Thermodynamic Stabilities of Internal Loops with GU Closing Pairs in RNA†

Susan J. Schroeder and Douglas H. Turner*

Department of Chemistry, RC Box 270216, University of Rochester, Rochester, New York 14627-0216

Received March 9, 2001; Revised Manuscript Received July 6, 2001

ABSTRACT: Many internal loops that form tertiary contacts in natural RNAs have GU closing pairs; examples include the tetraloop receptor and P1 helix docking site in group I introns. Thus, thermodynamic parameters of internal loops with GU closing pairs can contribute to the prediction of both secondary and tertiary structure. Oligoribonucleotide duplexes containing small internal loops with GU closing pairs were studied by optical melting, one-dimensional imino proton NMR, and one-dimensional phosphorus NMR. The thermodynamic stabilities of asymmetric internal loops with GU closing pairs relative to those of loops with GC closing pairs may be explained by hydrogen bonds. In contrast, the free energy increments for symmetric internal loops of two noncanonical pairs with GU closing pairs relative to loops with GC closing pairs show much more sequence dependence. Imino proton and phosphorus NMR spectra suggest that some GA pairs adjacent to GU closing pairs may form an overall thermodynamically stable but non-A-form conformation.

Internal loops with GU closing pairs occur in RNA secondary structures and participate in tertiary and quaternary interactions. In small and large subunit rRNA, 42% of GU pairs occur at loop−helix junctions with a strong preference for the motif 5′UG/3′GA (*1*). For example, helix 44 in the *Escherichia coli* 16S ribosomal subunit has a series of consecutive GU and GA pairs; this helix forms part of the interface between the two subunits and the binding site for initiation factor 1 (*2*−*6*). In group I introns, GU pairs are 80% conserved as a closing pair in the J4/5 loop (*7*). The J4/5 loop is the docking site for the P1 helix (*8*, *9*), and therefore participates in the first step of self-splicing (*10*−*13*). In the crystal structure of the P4−P6 module of the *Tetrahymena thermophila* group I intron, GU closing pairs form the base for the AA platforms observed in the J6/6a loop and the J6a/6b tetraloop receptor (*14*, *15*). In the tetraloop receptor, GU closing pairs participate in the coordination of a potassium ion (*16*). Thus, internal loops closed by GU pairs are often important for RNA structure and function.

Previous thermodynamic studies of internal loops include few examples with GU closing pairs. In two cases, the stabilities of tandem GA pairs with GU closing pairs are 0.8 and 4.4 kcal/mol less stable than similar loops with GC closing pairs (*17*). In 2 × 2 loops[1] with GC closing pairs, UU pairs are more stable than AA pairs; in 2 × 2 loops with GU closing pairs, however, the order of stability is reversed (*18*). Thus, GU closing pairs can affect the sequence dependence of the thermodynamic stability of an internal loop.

This paper presents thermodynamic parameters for a variety of internal loops with GU closing pairs. The difference in stability between loops with GU and GC closing pairs depends on loop sequence and asymmetry. In a 1 × 3 and several 2 × 3 internal loops, the average difference in stability between loops with GU and GC closing pairs is 1.5 kcal/mol, which can be explained by a hydrogen bond model (*19*, *20*). In 2 × 2 loops, however, the difference in stability between loops with GU and GC closing pairs is more sequence-dependent, which suggests significant changes in stacking interactions and/or backbone conformations, in addition to a change in the number of hydrogen bonds. A downfield-shifted phosphorus resonance is observed in some one-dimensional phosphorus NMR spectra, suggesting that GA pairs adjacent to GU closing pairs have a non-A-form conformation in some contexts (*21*).

The results provide models for predicting stabilities of internal loops closed by GU pairs. These parameters can be incorporated into programs such as *mfold* for predicting RNA secondary structure (*22*−*25*). The results also contribute to understanding molecular recognition involving internal loops in tertiary structure formation. For example, thermodynamic stabilities of internal loops that model the J4/5 loop in different group I introns are compared with measurements of tertiary interactions at this site (*8*, *26*, *27*). Thus, the results should facilitate prediction of RNA secondary and three-dimensional structure.

MATERIALS AND METHODS

Oligoribonucleotides were synthesized and purified as previously described (*20*). All oligomers were greater than 90% pure. Optical melting experiments and imino proton

† This work supported by NIH Grant GM22939.

* To whom correspondence should be addressed. Telephone: (716) 275-3207. Fax: (716) 473-6889. E-mail: Turner@chem.rochester.edu.

[1] Abbreviations: EDTA, ethylenediaminetetraacetic acid; HPLC, high-pressure liquid chromatography; MES, 2-(*N*-morpholino)ethanesulfonic acid; *m* × *n* loop, an internal loop containing *m* nucleotides opposite *n* nucleotides, where *m* and *n* are integers; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; TLC, thin-layer chromatography; T_M, melting temperature in kelvin; T_m, melting temperature in degrees Celcius.

