAA linker suggest it is not necessary to apply correction factors when comparing binding to tetramers and to hairpins.

GA Mismatches Are Stabilizing at Helix Interfaces. Thermodynamic parameters of coaxial stacking calculated without corrections for molecular weight are listed in Table 6. For sequences where disruption of an interaction at the 3′ end of the hairpin stem is required for complete binding of the oligonucleotides, Table 6 also lists in parentheses a value for coaxial stacking that includes a correction term for the disruption of these pre-existing interactions. Figure 3 illustrates the free energy increments with and without correction terms. As seen in Table 6 and Figure 3, GA mismatches between two helixes stabilize the complex. The free energy increments for the interfaces $\frac{5'\text{CG-C}}{3'\text{GA/G}}$ (1) and $\frac{5'\text{CA-C}}{3'\text{GG/G}}$ (7) are −2.4 and −2.5 kcal/mol, respectively, so that both orientations of the GA mismatch have similar stabilities. The free energy increments for $\frac{5'\text{C-GA}}{3'\text{G/AG}}$ (5) and $\frac{5'\text{C-GA}}{3'\text{G/AU}}$ (4) are −1.8 and −2.0 kcal/mol, respectively, and those for $\frac{5'\text{CG-A}}{3'\text{GA/G}}$ (6) and $\frac{5'\text{CG-A}}{3'\text{GA/U}}$ (3) are −1.4 and −2.1 kcal/mol, respectively. Thus an interface involving a tandem GA mismatch can be almost as stabilizing as a single GA mismatch.

A comparison of the free energies of GA complexes with the free energies of Watson−Crick helical junctions (Walter & Turner, 1994) is shown in Table 7. Flush GA mismatches between helixes are worth significantly less than the corresponding GC interfaces. The flush single-GA and flush double-GA mismatch interfaces are on average 1.3 and 2.0 kcal/mol, respectively, less favorable than their corresponding GC counterparts. The flush GA mismatch interfaces are more on par with those of UA Watson−Crick interfaces. For example, the free energy increments of coaxial stacking for $\frac{5'\text{CG-C}}{3'\text{GA/G}}$ (1) and $\frac{5'\text{CG-A}}{3'\text{GA/U}}$ (3) are −2.4 and −2.1 kcal/mol, respectively, while those for $\frac{5'\text{U-C}}{3'\text{A/G}}$ and $\frac{5'\text{U-A}}{3'\text{A/U}}$ are also −2.4 and −2.1 kcal/mol, respectively (Walter & Turner, 1994).

CC Mismatches Are Also Stabilizing at Helix Interfaces. CC mismatch interfaces were studied as a control since they are one of the least frequently occurring motifs in known RNA secondary structures (Table 2). They are also the least frequently occurring mismatch between the D stem and the anticodon stem of tRNAs (R. Gutell, personal communication). Despite their rare occurrence, CC mismatches are surprisingly stable (Table 6 and Figure 3). The average $\Delta\Delta G°_{37}$ of −1.4 kcal/mol for the flush CC interfaces is equivalent to the $\frac{5'\text{G-A}}{3'\text{A/G}}$ (6) interface. When measured at pH 5.5, the $\frac{5'\text{CC-C}}{3'\text{GC/G}}$ (9) interface is even more stable at −2.3 kcal/mol. This extra stabilization may be due to formation of a C-C$^+$ base pair. C-C$^+$ base pairs have been observed in several other contexts (Guschlbauer, 1975, and references cited therein; SantaLucia et al., 1991a; Antao & Gray, 1993; Chen et al., 1994; Kang et al., 1994). Evidently, the thermodynamics of coaxial stacking involving mismatches is not strongly dependent on the mismatch.

