Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 9218–9222, September 1994
Biochemistry

Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding

AMY E. WALTER*, DOUGLAS H. TURNER*†, JAMES KIM*, MATTHEW H. LYTTLE‡§, PETER MÜLLER¶, DAVID H. MATHEWS*, AND MICHAEL ZUKER‖

*Department of Chemistry, University of Rochester, Rochester, NY 14627-0216; ‡Millipore Corp., Bedford, MA 01730; ¶Dr. Karl Thomae, GmbH, Biberach an der Riss, 7950 Biberach 1, Germany; and ‖Institute for Biomedical Computing, Washington University, St. Louis, MO 63110

Communicated by Olke C. Uhlenbeck, May 26, 1994 (received for review September 28, 1993)

ABSTRACT **An RNA model system consisting of an oligomer binding to a 4-nt overhang at the 5′ end of a hairpin stem provides thermodynamic parameters for helix–helix interfaces. In a sequence-dependent manner, oligomers bind up to 1000-fold more tightly adjacent to the hairpin stem than predicted for binding to a free tetramer at 37°C. For the interface (/) in 5′-GGAC G-**
ACCUG/C-, where X and Z are unpaired nucleotides, the
X Z
additional free energy change, $\Delta\Delta G^\circ_{37}$, for binding is roughly the nearest-neighbor ΔG°_{37} for propagation of an uninterrupted helix of equivalent sequence, $\frac{CG}{GC}$. When X and Z are omitted, the $\Delta\Delta G^\circ_{37}$ is even more favorable by ≈1 kcal/mol (1 cal = 4.184J). On average, predictions of 11 RNA secondary structures improve from 67 to 74% accuracy by inclusion of similar stacking contributions.

RNA is important for translation, regulation, and catalysis and as a target for therapeutics (1). Understanding and controlling these functions require knowledge of RNA structure. Although RNA sequences are being discovered rapidly, only three-dimensional structures of tRNA are known (2, 3). Thus, various approaches have been devised for modeling RNA structure (4–15). One approach used independently or in combination with others is free-energy minimization (10–15). This approach provides a first-approximation prediction of secondary structure from sequence. One motif that is often predicted poorly is the multibranch loop or junction, where three or more helixes converge (15). Coaxial stacking of helixes in multibranch loops is seen in the crystal structures of tRNAs (2, 3). Coaxial-stacking interactions, however, are not included in previous algorithms for free-energy minimization (12–14). Here we provide evidence that coaxial stacking provides large, favorable free-energy contributions. Inclusion of such interactions leads to improved predictions of RNA secondary structure.

MATERIALS AND METHODS

Oligonucleotides. Oligoribonucleotides were synthesized by standard methods (16–18), and one hairpin was extended by adding pUp with the T4 RNA ligase reaction (19, 20). Synthetic details will be provided elsewhere.

Measurement of Thermodynamic Parameters. The buffer for melting experiments was 1 M NaCl/10 mM sodium cacodylate/0.5 mM EDTA. Absorbance versus temperature curves were measured at 280 nm with a heating rate of 1°C/min. These curves were analyzed by fitting to a two-state model with sloping base lines through use of a nonlinear least-squares program (21).

Predictions of Secondary Structure. Sets of 20 secondary structures were generated with the program of Zuker (12),

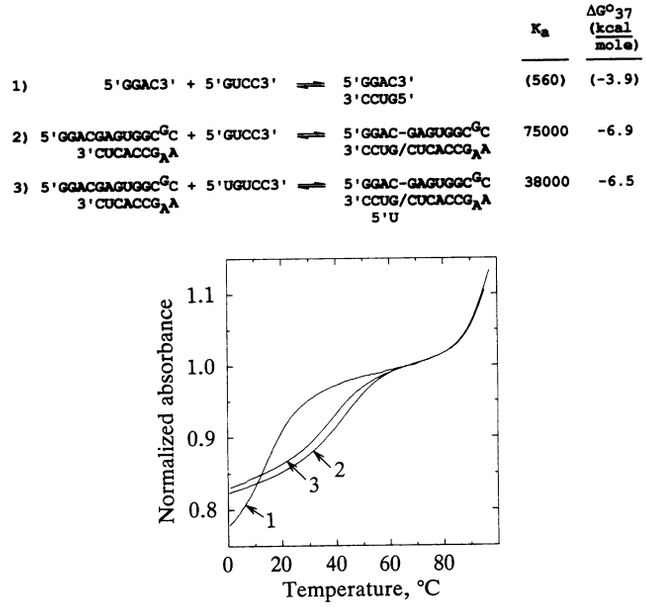

FIG. 1. Typical melting curves. Absorbances were normalized by dividing by the absorbance at 67°C. Sequences are shown above. Curve 1 has one transition for the tetramer duplex melt. The first transition of curves 2 and 3 represents melting of the short oligomer from the hairpin. This transition is fit to obtain the coaxial stacking parameters. The second transition is the melting of the hairpin with a $t_m > 100$°C. The concentrations and t_m values for the curves are as follows: 1.20×10^{-4} M and 12.6°C for curve 1, 1.59×10^{-4} M and 42.3°C for curve 2, and 1.18×10^{-4} M and 37.7°C for curve 3.