10.1021/bi010489o CCC: $20.00
Published on Web 08/30/2001

Table 1: Thermodynamic Parameters for Internal Loop and Duplex Formation at 37 °C[a]

	$\Delta G^\circ_{loop,37}$ (kcal/mol)	$-\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m (°C) (1.0×10^{-4} M)
3X3 Loop					
5'CCAG**CCAAG**UCCU [b,c,f,i] 3'GGUU**GACGU**AGGA	0.1±0.3 (1.6)	10.1±0.1	94.5±1.9	272.1±5.9	49.2
2X3 Loops					
5'GAGU **AA** CGAC [i] 3'CUC**GAAGG**CUG	1.8±0.5 (1.9)	8.0±0.1	83.7±2.1	244.2±6.6	42.4
5'GAGU **GA** UGAC [b,e,i,j] 3'CUC**GAAGG**CUG	1.9±1.0 (2.0)	7.0±0.1	75.6±2.8	221.2±9.2	38.8
5'GAGC **AA** UGAC [e] 3'CUC**GAAGG**CUG	2.2±0.5 (2.3)	7.5±0.1	79.5±2.6	232.2±8.2	40.9
5'CUGU **AU** GACG [d,e,i] 3'GAC**GAAUC**UGC	3.0±0.5 (2.9)	6.6±0.1	77.7±2.2	229.3±7.0	37.2
5'GAGU **AA** CGAC [e,i,j] 3'CUC**GAAAG**CUG	3.2±0.5 (3.2)	6.6±0.1	71.1±3.6	207.9±11.6	37.4
5'GAGU **AA** UGAC [e] 3'CUC**GAAGG**CUG	3.2±1.0 (3.0)	5.6±0.1	60.7±2.5	177.4±8.2	32.5
5'GAGU **GA** UGAC [f] 3'CUC**GAAAG**CUG	3.3±1.0 (3.0)	5.5±0.1	50.0±3.0	143.4±10.0	30.9
5'GAGC **AA** UGAC 3'CUC**GAAAG**CUG	3.6±0.5 (3.2)	6.2±0.1	64.1±2.7	186.6±8.8	35.4
5'GAGU **AG** UGAC [f] 3'CUC**GAAAG**CUG	4.0±1.0 (3.0)	4.8±0.1	46.1±1.1	133.2±3.8	26.0
5'GAGU **AA** UGAC 3'CUC**GAAAG**CUG	4.1±1.0 (3.8)	4.7±0.1	43.5±1.0	125.1±3.5	24.6
1X3 Loop					
5'GAGU **G** UGAC 3'CUC**G AAGG**CUG	3.8±1.0 (3.8)	5.1±0.1	51.5±2.3	149.7±7.5	28.6
2X2 Loops					
5'CU**GUGAU**GAC [f,j] 3'GAC**GAGG**CUG	0.2±1.4 (0.8)	8.4±0.1	89.8±2.3	262.3±7.2	43.7
5'GG**UAG**ACC 3'CC**AGAU**GG	0.9±0.3 (0.7)	5.1±0.1	66.6±2.6	198.2±8.5	34.4
5'GG**AAGU**CC [e,j] 3'CC**UGAA**GG	1.7±0.3 (0.3)	4.6±0.1	55.1±1.8	162.7±5.8	31.2
5'GAG**UAAC**GAC 3'CUC**GAAG**CUG	1.9±0.5 (1.8)	7.9±0.1	86.8±3.9	254.5±12.5	41.9
5'GAG**UGAU**GAC [e,j] 3'CUC**GAAG**CUG	2.2±1.0 (1.1)	6.7±0.1	77.4±2.8	227.9±8.9	37.6
5'GC**GAGU**GC [e,j] 3'CG**UGAG**CG	2.6±0.6 (0.3)	3.8±0.1	47.7±1.6	141.4±5.2	25.5
5'GAG**CAAU**GAC 3'CUC**GAAG**CUG	2.7±0.5 (2.0)	7.1±0.1	77.4±2.9	226.6±9.3	42.8
5'CU**GUAGG**CAG [e,j] 3'GAC**GGAU**GUC	3.4±0.8 (0.7)	5.5±0.1	82.2±4.1	247.3±13.3	36.1
5'GAG**UAGU**GAC 3'CUC**GAAG**CUG	3.6±1.0 (2.3)	5.2±0.1	61.6±2.9	181.9±9.6	30.6
5'GAG**UAAU**GAC [i] 3'CUC**GAAG**CUG	3.6±1.0 (2.5)	5.2±0.1	62.1±1.4	183.4±4.7	30.6
Stem duplexes					
5'CCAGCGUCCU 3'GGUUGUAGGA		11.6±0.2	87.9±3.4	246.0±10.3	55.9
5'GAGGAC 3'CUCCUG		8.5±0.1	57.3±1.4	157.1±4.5	48.3
5'GAGUUGAC [g] 3'CUCGGCUG		(8.4±0.2)	(74.8±7.5)	(214.2±5.9)	(44.8)
5'GAGUUGAG [h] 3'CUCGGCUC		8.2	73.9	211.8	44.2