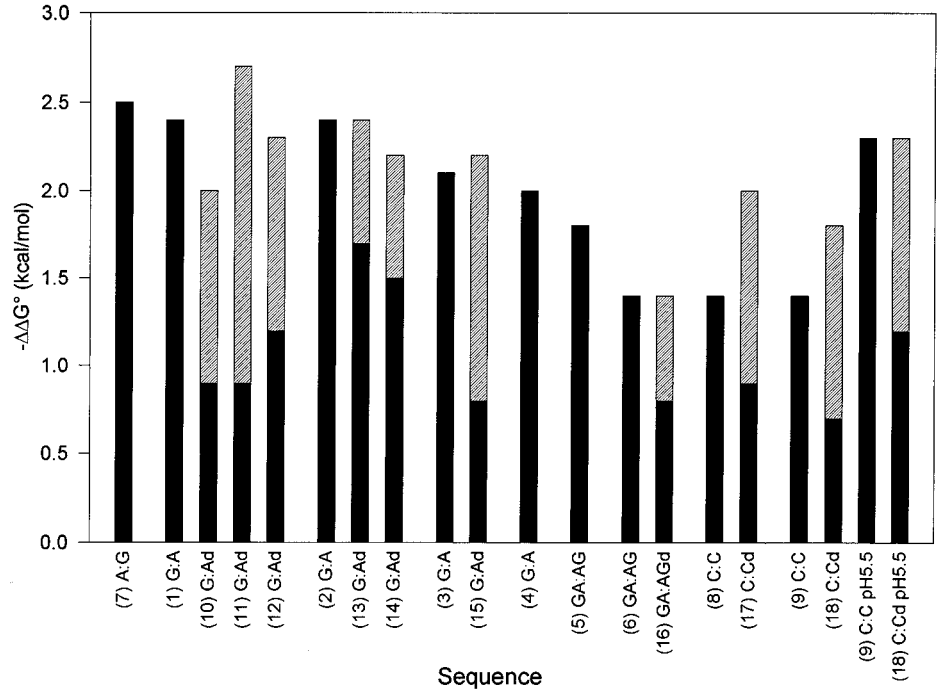

FIGURE 3: Free energy increments of coaxial stacking corresponding to Table 6. The abscissa labels indicate the mismatch, the sequence number (see Table 4), and if the sequence has dangling ends (d). The hashed bars for sequences with dangling ends are the corrected free energy increments for initially breaking the 3′ dangling end (see text).

Effect of Extending the Chains beyond the Helix−Helix Interface. Since naturally occurring helical interfaces in RNA are often followed by single-stranded regions, unpaired nucleotides were added to extend the chains and thus more closely mimic naturally occurring interfaces. Such extensions decrease the stabilities of the oligonucleotide−hairpin complexes by an average of ∼1 kcal/mol with a maximum decrease of 1.5 kcal/mol. For similar complexes with GC pairs at the interfaces, the average decrease in stability is ∼2.3 kcal/mol with a maximum decrease of 2.8 kcal/mol (Walter et al., 1994). Thus interfaces with GA and CC mismatches are destabilized less by chain extension than interfaces with GC pairs.

When 3′ dangling end corrections are made, the corrected $\Delta\Delta G^\circ_{37}$ values for the extended GA and CC systems are nearly identical with the $\Delta\Delta G^\circ_{37}$ values for the flush interfaces (Figure 3). For example, the corrected $\Delta\Delta G^\circ_{37}$ for $\begin{matrix}5'\text{GGAC-GC}\\ 3'\text{CCUG/AG}\\ \quad\ \text{P A}\end{matrix}$ (13) is −2.4 kcal/mol, which is identical to the $\Delta\Delta G^\circ_{37}$ of $\begin{matrix}5'\text{GGAC-GC}\\ 3'\text{CCUG/AG}\end{matrix}$ (2). This is not the case with interfaces containing only GC base pairs. When the 3′ dangling end correction is applied in these cases, the corrected $\Delta\Delta G^\circ_{37}$ values are consistently less stable than their corresponding flush interfaces by at least 1 kcal/mol. For instance, the $\Delta\Delta G^\circ_{37}$ of $\begin{matrix}5'\text{GGAG-C}\\ 3'\text{CCUC/G}\end{matrix}$ is −4.3 kcal/mol, whereas the corrected $\Delta\Delta G^\circ_{37}$ of $\begin{matrix}5'\text{GGAG-C}\\ 3'\text{CCUC/G}\\ \quad\ \text{U A}\end{matrix}$ is −2.6 kcal/mol. As shown in Table 7, the net effect is that GA interfaces are only an average of 0.6 kcal/mol less favorable than all GC interfaces when the hairpin or both chains are extended. While the chain extensions studied have different sequences for the two cases, the trend is clear.