and the ΔG°_{37} for each structure was then recalculated with possible coaxial stacking and a Jacobson–Stockmayer function for multibranch loops larger than 6 nt [ΔG°_{37} (kcal per mol; 1 cal = 4.184J) = 4.6 + 0.1 (no. of helixes) + 2.4 + 1.1 ln (no. of unpaired nt/6)]. Thermodynamic parameters were those listed by Jaeger *et al.* (14) with the following exceptions. (*i*) G·U pairs were from He *et al.* (22). (*ii*) The asymmetry penalty for internal loops with branches of N_1 and N_2 nucleotides was the minimum of 3.0 or 0.3 $|N_1 - N_2|$ kcal/mol, and internal loops of three were given a ΔG°_{37} for loop closure of 5.1 + 0.3 = 5.4 kcal/mol (23). (*iii*) Stacking and pairing of terminal mismatches were included in calculating free energies of internal loops in the following manner. Single mismatches were given ΔG°_{37} = 0.8 kcal/mol. For larger internal loops, terminal GA, UU, and other mismatches adjacent to C·G pairs were given favorable free energy increments of −2.7, −2.5, and −1.5 kcal/mol, respectively (24). These mismatches adjacent to A·U pairs were given ΔG° values of −2.2, −2.0, and −1.0 kcal/mol. (*iv*) For

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

†To whom reprint requests should be addressed.
§Present address: Terrapin Technologies, 750 H Gateway Boulevard, South San Francisco, CA 94080.

hairpin loops of 3, 4, 5, 6, 7, 8, and 9, ΔG°_{37} for loop closure was 4.1, 4.9, 4.4, 4.7, 5.0, 5.1, and 5.2 kcal/mol, respectively (25–29). UCCG, GUAA, CUUG, AUUU, and UUUA were added to the list of extra-stable tetraloops. (*v*) Hairpin loops with >4 nt were made more favorable by 0.7 kcal/mol if GA is the first mismatch (30). (*vi*) Terminal mismatches $\frac{5'\text{-CG}}{3'\text{-GA}}$ $\frac{5'\text{-CA}}{3'\text{-GA}}$, and $\frac{5'\text{-CG}}{3'\text{-GG}}$ were given $\Delta G^\circ_{37} = -1.4$ kcal/mol (25). The structures with lowest free energy before (optimal) and after recalculation (coaxial stacking plus Jacobson–Stockmayer reordered optimal) were compared with phylogenetically determined secondary structures as described by Jaeger *et al.* (14). For folding, RNAs were divided into domains as described by Jaeger *et al.* (14), except domain 3 of *Escherichia coli* was taken as nt 920–1397, and B1 was divided as nt 1–385 and 388–768. Group II introns were also folded without being split into two domains, and the results are given in parentheses. Structures were compared with those in refs. 6–8 and those listed in ref. 14.

RESULTS AND DISCUSSION

The effect of coaxial stacking on helix stability was determined by measuring thermodynamic parameters for helix formation when the helix is formed adjacent to a preexisting helix and comparing them to parameters predicted or measured for an isolated helix. The model system for adjacent helixes consists of GUCC binding to an unpaired GGAC overhang at the 5′ end of a hairpin stem (Fig. 1). The hairpin is designed to be stable below 80°C so that the transition involving GUCC binding can be followed separately. As shown in Fig. 1 and Table 1, the melting temperature, t_m, for the complex with hairpin is ≈20°C higher than the t_m for the $\frac{\text{GGAC}}{\text{CCUG}}$ duplex without the adjacent helix. For the $\frac{\text{C-G}}{\text{G/C}}$