[a] Melting experiments were carried out in 1 M NaCl, 10 or 20 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 7 buffer. For each type of loop, sequences are listed in order of internal loop free energy, calculated according to eq 1a. Thermodynamic parameters used in eq 1a and this table were obtained by fitting plots of inverse melting temperature, T_M^{-1}, versus natural $\log(C_T/n)$ to the equation (*73*) $T_M^{-1} = (R/\Delta H^\circ)\ln(C_T/n) + \Delta S^\circ/\Delta H^\circ$, where R is the gas constant, C_T is the total strand concentration, and n is 1 or 4 for self-complementary or non-self-complementary duplexes, respectively. Predicted values for ΔG°_{loop} are listed in parentheses below the experimental values and calculated according to the values in Table 2 and the rules in ref *22* for 2 × 2 loops and in ref *20* for 2 × 3 and 1 × 3 loops (see the discussion in the text). Listed errors are standard deviations from reported measurements assuming no correlation of errors in the slope and intercept and are therefore overestimates of this source of error. Estimated errors from all sources are ±10%, ±10%, ±2%, and ±1 °C for ΔH°, ΔS°, ΔG°, and T_m, respectively. Significant figures beyond error estimates are given to allow accurate calculation of T_M and other parameters. Three duplexes containing other 1 × 3 loops with GU closing pairs, 5′GAG**UAU**GAC/3′CUC**GAAGG**CUG, 5′GAG**UAU**GAC/3′CUC**GAAAG**CUG, and 5′GAG**UGU**GAC/3′CUC**GAAAG**CUG, do not exhibit two-state melting transitions. Stabilities for the 5′U**AA**U/3′G**GAA**G and 5′U**AG**U/3′G**GAA**G loops are not reported because the stability of self-complementary duplex formation for the 3′CUC**GGAAG**CUG sequence is greater than the stabilities for the combinations of two non-self-complementary oligomers. [b] These duplexes were also melted in 10 mM $MgCl_2$, 0.15 M KCl, 10 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 7.0 buffer at one or more concentrations. No significant change in duplex stability relative to that in 1 M NaCl buffer was observed. [c] These duplexes were also melted in 1 M NaCl, 50 mM MES, 0.5 mM Na_2EDTA, pH 5.5 buffer at one or more concentrations. No significant change in duplex stability relative to that in pH 7.0 buffer was observed. [d] These duplexes were also melted in 1 M KCl, 20 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 7 buffer at one or more concentrations. No significant change in duplex stability relative to that in 1 M NaCl buffer was observed. [e] These duplexes were studied by variable-temperature imino proton NMR (500 MHz) and one-dimensional difference NOE NMR in 80 mM NaCl, 10 mM phosphates, 0.5 mM Na_2EDTA, pH 6.7 buffer and 10% D_2O. [f] These duplexes were studied by variable temperature imino proton NMR (500 MHz) in 80 mM NaCl, 10 mM phosphates, 0.5 mM Na_2EDTA, pH 6.7 buffer and 10% D_2O. [g] Melting curves for the 5′GAGUUGAC/3′CUCGGCUG duplex do not exhibit two-state behavior. Therefore, the thermodynamic parameters are estimated from the experimental data for a similar duplex, 5′GAGUUGAG/3′CUCGGCUC (*32*), by adding the predicted difference (*19*) between the last nearest neighbor pair. [h] From ref *32*. [i] These loops have the same sequence as J4/5, J6/6a, or J6a/6b loops that occur naturally in the *T. thermophila*, *P. carinii*, or *C. albicans* group I introns (*7*, *26*, *74*, *75*). [j] These duplexes were studied by one-dimensional phosphorus NMR in 80 mM NaCl, 10 mM phosphates, 0.5 mM Na_2EDTA, pH 6.7 buffer and 99% D_2O.

NMR experiments were conducted as previously described (*20*). The standard melt buffer consisted of 1 M NaCl, 10 or 20 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 7 buffer. Some samples were also melted in 1 M NaCl, 50 mM MES, 0.5 mM Na_2EDTA, pH 5.5 buffer; 1 M KCl, 10 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 7 buffer; or 10 mM $MgCl_2$, 0.15 M KCl, 10 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 7 buffer. The ^{1}H-decoupled one-dimensional phosphorus spectra were collected at 202 MHz with a sweep width of 2.4 kHz and WALTZ decoupling. Samples in D_2O were referenced to the phosphate buffer peak (pH 6.7) at 0.0 ppm; the data were multiplied by a 5−12 Hz exponential line-broadening function. Additional experimental details are provided in the Supporting Information.

RESULTS

Thermodynamic Parameters. Table 1 lists thermodynamic parameters for the formation of duplexes and for internal loop formation. Duplexes are grouped by the type of loop and listed in order of decreasing loop stability. The thermo-

dynamic parameters from T_M^{-1} versus ln(C_T/n) plots and from curve fits of the data agree within 15% (see the Supporting Information), which is consistent with the two-state model assumption (*28*−*30*). As noted in footnotes b−d of Table 1, duplexes melted in other buffers exhibit no difference in stability as compared to those in 1 M NaCl, pH 7 buffer.

The thermodynamic parameters for formation of the internal loops are calculated according to the following equation (*31*):

$$\Delta G^\circ_{\text{loop}} = \Delta G^\circ_{\text{duplex with loop}} - (\Delta G^\circ_{\text{duplex without loop}} - \Delta G^\circ_{\text{interrupted base pair}}) \quad (1a)$$

For example,

$$\Delta G^\circ \begin{smallmatrix} 5'U\,{}^{GA}\,U \\ 3'G_{AAG}G \end{smallmatrix} = \Delta G^\circ \begin{smallmatrix} 5'GAGU\,{}^{GA}\,UGAC \\ 3'CUCG_{AAG}GCUG \end{smallmatrix} - \left(\Delta G^\circ \begin{smallmatrix} 5'GAGUUGAC \\ 3'CUCGGCUG \end{smallmatrix} - \Delta G^\circ \begin{smallmatrix} 5'UU \\ 3'GG \end{smallmatrix} \right) \quad (1b)$$

Values for $\Delta H^\circ_{\text{loop}}$ and $\Delta S^\circ_{\text{loop}}$ are similarly derived and included in the Supporting Information. The last term in eq 1a accounts for the nearest neighbor stacking interaction present in the uninterrupted helix but absent in the helix containing an internal loop. These nearest neighbor parameters, taken from an analysis of nearest neighbor GU pairs (*22*, *32*), have errors as large as 0.96 kcal/mol, which accounts for the large error in some $\Delta G^\circ_{\text{loop}}$ parameters.