Position of the Break in the Sugar-Phosphate Backbone Has Little Effect on Stability. The mismatched interface systems seem to have no energetic preference for a 5′ or 3′ break in the phosphate backbone adjacent to the GA mismatch. Comparison of $\begin{matrix}5'\text{CG-C}\\ 3'\text{GA/G}\end{matrix}$ (1) with $\begin{matrix}5'\text{C-GC}\\ 3'\text{G/AG}\end{matrix}$ (2) and $\begin{matrix}5'\text{CG-A}\\ 3'\text{GA/U}\end{matrix}$ (3) with $\begin{matrix}5'\text{CG-A}\\ 3'\text{G/AU}\end{matrix}$ (4) shows that the $\Delta\Delta G^\circ_{37}$ values are virtually the same (see Figure 3). A further comparison between $\begin{matrix}5'\text{CG-C}\\ 3'\text{GA/G}\\ \quad\ \text{A P}\end{matrix}$ (10) and $\begin{matrix}5'\text{C-GC}\\ 3'\text{G/AG}\\ \quad\ \text{P A}\end{matrix}$ (13) shows that even with the addition of dangling nucleotides to the helixes, the position of the backbone break makes little difference in the coaxial stacking free energies. Inspection of the double-GA mismatches shows a small decrease in free energy increment (from −1.8 to −1.4 kcal/mol) when the position of the phosphodiester backbone break is changed from $\begin{matrix}5'\text{C-GA}\\ 3'\text{G/AG}\end{matrix}$ (5) to $\begin{matrix}5'\text{CG-A}\\ 3'\text{GA/G}\end{matrix}$ (6). This difference, however, is within the experimental uncertainty.

Switch from CG to an AU Base Pair at the Interface Has Little Effect on Stability. There is only a 0.3 kcal/mol difference in stability between $\begin{matrix}5'\text{CG-C}\\ 3'\text{GA/G}\end{matrix}$ (1) and $\begin{matrix}5'\text{CG-A}\\ 3'\text{GA/U}\end{matrix}$ (3), which is within experimental error. This suggests the free energy increment for a GA mismatch will be relatively insensitive to the Watson−Crick base pair at the interface.

No Effect Is Observed from Potential Base Triples. Base triples involving coaxially stacked helixes are known to occur in tRNA (Holbrook et al., 1978) and group I introns (Michel & Westhof, 1990; Michel et al., 1990; Chastain & Tinoco, 1992, 1993). As shown in Figure 4, sequence 14 contains a motif that is related to the D stem and the anticodon stem in yeast phenylalanine tRNA where two G's form base triples (Kim et al., 1974; Robertus et al., 1974). The differences are that the model system has no modified nucleotides and the middle two base pairs are interchanged. Sequence 14 was studied to see if a free energy increment could be detected for potential base triples. The $\Delta\Delta G^\circ_{37}$ of this

Table 7. Comparison of Free Energy Increments between Mismatch Interfaces and Watson−Crick Interfaces