Table 1. Thermodynamic parameters for binding to hairpins in 1 M NaCl

Sequence at interface	$-\Delta H^\circ$,* kcal·mol^{-1}	$-\Delta S^\circ$, cal·K^{-1}·mol^{-1}	$-\Delta G^\circ_{37}$,* kcal·mol^{-1}	t_m, °C	$-\Delta\Delta G^\circ_{37}$, kcal·mol^{-1} measured† (Corr.)‡
Hairpins with CG interface					
5′–GGAC–G– / CCUG/C–	42.7 ± 1	115.3 ± 4	6.92 ± 0.03	40	3.0
	38.2 ± 2	100.5 ± 7	7.00 ± 0.2	41	(3.0)
5′–GGAC–G– / CCUG/C– (A)	40.7 ± 2	109.9 ± 5	6.65 ± 0.04	38	2.8
	36.7 ± 3	96.4 ± 9	6.78 ± 0.1	39	(2.8)
5′–GGAC–G– / CCUG/C– (U)	42.9 ± 2	117.5 ± 6	6.50 ± 0.05	37	2.6
	40.3 ± 3	108.4 ± 12	6.62 ± 0.1	38	(2.6)
5′–GGAC–G– / CCUG/C– (A)	29.1 ± 1	77.5 ± 3	5.08 ± 0.04	22	1.2
	37.4 ± 7	104.5 ± 22	4.95 ± 0.2	24	(2.9)
5′–GGAC–G– / CCUG/C– (Up)	32.9 ± 3	88.7 ± 6	5.37 ± 0.08	26	1.5
	28.9 ± 3	75.1 ± 10	5.58 ± 0.2	27	(2.7)
5′–GGAC–G– / ACCUG/C– (U)	48.1 ± 1	131.8 ± 5	7.19 ± 0.03	41	1.5
	40.9 ± 2	108.5 ± 5	7.19 ± 0.2	42	(2.7)
5′–GGAC–G– / ACCUG/C– (A A)	40.1 ± 2	109.6 ± 8	6.11 ± 0.08	34	0.4
	36.5 ± 4	97.6 ± 12	6.27 ± 0.2	35	(2.1)
5′–GGAC–G– / ACCUG/C– (A U)	44.3 ± 2	121.4 ± 7	6.59 ± 0.05	37	0.9
	39.7 ± 3	106.4 ± 9	6.70 ± 0.2	38	(2.1)
5′–GGAC–G– / ACCUG/C– (U A)	42.6 ± 2	118.9 ± 6	5.76 ± 0.07	31	0.1
	38.4 ± 4	104.5 ± 14	5.97 ± 0.2	33	(1.8)
5′–GGAC–G– / ACCUG/C– (U U)	41.6 ± 1	115.7 ± 5	5.77 ± 0.05	31	0.1
	35.4 ± 5	94.8 ± 17	6.04 ± 0.3	33	(1.3)
5′–GGAC–G– / ACCUG/C– (C A)	41.1 ± 2	111.2 ± 6	6.58 ± 0.05	37	0.9
	35.4 ± 4	92.5 ± 12	6.75 ± 0.2	39	(2.6)
5′–GGAC–G– / ACCUG/C– (C U)	39.6 ± 4	106.4 ± 12	6.66 ± 0.09	38	1.0
	34.8 ± 2	90.5 ± 7	6.77 ± 0.2	39	(2.2)
Hairpins with GC interface§					
5′–GGAG–C– / CCUC/G–	43.3 ± 1	114.4 ± 3	7.80 ± 0.02	46	4.3
	46.2 ± 2	123.7 ± 6	7.86 ± 0.1	46	(4.3)
5′–GGAG–C– / CCUC/G– (A)	53.2 ± 3	148.0 ± 11	7.29 ± 0.05	41	3.8
	39.6 ± 3	104.3 ± 9	7.29 ± 0.2	43	(3.8)
5′–GGAG–C– / CCUC/G– (U)	44.3 ± 1	118.6 ± 4	7.56 ± 0.03	44	4.1
	43.3 ± 2	115.1 ± 7	7.58 ± 0.1	45	(4.1)
5′–GGAG–C– / CCUC/G– (A)	49.3 ± 3	140.1 ± 10	5.80 ± 0.09	32	2.3
	35.8 ± 2	95.7 ± 8	6.18 ± 0.3	34	(3.4)
5′–GGAG–C– / CCUC/G– (U A)	34.7 ± 2	95.6 ± 6	5.03 ± 0.1	24	1.5
	38.1 ± 7	106.8 ± 25	5.01 ± 0.3	25	(2.6)
Oligomers¶					
5′–GGAC–3′ / CCUG	(35.7)	(102.2)	(3.9)	(16.4)	
5′–GGAC / CCUGU	(36.1)	(103.4)	(3.9)	(16.8)	
5′–GGAC / ACCUGU	37.7 ± 0.7	103.4 ± 2	5.67 ± 0.02	30	
	39.1 ± 4	107.6 ± 12	5.71 ± 0.07	31	
	(45.1)	(126.8)	(5.6)	(31.8)	
5′–GGAG / CCUC	(33.1)	(95.2)	(3.5)	(11.5)	

*For each sequence, the top number is derived from t_m^{-1} vs. log (C_t/4) plots, and the bottom number is derived from averaging fits to melting curves (21). Significant figures are given beyond error estimates to allow accurate calculation of t_m and other parameters.