Listed in parentheses in Table 1 are the free energies of loop formation predicted by the models described below. For 3 × 3, 2 × 3, and 1 × 3 loops, the thermodynamic stability of an internal loop is predicted with the following equation (*20*, *22*, *33*):

$$\Delta G^\circ_{\text{loop (predicted)}} = \Delta G^\circ_{\text{loop initiation}} + \Delta G^\circ_{\text{asymmetry penalty}} + \Delta G^\circ_{\text{AU/GU penalty}} + \Delta G^\circ_{\text{GA/GG/UU bonus}} \quad (2)$$

The loop initiation term accounts for the penalty of forming a loop closed by two GC pairs and depends on loop size (*20*, *22*). The asymmetry penalty increases with an increasing difference in the number of nucleotides on each side of the loop (*34*). The AU/GU penalty accounts for the different number of hydrogen bonds in AU or GU pairs versus GC pairs (*19*, *20*) and has a value of 0.2 kcal/mol in addition to a 0.45 kcal/mol penalty for terminating a helix with an AU pair (*18*, *22*). The GA/GG/UU bonus accounts for potential stabilizing effects of GA, GG, and UU pairs when these noncanonical pairs are adjacent to the closing pairs in an internal loop (*18*, *33*, *35*, *36*). This term is zero for 1 × 3 loops (*20*). The rules for approximating stabilities of symmetric 2 × 2 loops have been previously summarized (*18*, *22*, *37*, *38*). The loop free energies for 2 × 2 loops presented in this paper suggest that the dependence on the GU and AU closing pairs may be more idiosyncratic than previously predicted, however.

The values in Table 1 for $\Delta G^\circ_{\text{loop}}$ can be compared with values for previously measured loops with one or two GC closing pairs and identical loop sequences (see the Supporting Information and refs *18*, *20*, *22*, and *37*). Ten 2 × 3 loops with one or two GU closing pairs are less stable than similar loops with two GC closing pairs by average values of 0.6 ± 0.2 or 1.5 ± 0.2 kcal/mol, respectively. The AU/GU penalty adds a predicted value of 0.65 or 1.3 kcal/mol for loops with one or two GU closing pairs, respectively, which is within experimental error of the measured values. Similarly, the 1 × 3 loop with two GU pairs is 1.4 kcal/mol less favorable than the average value for 1 × 3 loops with GC closing pairs (see the Supporting Information and ref *20*). For eight 2 × 2 loops, however, the effect of changing the two closing pairs from GU to GC varies from 0.8 to 4.4 kcal/mol. Thus, for 2 × 2 loops, the changes in stability between GU and GC closing pairs may be due to not only one fewer hydrogen bond per closing base pair but also changes in base stacking interactions and/or backbone conformations.

Table 2: Free Energy Increments for Symmetric 2 × 2 Purine Loops[a]

closing base pair	$\begin{smallmatrix}GA\\AG\end{smallmatrix}$	$\begin{smallmatrix}AG\\GA\end{smallmatrix}$	$\begin{smallmatrix}GG\\GG\end{smallmatrix}$	$\begin{smallmatrix}AA\\AA\end{smallmatrix}$
$\begin{smallmatrix}5'G\\3'C\end{smallmatrix}$	−2.6[b]	−1.3[b]	(0.8)[b]	1.5[b]
$\begin{smallmatrix}5'C\\3'G\end{smallmatrix}$	−0.7[b]	−0.7[b]	0.8[b]	1.3[b]
$\begin{smallmatrix}5'A\\3'U\end{smallmatrix}$	0.3[b]	1.7	1.9[c]	2.8[b]
$\begin{smallmatrix}5'U\\3'A\end{smallmatrix}$	0.7[b]	0.9	2.3[c]	2.8[b]
$\begin{smallmatrix}5'G\\3'U\end{smallmatrix}$	1.8[d]	2.6	(2.1)	(4.1)
$\begin{smallmatrix}5'U\\3'G\end{smallmatrix}$	0.1[d]	3.4	(2.1)	2.1[e]

[a] See footnote a of Table 1. Values in parentheses are predicted values. The predicted free energy increments for 5′G**GG**U/3′U**GG**G and 5′U**GG**G/3′G**GG**U are the average of the free energy increments for 5′A**GG**U/3′U**GG**A and 5′U**GG**A/3′A**GG**U. The prediction for 5′G**AA**U/3′U**AA**G is based on the free energy increments for related nonsymmetric 2 × 2 loops (Table 1) and equations in refs *33* and *37*. The (5′GC**GAAU**GC)$_2$ sequence did not exhibit a transition in the melting curve at 1.4 × 10^{-4} M, indicating the duplex is less stable than the (5′GC**UAAG**GC)$_2$ duplex (*18*). [b] Values are from ref *22*. Original data are from refs *17*, *18*, and *35* but recalculated with nearest neighbors from the INN-HB parameters (*19*). When there is more than one measurement of a loop, the average value is given. Values in parentheses are predicted stabilities. [c] From ref *36*. [d] Values are from ref *17* and corrected for the new nearest neighbor parameters (*19*). [e] Values are from ref *18* and corrected for the new nearest neighbor parameters (*19*).