sequence no.	interface	$-\Delta\Delta G°_{37}$ (kcal mol^{-1}) (corr.)[c]	interface[a]	$-\Delta G°_{37}$[b] (kcal mol^{-1}) (corr.)[c]
(1)	GGACG-C- CCUGA/G-	2.4 ± 0.1	GGAG-C- CCUC/G-	4.3 ± 0.2
(2)	GGAC - GC- CCUG / AG-	2.4 ± 0.2	GGAC - G- CCUG / C-	3.0 ± 0.2
(3)	GGACG-A- CCUGA/U-	2.1 ± 0.1	GGAG-A- CCUC/U-	3.9 ± 0.2
(4)	GGAC-GA- CCUG/AU-	2.0 ± 0.2	GGAC - G- CCUG / C-	3.0 ± 0.2
(5)	GGAC-GA- CCUG/AG-	1.8 ± 0.2	GGAC - G- CCUG / C-	3.0 ± 0.2
(6)	GGACG-A- CCUGA/G-	1.4 ± 0.1	GGAG-C- CCUC/G-	4.3 ± 0.2
(7)	GGACA - C- CCUGG / G-	2.5 ± 0.2	GGAC - C- CCUG / G-	4.1 ± 0.2
(10)	GGACG - C- CCUGA / G- A P	0.9 ± 0.1 (2.0 ± 0.2)	GGACG - C- CCUGC / G- U A	1.5 ± 0.2 (2.6 ± 0.3)
(11)	GGACG-C- CCUGA/G- A A	0.9 ± 0.1 (2.7 ± 0.3)	GGAG - C- CCUC / G- A	2.3 ± 0.2 (3.4 ± 0.3)
(12)	GGACG-C- CCUGA/G- P A	1.2 ± 0.1 (2.3 ± 0.2)	GGAG - C- CCUC / G- A	2.3 ± 0.2 (3.4 ± 0.3)
(13)	GGAC - GC- CCUG / AG- P A	1.7 ± 0.2 (2.4 ± 0.2)	GGAC - G- CCUG / C- A A	0.4 ± 0.2 (2.1 ± 0.3)
(14)	GGAC - GC- CCUG /AG- G G	1.5 ± 0.3 (2.2 ± 0.3)	GGAC - G CCUG /C- A	1.2± 0.2 (2.9 ± 0.3)
(15)	GGACG - A- CCUGA / G- A P	0.8 ± 0.2 (1.4 ± 0.2)	GGACG - C- CCUGC / G- U A	1.5 ± 0.2 (2.6 ± 0.3)

[a] Walter et al. (1994); Walter and Turner (1994). [b] Measured $\Delta\Delta G°_{37}$ = $\Delta G°_{37}$(hairpin + short oligomer) − $\Delta G°_{37}$(short oligomer duplex). The nearest neighbor $\Delta G°_{37}$ for the short oligomer duplex was used (Freier et al., 1986; SantaLucia et al., 1991b) except for $\frac{5'GGACG}{3'CCUGA'}$ which was measured experimentally. [c] Values in parentheses are corrected for stacking of any 3′ unpaired nucleotide on the hairpin stem such that $\Delta\Delta G°_{37}$(corr.) = $\Delta\Delta G°_{37}$(uncorrected) − $\Delta G°_{37}$(predicted for displacing 3′ dangling end), where $\Delta G°_{37}$(predicted for displacing 3′ dangling end) is the nearest neighbor $\Delta G°_{37}$ for the 3′ dangling nucleotide of the hairpin stem. For $\Delta G°_{37}$(predicted for displacing 3′ dangling end), see Results.

system is −1.5 kcal/mol, which is less favorable than the increment of -2.4 kcal/mol for $\frac{5'GGAC-GC}{3'CCUG/AG}$ (2) and nearly identical to the free energy increment of −1.7 kcal/mol for

5′GGAC-GC
3′CCUG/AG (13). Thus, either the base triples do not
P A

form in this model or they do not contribute any thermodynamic stability. There are several possible reasons for this result. The base methylations or the sequence of the middle two base pairs of the tetramer helix may be crucial. Another hypothesis is that small oligonucleotide models have more flexibility than their corresponding sequence in tRNA and thus can sample a greater conformational space. This favorable entropy may outweigh any stability gained from the formation of base triples. These small models are also more exposed to the aqueous solvent than the corresponding

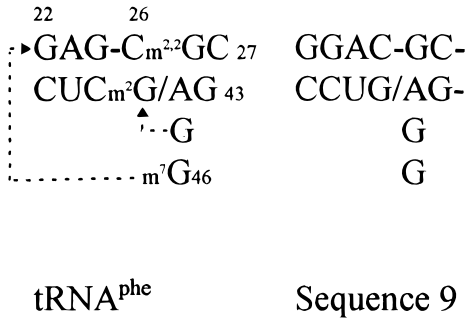

FIGURE 4: (Left) Juncture of the D and anticodon stems of yeast phenylalanine tRNA. The dashed arrows indicate triple base pair interactions. m^2G = N^2-methylguanosine; $m^{2,2}G$ = N^2-dimethylguanosine; m^7G = 7-methylguanosine. (Right) Juncture of sequence 9 (see Table 4) from this study.

motif in tRNA. This difference in environment may also be a factor. Recently, Sarkar et al. (1996) reported CD evidence for the formation of base triples in a small oligonucleotide model in the presence of Mg^{2+} or Mn^{2+}. Base triple formation was not seen in the absence of divalents, however, even at Na^+ concentrations of 0.5 M. Thus, the formation of base triples may be favorable in the context of a large RNA or in the presence of divalent cations even though they are not mimicked by the current model in 1 M Na^+.