†Measured $\Delta\Delta G^\circ_{37} = \Delta G^\circ_{37}$ (hairpin + short oligomer) − ΔG°_{37} (short oligomer duplex), e.g., $\Delta\Delta G^\circ_{37} = \Delta G^\circ_{37}$ ($\frac{\text{GGACGAGUGGC}^{\text{G}}_{\text{C}}}{\text{CCUG/CUCACCG}_{\text{A}}^{\text{A}}}$) − ΔG°_{37} ($\frac{\text{GGAC}}{\text{CCUG}}$). Measured ΔG°_{37} values are from t_m^{-1} vs. log(C_t/4) plots. Predicted ΔG°_{37} values (31) were used for $\frac{\text{GGAC}}{\text{CCUG}}$ and $\frac{\text{GGAG}}{\text{CCUC}}$ because experimental parameters were unreliable due to low melting temperatures. The experimental ΔG°_{37} for $\frac{\text{GGAC}}{\text{ACCUGU}}$ was used for this sequence and for $\frac{\text{GGAC}}{\text{ACCUG}}$ because the predicted ΔG°_{37} of $^{5'\text{-UG-3'}}_{\text{C}}$ is 0.0 kcal/mol (13). For oligomers with a 5′ unpaired nucleotide, the duplex was subtracted without adding the ΔG°_{37} for the 5′ dangling end because this interaction is expected to be different between binding of a short oligomer to another short oligomer and to the hairpin. This ΔG°_{37} is expected to be only ≈0.2 kcal/mol (13) and thus will not affect any conclusions. Errors in $\Delta\Delta G^\circ_{37}$ are estimated as 0.2 and 0.1 kcal/mol, respectively, when predicted and measured ΔG°_{37} values are used.

‡Values in parentheses are corrected (Corr.) for stacking of any unpaired 3′ nucleotide on the hairpin stem: $\Delta\Delta G^\circ_{37}$ (corrected) = $\Delta\Delta G^\circ_{37}$ (uncorrected) + ΔG°_{37} (predicted for displacing 3′ dangling end), where ΔG°_{37} (predicted for displacing 3′ dangling end) is the nearest-neighbor ΔG°_{37} for the 3′ mismatched nucleotide of the hairpin stem (13).

§The case of $\frac{\text{-GC-}}{\text{-C/G-}}$ (U) was attempted but did not exhibit a first transition melt. This was probably due to the dangling 3′-U forming a GU mismatch at the end of the hairpin, precluding binding of tetramer to the hairpin.

¶Numbers in parentheses are predicted values.

helix–helix interface illustrated in Fig. 1, this corresponds to 3.0 ± 0.2 kcal/mol more favorable free energy of duplex formation at 37°C, $\Delta\Delta G^\circ_{37}$, over that expected for the isolated (GGAC)·(GUCC) duplex (Table 1). The (GGAC)·(GUCC) duplex melts at too low a temperature and with too broad a transition for measurement of accurate thermodynamic parameters (see Fig. 1). Thus ΔG°_{37} was predicted from known nearest-neighbor parameters (31). The predicted ΔG°_{37} is consistent with parameters derived from fitting melting curves for (GGAC)·(GUCC) at ≈1 mM oligomer concentration and with parameters measured for the more stable duplex (GGAC)·(UGUCCA) (Table 1).

The sequence dependence of coaxial stacking was tested by switching the sequence at the interface to $\frac{\text{G-C}}{\text{C/G}}$, while keeping the rest of the sequence the same. In this case, binding of CUCC to the GGAG overhang adjacent to the hairpin stem is 4.3 kcal/mol more favorable at 37°C than expected for binding to the isolated GGAG tetramer (Table 1), a 1000-fold increase in binding constant. When contained in an intact RNA double helix, $\frac{\text{C-G}}{\text{G-C}}$ and $\frac{\text{G-C}}{\text{C-G}}$ nearest-neighbor interactions contribute −2.0 and −3.4 kcal/mol, respectively, to the free energy of duplex formation (31). Thus, in both cases, the interactions at the helix–helix interfaces contribute ≈1 kcal/mol more favorable free energy than if all nucleotides were connected by a phosphodiester linkage. Evidently, the base pairs at the helix–helix interface stack coaxially, and the lack of the phosphodiester linkage gives the nucleotides more freedom to maximize intermolecular interactions and/or reduces charge repulsion.

In a large RNA, the polynucleotide chains usually extend beyond a helix–helix interface. To determine the effect of chain extension on the thermodynamics of coaxial stacking, an unpaired nucleotide was appended to either or both sides of the $\frac{\text{C-G}}{\text{G/C}}$ interface. When GUCC is extended at the 5′ end to AGUCC or UGUCC, $\Delta\Delta G^\circ_{37}$ associated with duplex formation adjacent to the hairpin stem is −2.8 and −2.6 kcal/mol, respectively (Table 1). On average, this is only 0.3 kcal/mol less stable than for GUCC. When the hairpin stem is extended by an adenosine or uridine in the 3′ direction to form a CA or CU mismatch at the base of the stem, $\Delta\Delta G^\circ_{37}$ is reduced to −1.2 and −1.5 kcal/mol, respectively (Table 1). This can be accounted for by known favorable interactions of 3′ unpaired or mismatched nucleotides with an adjacent helix (13). Presumably, these interactions must be broken to allow the oligomer to bind tightly. Data for stacking of 3′ unpaired nucleotides can be used to estimate the ΔG°_{37} associated with breaking these interactions. These correction terms are 1.7 and 1.2 kcal/mol for $_{3'\text{-AC}}^{\text{G}}$ and $_{3'\text{-UC}}^{\text{G}}$ ends, respectively (13). Thus the corrected $\Delta\Delta G^\circ_{37}$ for coaxial stacking in the presence of a 3′ unpaired adenosine or uridine at the helix–helix interface is −2.9 and −2.7 kcal/mol, respectively, close to the values observed in the absence of a 3′ unpaired nucleotide. This corrected $\Delta\Delta G^\circ_{37}$ represents the magnitude of the favorable interactions at the helix–helix interface. An alternative model in which the 3′ unpaired nucleotide remains stacked by intercalation between the base pairs at the interface gives the same total free energy for the sum of the 3′ nucleotide and coaxial stacking.