Table 2 summarizes the experimental thermodynamic parameters for symmetric 2 × 2 purine internal loops with different closing pairs. When the closing pairs change from $\begin{smallmatrix}5'C\\3'G\end{smallmatrix}$ to $\begin{smallmatrix}5'U\\3'A\end{smallmatrix}$, the changes in loop free energy for all the symmetric purine 2 × 2 loops are consistent with the hydrogen bonding model. In contrast, when the closing pairs are changed from $\begin{smallmatrix}5'G\\3'C\end{smallmatrix}$ to $\begin{smallmatrix}5'A\\3'U\end{smallmatrix}$ or $\begin{smallmatrix}5'G\\3'U\end{smallmatrix}$, the changes in thermodynamic stability are not consistent with a simple hydrogen bonding model. For $\begin{smallmatrix}5'GA\\3'AG\end{smallmatrix}$ and $\begin{smallmatrix}5'AG\\3'GA\end{smallmatrix}$, the loop free energy changes by 3 and 4 kcal/mol on average when the closing pairs change from $\begin{smallmatrix}5'G\\3'C\end{smallmatrix}$ to $\begin{smallmatrix}5'A\\3'U\end{smallmatrix}$ and from $\begin{smallmatrix}5'G\\3'C\end{smallmatrix}$ to $\begin{smallmatrix}5'G\\3'U\end{smallmatrix}$, respectively. Thus, base stacking and electrostatic interactions also influence the loop free energies of symmetric 2 × 2 loops closed with a 5′-purine.

Imino Proton NMR. Imino protons in AU and GC pairs typically resonate in the regions between 13.5 and 15 ppm and between 11 and 13.5 ppm, respectively (*39*). Formation of expected Watson−Crick pairs was confirmed by imino proton NMR, and tentative assignments were made by one-dimensional NOE difference spectra (some examples are shown in Figures 1−3; also see footnotes e and f of Table 1 and the Supporting Information). GU pairs in a helix contain two imino protons that typically resonate between 10.5 and 12 ppm and give a strong NOE (*32*, *40*, *41*). Imino protons in imino hydrogen-bonded GA pairs and sheared GA pairs in RNA internal loops typically resonate near 12.5−

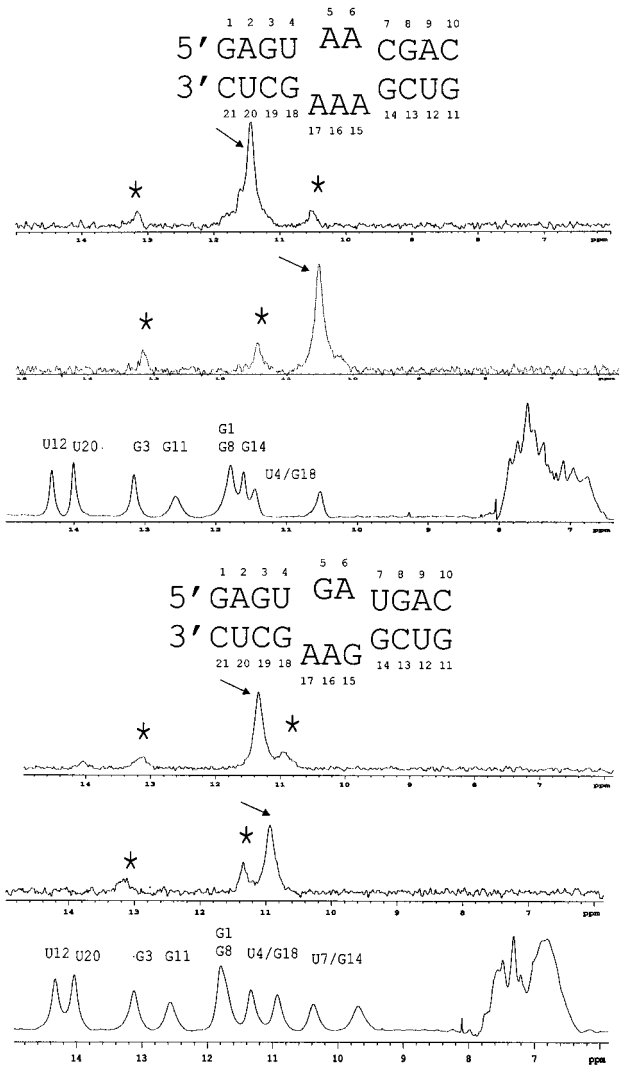

FIGURE 1: One-dimensional NOE difference NMR spectra (6−15 ppm) of 0.6 mM 5′GAG**UAAC**GAC/3′CUC**GAAAG**CUG at 1 °C and 2 mM 5′GAG**UGAU**GAC/3′CUC**GAAGG**CUG at 5 °C in 80 mM NaCl, 10 mM phosphate, 0.5 mM Na_2EDTA, pH 6.7 buffer. Asterisks denote observed NOE peaks when the resonance marked with an arrow is saturated. Tentative assignments are marked in the off-resonance spectrum. Spectra were acquired with 3 s saturation at the following frequencies: 11.45 and 10.51 ppm (top) and 11.32 and 10.94 ppm (bottom).

11.5 and 10 ppm, respectively (*17*, *42*). Thus, peaks in the upfield region of the spectra (<12.5 ppm) provide information about the GU closing pairs and the imino protons in the loops. The absence of a peak in the imino proton spectrum does not necessarily imply that a hydrogen-bonded GU or GA pair does not form; the imino protons in the pair can exchange with water if the structure is partly flexible.

Figure 1 shows some of the results of one-dimensional NOE experiments for the 5′U**AA**C/3′G**AAA**G and 5′U**GA**U/3′G**AAG**G loops. The peaks at 10.5 and 11.4 ppm (5′U**AA**C/3′G**AAA**G) or 10.9 and 11.3 ppm (5′U**GA**U/3′G**AAG**G) give mutual NOEs, which is characteristic of a GU pair (*40*, *43*−*45*). Each of these peaks has an additional NOE to a peak that is tentatively assigned to the adjacent GC pair in the stem. In a hairpin construct containing the 5′U**GA**U/3′G**AAG**G loop, weak NOEs were observed between G14 and U7 imino protons in a two-dimensional NOESY experiment (data not shown). These NOEs confirm the formation of the GU closing pairs in these 2 × 3 internal loops.