Comparison with Parameters Used To Predict RNA Secondary Structure. Walter et al. (1994) added coaxial stacking in a preliminary way to the Zuker (1989) method for prediction of RNA secondary structure. In that study, stabilities of potential secondary structures were recalculated with coaxial stacking of Watson−Crick interfaces approximated with Watson−Crick nearest neighbor parameters. A special bonus of −1.0 kcal/mol was also included as an approximation for the GA or AG mismatch interface. No bonus was included for any other mismatch at the helix−helix interface, however. The averages of the corrected $\Delta\Delta G°_{37}$ values measured for the single-GA and -CC mismatch interfaces with extended chains are −2.3 and −1.9 kcal/mol, respectively (Table 6). Thus the free energy increments used for mismatched interfaces in the folding algorithm are too small. The bonus of −1 kcal/mol does, however, approximate the relative insensitivity of GA mismatch interfaces to the adjacent base pair as well as the equal thermodynamic stabilities of GA and AG mismatches between helixes.

Thermodynamics of Coaxial Stacking Do Not Correlate with Frequencies of Natural Occurrence. The surprising stabilities of the CC mismatch interfaces do not correlate with the frequency of occurrence in known secondary structures (Table 2). CC mismatches are one of the most infrequent mismatches at helical interfaces, and yet their free energy increments are similar to those of GA mismatches, which occur frequently. This suggests that the major determinant of sequence specificity of the mismatch between helixes in secondary structures may not be the thermodynamics of the interface but rather the three-dimensional architecture of the RNA. This may depend on local structure of the mismatches or on tertiary interactions involving mismatches.

Coaxial stacking of helixes can act as a bridge between secondary and tertiary structure. As thermodynamic and structural information on this motif expands, presumably

rules will be obtained concerning their behavior. The localization of two or more helixes in three-dimensional space then provides an important step toward the prediction of three-dimensional RNA structure.

ACKNOWLEDGMENT

We thank Dr. Robin R. Gutell for providing the frequency of mismatches occurring between the D and anticodon stems of tRNA, Dr. Dudley Williams for discussions of the effects of molecular weight on binding constants, and a reviewer for suggesting Figure 3.