To determine the effect of extending both sides of the helix–helix interface by one unpaired nucleotide, the binding of 5′-XGUCCA-3′ sequences to hairpin stems beginning $^{5'\text{-GGACG-}}_{\text{ZC-}}$ was studied (Table 1). The 3′ unpaired adenosine on 5′-XGUCCA-3′ raises the t_m sufficiently to allow accurate measurements of binding to both the hairpin overhang and the GGAC tetramer. After correction for displacing Z, the $\Delta\Delta G^\circ_{37}$ associated with coaxial stacking ranged from −1.3 to −2.6 kcal/mol, with an average of −2.0 kcal/mol for six sequences (Table 1). One of the hexamers, CGUCCA, has a 5′ cytidine that could compete with the terminal Watson–Crick base pair of the hairpin stem, leaving a flush interface with the last two 3′ nucleotides of the hairpin stem unpaired. If the two complexes with this hexamer are omitted from the analysis, the average $\Delta\Delta G^\circ_{37}$ for the interface is −1.8 kcal/mol. Thus, on average, extension of the helix–helix interface by an unpaired nucleotide on each side reduces the corrected $\Delta\Delta G^\circ_{37}$ for the $\frac{\text{C-G}}{\text{G/C}}$ interface to approximately the value for a $\frac{\text{C-G}}{\text{G-C}}$ nearest-neighbor interaction in an intact helix (31).

The effects of chain extension on the thermodynamics of the $\frac{\text{G-C}}{\text{C/G}}$ interface were explored by measuring the binding of 5′-UCUCC-3′ and 5′-ACUCC-3′ to hairpin stems beginning $^{5'\text{-GGAGC-}}_{\text{G-}}$ and of 5′-UCUCC-3′ binding to $^{5'\text{-GGAGC-}}_{\text{AG-}}$ (Table 1). In the latter case, the corrected $\Delta\Delta G^\circ_{37}$ for UCUCC binding is −2.6 kcal/mol, somewhat less than the ΔG°_{37} of −3.4 kcal/mol for a $\frac{\text{G-C}}{\text{C-G}}$ interaction in an intact helix (31). The corrected $\Delta\Delta G^\circ_{37}$ for 5′-CUCC-3′ binding to the overhang with the GA mismatch is −3.4 kcal/mol, compared with −4.3 kcal/mol in the absence of the adenosine. Thus the 3′ adenosine terminating the hairpin stem may affect the interactions at the helix–helix interface. Nevertheless, for both interfacial sequences studied, the ΔG°_{37} for base pairs in intact helixes are reasonable approximations for stacking interactions in the presence of extended chains at the interface.

Phylogenetic structures for ribosomal RNA (6, 32) were studied to determine the number of occurrences of helix–helix interfaces. $\frac{\text{C-G}}{\text{G/C}}$ occurs 63 times, and $\frac{\text{G-C}}{\text{C/G}}$ occurs 33 times. A 5′ unpaired uridine extending from the $\frac{\text{C-G}}{\text{G/C}}$ interface occurs 57 times, and an adenosine four times. As the 3′ extension of the $\frac{\text{C-G}}{\text{G/C}}$ interface, adenosine is present four times, and uridine is present 12 times. The $\frac{\text{G-C}}{\text{C/G}}$ helix–helix interface is extended at the 5′ position by an adenosine 17 times and by a uridine

Table 2. Comparison of secondary structures predicted by computer and determined by phylogenetic comparisons

RNA	Phylogenetic helixes predicted, %		
	Helixes, no.	Optimal	Coaxial stacking + J–S reordered optimal
Small subunit rRNAs			
E. coli	65	57	72
Rat mitochondria	35	60	63
H. volcanii	62	65	74
C. r. chloroplast	66	58	74
Introns			
LSU (I)	17	88	88
Yeast OX5ᵉ(I)	11	73	82
ND1 (I)	23	43	43
T4D (I)	12	67	83
Yeast A1 (II)	32 (31)	69 (48)	63 (48)
Yeast A5 (II)	30 (29)	87 (86)	87 (86)
Yeast B1 (II)	31 (31)	71 (71)	81 (81)
Average, %		67*	74
Total helixes	384 (382)	247 (239)	279 (273)
Helixes predicted, %		64 (63)	73 (71)

(*H.*) *Halobacterium volcanii*; *C. r.*, *Chlamydomonas reinhardtii*; LSU, self-splicing group I LSU intron in *Tetrahymena thermophila*; ND1, *Podospora anserina* ND1; T4D, bacteriophage T4 *td*; J-S, Jacobson–Stockmayer.