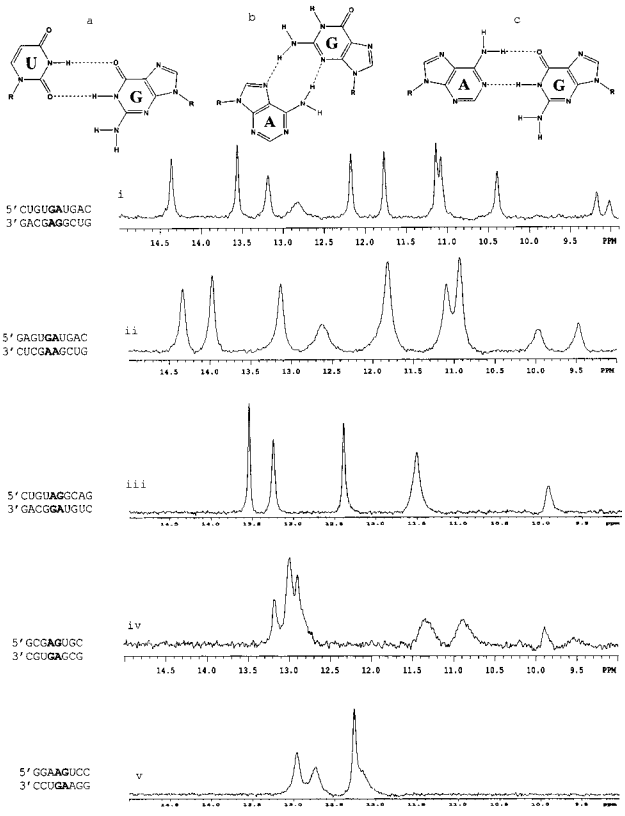

FIGURE 2: Diagrams of (a) a GU wobble pair, (b) a sheared GA pair, and (c) an imino GA pair (*77*, *78*). Imino protons in GU pairs, sheared GA pairs, and imino hydrogen-bonded GA pairs in RNA internal loops typically resonate near 10.5−12, 10, and 11.5 ppm, respectively (*17*, *32*, *40*−*42*). Imino proton NMR spectra (9−15 ppm) in 80 mM NaCl, 10 mM phosphate, 0.5 mM Na_2EDTA, pH 6.7 buffer. The single spectrum shown from a variable-temperature study was selected for the best dispersion: (i) 5′CUG**UGAU**GAC/3′GAC**GAGG**CUG, 0.2 mM, 1 °C, $\Delta G°_{loop\ 37} = 0.2$ kcal/mol; (ii) 5′GAG**UGAU**GAC/3′CUC**GAAG**CUG, 0.6 mM, 1 °C, $\Delta G°_{loop\ 37} = 2.2$ kcal/mol; (iii) $(5′CUG\mathbf{UAGG}CAG)_2$, 0.6 mM, 1 °C, $\Delta G°_{loop\ 37} = 3.4$ kcal/mol; (iv) $(5′GC\mathbf{GAGU}GC)_2$, 0.4 mM, 10 °C, $\Delta G°_{loop\ 37} = 2.6$ kcal/mol; and (v) $(5′GG\mathbf{AAGU}CC)_2$, 0.5 mM, 5 °C, $\Delta G°_{loop\ 37} = 1.7$ kcal/mol.

Figure 2 shows the structures of GU, sheared GA, and imino hydrogen-bonded GA pairs and imino proton spectra of several duplexes containing 2 × 2 loops with GA pairs and GU or AU closing pairs. The upfield resonances provide some information about the structures of the internal loops. The spectra of duplexes with 5′U**GA**U/3′G**AG**G and 5′U**GA**U/3′G**AA**G loops (spectra i and ii of Figure 2) exhibit resonances near 10 ppm that are consistent with either a proton in a sheared GA pair or a GU closing pair. The peaks below 10 ppm are difficult to assign but may reflect an unusual amino proton (*46*). For self-complementary duplexes, each resonance can be attributed to two imino protons due to symmetry. In spectrum v of Figure 2, the U imino protons in AU closing pairs adjacent to tandem imino GA pairs have an unusual upfield chemical shift (12.7 ppm); this unusual chemical shift was also observed for the U imino protons in AU closing pairs adjacent to tandem sheared GA pairs (*42*). In the $(5′CUG\mathbf{UAGG}CAG)_2$ duplex, NOEs between the resonances at 10 and 11.5 ppm confirmed the formation of the GU closing pairs (Figure 3); however, the G6 imino resonances are not seen, prohibiting the assignment of sheared or imino GA pairs. The G6 imino proton apparently exchanges with water; thus, the loop may be partly flexible.