REFERENCES

Antao, V. P., & Gray, D. M. (1993) *J. Biomol. Struct. Dyn. 10*, 819−839.
Biou, V., Yaremchuk, A., Tukalo, M., & Cusack, S. (1994) *Science 263*, 1404−1410.
Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. 1, p 597, CRC Press, Cleveland, OH.
Borer, P. N., Dengler, B., Tinoco, I., Jr., & Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843−853.
Chastain, M., & Tinoco., I., Jr. (1992) *Biochemistry 31*, 12733−12741.
Chastain, M., & Tinoco., I., Jr. (1993) *Biochemistry 32*, 14220−14228.
Chen, L., Cai, L., Zhang, X., & Rich, A. (1994) *Biochemistry 33*, 13540−13546.
Chou, S.-H., Flynn, P., & Reid, B. R. (1989) *Biochemistry 28*, 2422−2435.
Ekland, E. H., Szostak, J. W., & Bartel, D. P. (1995) *Science 269*, 364−370.
Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., & Turner, D. H. (1983) *Biochemistry 22*, 4533−4539.
Freier, S. M., Petersheim, M., Hickey, D. R., & Turner, D. H. (1984) *J. Biomol. Struct. Dyn. 1*, 1129−1242.
Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.
Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301−319.
Gultyaev, A. P., van Batenburg, F. H. D., & Pleij, C. W. A. (1995) *J. Mol. Biol. 250*, 37−51.
Guschlbauer, W. (1975) *Nucleic Acids Res. 2*, 353−360.
Gutell, R. R., Weiser, B., Woese, C. R., & Noller, H. F. (1985) *Prog. Nucleic Acids Res. 32*, 155−195.
Gutell, R. R., Schnare, M. N., & Gray, M. W. (1992) *Nucleic Acids Res. Suppl. 20*, 2095−2109.
Gutell, R. R., Larsen, N., & Woese, C. R. (1994) *Microbiol. Rev. 58*, 10−26.
Hall, R. (1971) *The Modified Nucleosides in Nucleic Acids*, Chapter 4, Columbia University Press, New York.
Holbrook, S. R., Sussman, J. L., Warrant, R. W., & Kim, S.-H. (1978) *J. Mol. Biol. 123*, 631−660.
Huang, W. M., & Ts'o, P. O. P. (1966) *J. Mol. Biol. 16*, 523−543.
Jaeger, J. A., Turner, D. H., & Zuker, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7706−7710.
Kang, C., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R., & Rich, A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11636−11640.
Kim, S. H., & Cech, T. R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8788−8792.
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.
Laing, L. G., & Draper, D. E. (1994) *J. Mol. Biol. 237*, 560−576.
Marino, J. P., Gregorian, R. S., Jr., Csankovszki, G., & Crothers, D. M. (1995) *Science 268*, 1446−1454.
Michel, F., & Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.
Michel, F., Ellington, A. D., Couture, S., & Szostak, J. W. (1990) *Nature 347*, 578−580.
Nickrent, D. L., & Sargent, M. L. (1991) *Nucleic Acids Res. 19*, 227−235.
Ogilvie, K. K., Usman, N., Nicoghosian, K., & Cedergren, R. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5764−5768.
Pace, N. R., Smith, D. K., Olsen, G. J., & James, B. D. (1989) *Gene 82*, 65−75.
Petersheim, M., & Turner, D. H. (1983) *Biochemistry 22*, 256−263.
Pley, H. W., Flaherty, K. M., & McKay, D. B. (1994) *Nature 372*, 68−74.
Richards, E. G. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. 1, p 579, CRC Press, Cleveland, OH.
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.
Sarkar, M., Sigurdsonn, S., Tomac, S., Srikanta, S., Rozners, E., Sjöberg, B.-M., Strömberg, R., & Gräslund, A. (1996) *Biochemistry 35*, 4678−4688.
Schmitt, M. E., Bennett, J. L., Dairaghi, D. J., & Clayton, D. A. (1993) *FASEB J. 7*, 208−213.
Schmitz, M., & Steger, G. (1996) *J. Mol. Biol. 255*, 254−266.
Scott, W. G., Finch, J. T., & Klug, A. (1995) *Cell 81*, 991−1002.
Sprinzl, M., Hartman, T., Weber, J., Blank, J., & Zeidler, R. (1989) *Nucleic Acids Res. Suppl. 17*, r1−r172.
Stern, S., Weiser, B., & Noller, H. F. (1988) *J. Mol. Biol. 204*, 447−481.
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. J. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.
Walter, A. E., & Turner, D. H. (1994) *Biochemistry 33*, 12715−12719.
Walter, A. E., Turner, D. H., Kim, J., Lyttle, M. H., Muller, P., Mathews, D. H., & Zuker, M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9218−9222.
Wang, Y. Y., Lyttle, M. H., & Borer, P. N. (1990) *Nucleic Acids Res. 18*, 2069−2076.
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.
Westhof, E., Dumas, P., & Moras, D. (1985) *J. Mol. Biol. 184*, 119−145.
Williams, D. H., Cox, J. P. L., Doig, A. J., Gardner, M., Gerhard, U., Kaye, P. T., Lal, A. R., Nicholls, I. A., Salter, C. J., & Mitchell, R. C. (1991) *J. Am. Chem. Soc. 113*, 7020−7030.
Wilson, C., & Szostak, J. W. (1995) *Nature 374*, 777−782.
Wu, T., Ogilvie, K. K., & Pon, R. T. (1989) *Nucleic Acids Res. 17*, 3501−3517.
Zuker, M. (1989) *Science 244*, 48−52.
Zuker, M., Jaeger, J. A., & Turner, D. H. (1991) *Nucleic Acids Res. 19*, 2707−2714.

BI960913Z