*This average was 66% when the parameters of Jaeger *et al.* (14) were used.

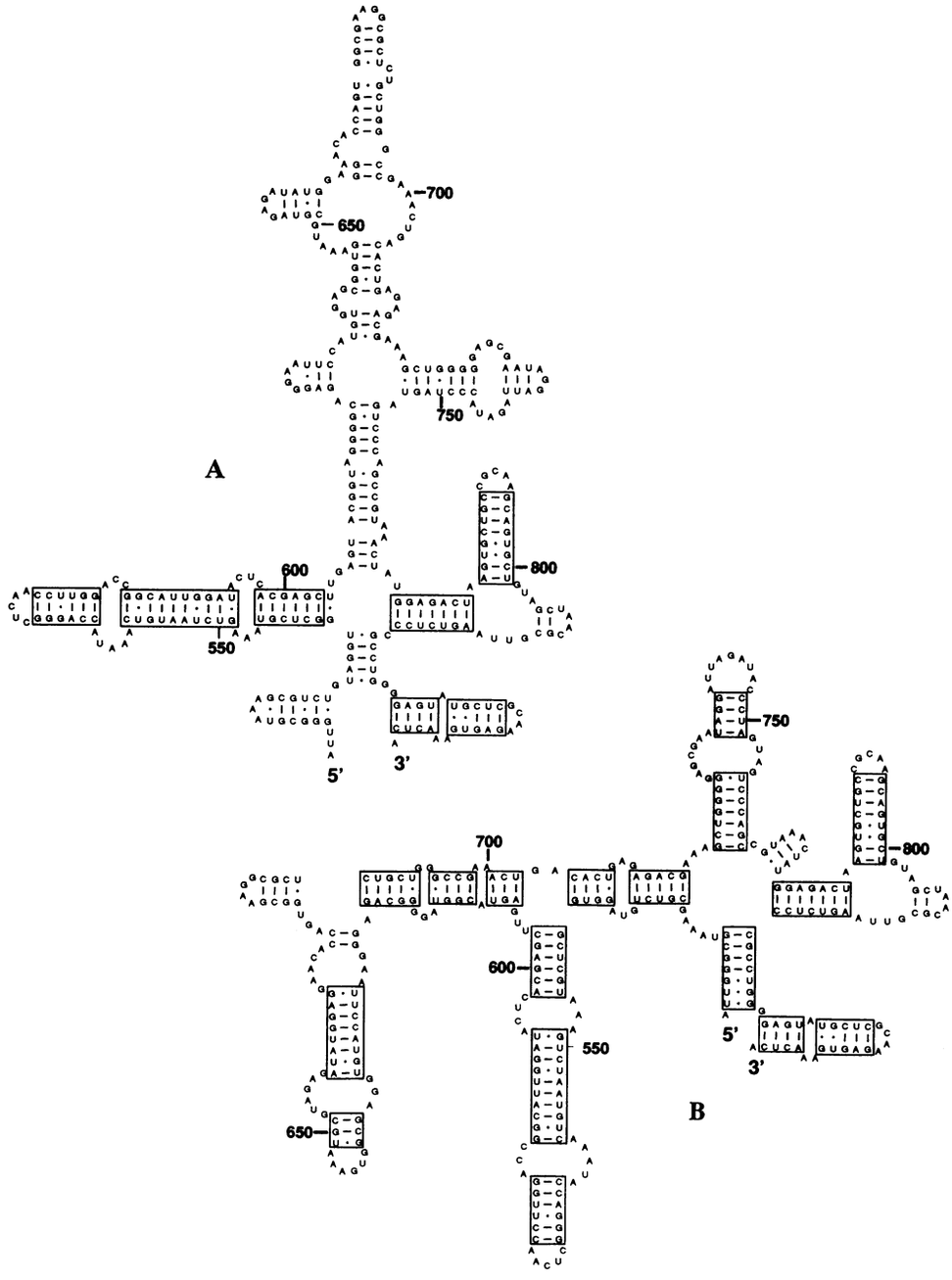

FIG. 2. Effect of coaxial stacking on prediction of structure for domain 2 of *Chlamydomonas reinhardtii* 16S-like rRNA. (*A*) Predicted structure in absence of coaxial stacking. (*B*) Predicted structure when coaxial stacking is included. Boxed regions are helixes of at least 3 bp in the phylogenetically determined structure (6).

14 times. Adenosine as the 3′ extension of this helix–helix interface occurs 11 times. Thus the interfaces listed in Table 1 are representative of those that occur naturally.