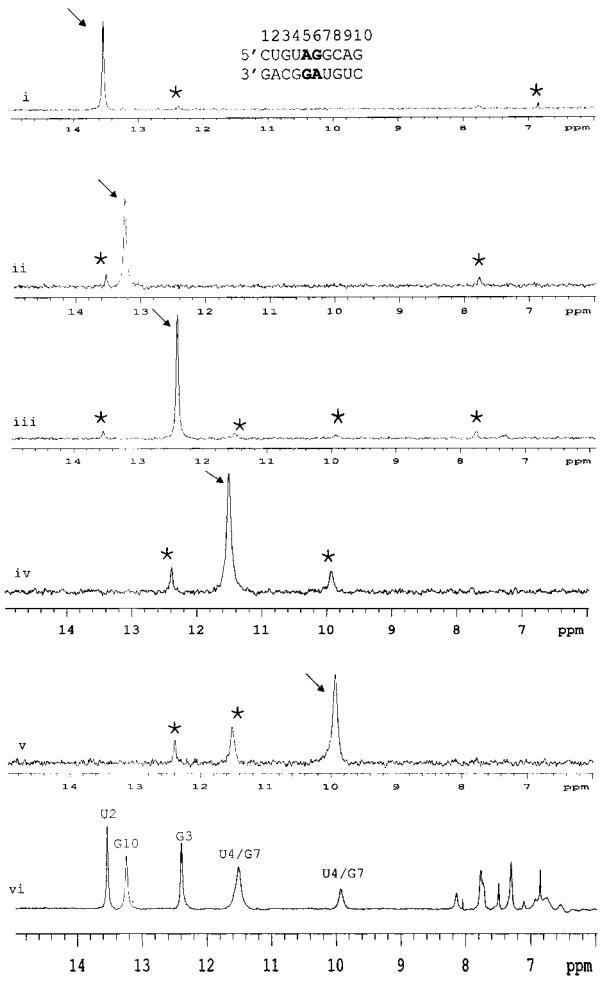

FIGURE 3: One-dimensional NOE difference NMR spectra (6−15 ppm) of 0.6 mM (5′CUG**UAGG**CAG)$_2$ at 1 °C in 80 mM NaCl, 10 mM phosphate, 0.5 mM Na_2EDTA, pH 6.7 buffer. Asterisks denote observed NOE peaks when the resonance marked with an arrow is saturated. Spectra were acquired with 3 s saturation at the following frequencies: (i) 13.52, (ii) 13.24, (iii) 12.38, (iv) 11.49, and (v) 9.93 ppm and (vi) off-resonance. Tentative assignments are marked in spectrum vi, and numbering of the bases is shown at the top.

The (5′U**AG**G)$_2$ loop in (5′CUG**UAGG**CAG)$_2$ is 4.1 kcal/mol less stable than the (5′C**AG**G)$_2$ loop (Table 2); this large thermodynamic difference may also imply a less structured, more flexible loop.

Phosphorus NMR. Figure 4 shows sample one-dimensional ^{1}H-decoupled ^{31}P NMR spectra of duplexes with GU closing pairs. Typical A-form RNA shows a dispersion of ~1.5 ppm (*47*); an example is shown in spectrum ii (also see footnote j of Table 1 and the Supporting Information). Some duplexes with GU closing pairs and GA pairs, however, have unusual downfield phosphorus resonances and thus a wider than normal dispersion of resonances. For example, the 5′U**GA**U/3′G**AG**G loop produces the widest dispersion (4.7 ppm; spectrum iii of Figure 4).

Changes in torsion angles, bond angles, and hydrogen bonding can affect the phosphorus chemical shift dispersion, which is approximately 1.5 ppm for the gauche conformations (60° or −60°) of the R−O−P−O(R′) torsion angle in an A-form helix (*21*, *47*). A trans (180° or −180°) conformation gives a dispersion of approximately 2.6 ppm (*21*). The trans−trans conformation of the torsion angles also has

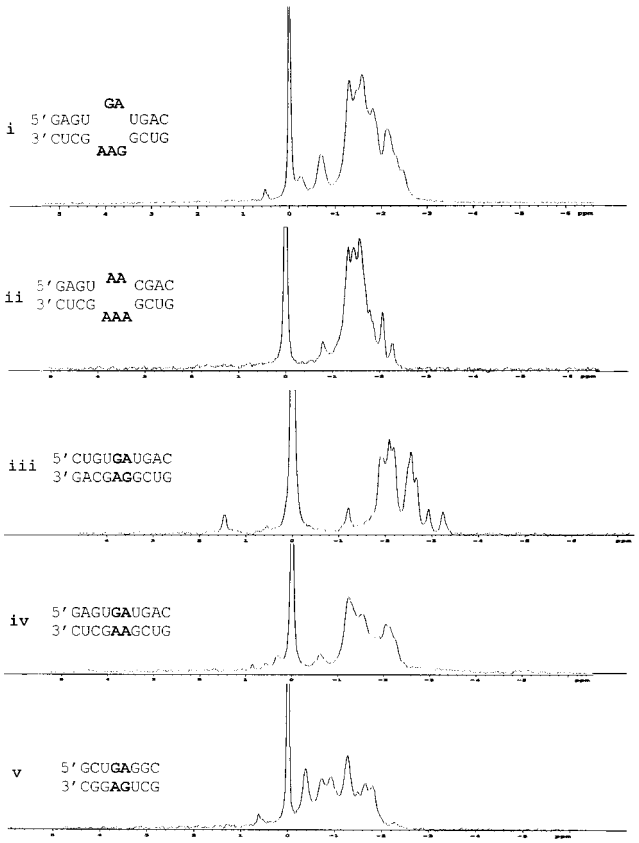

FIGURE 4: One-dimensional ^{1}H-decoupled ^{31}P spectra (−5 to 5 ppm) at 25 °C for i−iv (v at 10 °C) in 80 mM NaCl, 10 mM phosphate, 0.5 mM Na_2EDTA, pH 6.7 buffer and 99.99% D_2O. Spectra are listed by type of loop and then decreasing chemical shift dispersion. The largest peak is the resonance for the phosphate buffer: (i) 5′GAG**UGAU**GAC/3′CUC**GAAGG**CUG, 2 mM; (ii) 5′GAG**UAAC**GAC/3′CUC**GAAAG**CUG, 0.6 mM; (iii) 5′CU-G**UGAU**GAC/3′GAC**GAGG**CUG, 0.2 mM; (iv) 5′GAG**UGAU**-GAC/3′CUC**GAAG**CUG, 0.6 mM; and (v) (5′GC**UGAG**GC)$_2$, 2.0 mM.