The experimental results suggest coaxial stacking at a helix–helix interface of Watson–Crick pairs provides a favorable free energy. Presumably, coaxial stacking involving a GA mismatch like that observed between the anticodon and D stems of tRNA (2, 3) will also provide a favorable free energy because this motif occurs commonly in phylogenetically determined secondary structures (6–8, 32). The effect of coaxial stacking of adjacent helixes and helixes separated by a GA mismatch on prediction of RNA secondary structure was tested in a preliminary way by including favorable free energies for these motifs in a program that recalculates free energies for sets of secondary structures. In these calculations, helix–helix interfaces with only Watson–Crick pairs in multibranch loops were assigned a bonus free energy equivalent to that for the same nearest-neighbor interaction in an intact helix (31). Helix–helix interfaces with a GA mismatch in multibranch loops were assigned a bonus free energy of −1 kcal/mol in addition to the usual free energy for stacking. This is based on stabilities of GA mismatches in internal loops and hairpins (25, 30, 33). A Jacobson–Stockmayer function (34) was also used to calculate ΔG° for multibranch loops having more than six unpaired nucleotides. For each RNA sequence, a comparison was made between the secondary structure deduced from phylogenetic analysis and that predicted by free energy minimization with the Zuker algorithm (12), which does not include coaxial stacking (optimal) or with that predicted after reordering a set of 20 structures generated by the Zuker algorithm. As shown in Table 2, incorporation of coaxial stacking and the Jacobson–Stockmayer function increased average agreement between predicted and phylogenetically deduced secondary structures of 11 RNAs from 67 to 74% (e.g., see Fig. 2). The percentage of helixes found in the whole data set increased from 64 to 73%; this percentage was 64 and 71%, respectively, when only the Jacobson–Stockmayer function or only coaxial stacking of Watson–Crick pairs was included in the recalculation. This is the largest single improvement in pre-

dictions since the complete set of free energy parameters for nearest-neighbor base pair combinations was measured (31, 35). When parameters for coaxial stacking were made 1 kcal/mol less favorable, predictions for small subunit rRNAs, the LSU group I intron (I), and yeast B1 (II) were intermediate between the results with and without coaxial stacking shown in Table 2. When parameters for coaxial stacking were made 1 kcal/mol more favorable, predictions for these six RNAs were essentially the same as with the nearest-neighbor parameters.

The results suggest that coaxial stacking is important for determining RNA structure. It is probably particularly important for the stabilities of multibranch loops and may also be important for pseudoknots (36). Coaxial stacking may also be important for determining DNA structure, for example, at junctions (37, 38). A recent study has shown that DNA junctions are more stable when they contain unpaired nucleotides (38). The unpaired nucleotides could provide flexibility to allow the most favorable coaxial stacking. Presumably, there are also other interactions at helix–helix interfaces that will be important. For example, Michel and Westhof (9) have proposed base triples at helix junctions in the structures of group I introns. The model system used here can also provide thermodynamic parameters for such interactions. Once a reasonable model for secondary structure is available, knowledge of interactions at helix–helix interfaces should also aid modeling of three-dimensional structure because this depends largely on positioning of helixes.

Thermodynamic parameters for helix–helix interfaces may be useful for more than prediction of structure from sequence. For example, it has been shown that circularly permuted sequences of tRNA often form the same structure, even when the sequence is cleaved in the middle of a helix (39). Coaxial stacking at the cleavage site is probably a factor in stabilizing these structures. Thus data base searches for circularly permuted sequences (39) may benefit from thermodynamic rules that permit prediction of sequences likely to maintain a given three-dimensional shape. It has been shown that a series of short DNA primers can replace a single long primer to prime synthesis by DNA polymerase (40). Presumably this is due to coaxial stacking between the primers (40). Related approaches should also be applicable to design of antisense reagents or therapeutics targeted at RNA (41). An antisense sequence can be designed to take advantage of coaxial stacking with a helix in the target (Fig. 1) or with a helix generated by binding of a second antisense sequence. This design would enhance both the binding and specificity of the antisense sequence, facilitating the use of short sequences (42). Knowledge of coaxial stacking and other helix–helix interactions should help in designing such sequences.

We thank Drs. T. R. Krugh, E. Kool, L. Lindahl, E. Phizicky, J. Zengel, P. Hammond, and J. B. Olmsted for comments on the manuscript. This work was supported by National Institutes of Health grant GM22939 to D.H.T. M.Z. is a Fellow of the Canadian Institute for Advanced Research, and D.H.T. is a Guggenheim Fellow and American Cancer Society Scholar.

1. Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A. & Weiner, A. M. (1987) *Molecular Biology of the Gene* (Benjamin/Cummings, Menlo Park, CA).
2. 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.
3. Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. D., Clark, B. F. C. & Klug, A. (1974) *Nature (London)* **250**, 546–551.
4. Kim, S.-H. & Cech, T. R. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8788–8792.
5. James, B. D., Olsen, G. J. & Pace, N. R. (1989) *Methods Enzymol.* **180**, 227–239.
6. Gutell, R. R., Weiser, B., Woese, C. R. & Noller, H. F. (1985) *Prog. Nucleic Acid Res. Mol. Biol.* **32**, 155–195.
7. Moazed, D., Stern, S. & Noller, H. F. (1986) *J. Mol. Biol.* **187**, 399–416.
8. Gutell, R. R., Larsen, N. & Woese, C. R. (1994) *Microbiol. Rev.* **58**, 10–26.
9. Michel, F. & Westhof, E. (1990) *J. Mol. Biol.* **216**, 585–610.
10. Tinoco, I., Jr., Uhlenbeck, O. C. & Levine, M. D. (1971) *Nature (London)* **230**, 362–367.
11. DeLisi, C. & Crothers, D. M. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 2682–2685.
12. Zuker, M. (1989) *Science* **244**, 48–52.
13. Turner, D. H., Sugimoto, N. & Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17**, 167–192.
14. Jaeger, J. A., Turner, D. H. & Zuker, M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7706–7710.
15. Zuker, M., Jaeger, J. A. & Turner, D. H. (1991) *Nucleic Acids Res.* **19**, 2707–2714.
16. Wu, T., Ogilvie, K. K. & Pon, R. T. (1989) *Nucleic Acids Res.* **17**, 3501–3517.
17. Wang, Y. Y., Lyttle, M. H. & Borer, P. N. (1990) *Nucleic Acids Res.* **18**, 3347–3352.
18. Kierzek, R., Caruthers, M. H., Longfellow, C. E., Swinton, D., Turner, D. H. & Freier, S. M. (1986) *Biochemistry* **25**, 7840–7846.
19. Uhlenbeck, O. C. & Cameron, V. (1977) *Nucleic Acids Res.* **4**, 85–98.
20. England, T. E. & Uhlenbeck, O. H. (1978) *Biochemistry* **17**, 2069–2076.
21. Petersheim, M. & Turner, D. H. (1983) *Biochemistry* **22**, 256–263.
22. He, L., Kierzek, R., SantaLucia, J., Jr., Walter, A. E. & Turner, D. H. (1991) *Biochemistry* **30**, 11124–11132.
23. Peritz, A. E., Kierzek, R., Sugimoto, N. & Turner, D. H. (1991) *Biochemistry* **30**, 6428–6436.
24. SantaLucia, J., Jr., Kierzek, R. & Turner, D. H. (1991) *Biochemistry* **30**, 8242–8251.
25. SantaLucia, J., Jr., Kierzek, R. & Turner, D. H. (1991) *J. Am. Chem. Soc.* **113**, 4313–4322.
26. Groebe, D. R. & Uhlenbeck, O. C. (1988) *Nucleic Acids Res.* **16**, 11725–11735.
27. Gralla, J. & Crothers, D. M. (1973) *J. Mol. Biol.* **73**, 497–511.
28. Riesner, D., Maass, G., Thiebe, R., Philippsen, P. & Zachau, H. G. (1973) *Eur. J. Biochem.* **36**, 76–88.
29. Coutts, S. M., Gangloff, J. & Dirheimer, G. (1974) *Biochemistry* **13**, 3938–3948.
30. Serra, M. J., Lyttle, M. H., Axenson, T. J., Schadt, C. A. & Turner, D. H. (1993) *Nucleic Acids Res.* **21**, 3845–3849.
31. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T. & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9373–9377.
32. Gutell, R. R., Schnare, M. N. & Gray, M. W. (1990) *Nucleic Acids Res.* **18**, 2319–2330.
33. SantaLucia, J., Jr., Kierzek, R. & Turner, D. H. (1992) *Science* **256**, 217–219.
34. Jacobson, Y. & Stockmayer, W. H. (1950) *J. Chem. Phys.* **18**, 1600–1606.
35. Turner, D. H., Sugimoto, N., Jaeger, J. A., Longfellow, C. E., Freier, S. M. & Kierzek, R. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52**, 123–133.
36. Wyatt, J. R., Puglisi, J. D. & Tinoco, I., Jr. (1990) *J. Mol. Biol.* **214**, 455–470.
37. Lu, M., Guo, Q., Marky, C. A., Seeman, N. C. & Kallenbach, N. R. (1992) *J. Mol. Biol.* **223**, 781–789.
38. Leontis, N. B., Kwok, W. & Newman, J. S. (1991) *Nucleic Acids Res.* **19**, 759–766.
39. Pan, T., Gutell, R. R. & Uhlenbeck, O. C. (1991) *Science* **254**, 1361–1364.
40. Kieleczawa, J., Dunn, J. J. & Studier, F. W. (1992) *Science* **258**, 1787–1791.
41. Ecker, D. J., Vickers, T. A., Bruice, T. W., Freier, S. M., Jenison, R. D., Manoharan, M. & Zounes, M. (1992) *Science* **257**, 958–961.
42. Bevilacqua, P. C. & Turner, D. H. (1991) *Biochemistry* **30**, 10632–10640.