a reduced bond angle. A downfield shift in the phosphorus chemical shift implies a smaller than average RO−P−OR′ bond angle, and an upfield shift implies a larger than average bond angle (*21*). Hydrogen bonding may also affect the phosphorus chemical shift by approximately 0.5 ppm (*21*). Chemical shifts downfield from the large phosphate buffer peak are unusual, but have been observed for the sheared GA pair in a GAAA tetraloop hairpin (*47*) and tandem sheared GA pairs with AU closing pairs (*42*). A similar peak is seen in spectra i and iii−v of Figure 4. The phosphorus spectra for the duplexes that model the J4/5 loops in *Pneumocystis carinii* and *Tetrahymena thermophila* group I introns (spectra i and ii, respectively, of Figure 4) exhibit different chemical shift dispersion, suggesting that the two loops have some different backbone conformations.

Table 3 lists in order of decreasing phosphorus chemical shift dispersion several RNA duplexes that contain 2 × 2 and 2 × 3 internal loops with GU closing pairs and/or GA pairs. All duplexes above the bold line have a phosphorus chemical shift dispersion greater than the approximately 1.5 ppm typical for A-form RNA. The loops that contain both a sheared GA pair and either GU or AU pairs have a phosphorus chemical shift dispersion of >1.5 ppm and an unusual downfield phosphorus resonance (Figure 4 and Table 3). The loops that contain GC closing pairs and GA pairs

Table 3: Comparison of Dispersion in ^{31}P NMR Spectra and Loop Free Energies at 37 °C

loop	Δppm[a]	GA mismatch[b]	closing pair	$\Delta G^{\circ}_{loop,37}$ (kcal/mol)[c]
5′CUG**UGAU**GAC 3′GAC**GAGG**CUG	4.7	1 or 2 sheared	2 GU	0.2 (2.3)
5′pGGGC**UGAA**GCCU 3′UCCG**AAGU**CGGGp[d,e]	3.9	2 sheared	2 AU	0.7 (1.7)
5′GAG**U GA U**GAC 3′CUC**GAAGG**CUG	3.0	1 or 2 sheared	2 GU	1.9 (3.3)
5′GAG**UGAU**GAC 3′CUC**GAAG**CUG	2.4	1 sheared	2 GU	2.2 (2.3)
5′GC**UGAG**GC 3′CG**GAGU**CG[f]	2.4	2 sheared	2 GU	0.1 (2.3)
5′GC**GGAU**GC 3′CG**UAGG**CG[f]	1.9	2 imino	2 GU	1.8 (2.3)
5′GAGU **AA C**GAC 3′CUC**GAAAG**CUG	1.5	none	1 GU 1 GC	3.2 (2.8)
5′GC**GAGU**GC 3′CG**UGAG**CG	1.4	2 imino	2 GU	2.6 (2.3)
5′GC**GGAC**GC 3′CG**CAGG**CG[f,g]	1.2	2 imino	2 GC	-2.6 (0.3)
5′CU**GUAGG**CAG 3′GAC**GGAU**GUC	1.1	nr	2 GU	3.4 (2.3)
5′GG**CAGG**CC 3′CC**GGAC**GG[h]	1.1	2 imino	2 GC	-0.5 (0.3)
5′GG**CGAG**CC 3′CC**GAGC**GG[i]	1.0	2 sheared	2 GC	-0.7 (0.3)
5′GG**AAGU**CC 3′CC**UGAAG**G	1.0	2 imino	2 AU	1.7 (1.7)
5′GAG**C GA C**GAC 3′CUC**GAAGG**CUG[j]	0.9	2 sheared	2 GC	0.2 (2.1)

[a] The dispersion is defined as the difference between the most upfield and downfield peaks in the ^{31}P spectrum at the temperature that gave the greatest dispersion but was still at least 10 °C below the T_m of the duplex. Sequences are listed in order of decreasing dispersion. [b] The type of GA pair is either determined from the solved structure or predicted on the basis of a one-dimensional imino proton spectrum with a chemical shift at approximately 10 ppm indicating a sheared GA pair and at 12.5−11.5 ppm indicating an imino GA pair (*17*, *52*, *53*). When no G imino resonance was observed for the GA pair, the notation nr is used. An alternative arrangement of the data sorted by type of closing pair is included in the Supporting Information. [c] See footnote a in Table 1 for details of the calculation of loop free energy. Listed in parentheses is the average value for all measured loops with the same loop size, asymmetry, and closing pairs. The average values for ten 2 × 2 loops with two GU closing pairs (*17*, *18*), fourteen 2 × 2 loops with two AU closing pairs (*17*, *18*, *22*, *36*), and seventy 2 × 2 loops with two CG closing pairs (*17*, *18*, *22*, *34*, *36*, *37*), are 2.3, 1.7, and 0.3 kcal/mol, respectively. The average values for five 2 × 3 loops with two GU closing pairs (this work), five 2 × 3 loops with one GU and one GC closing pair (this work), and eighteen 2 × 3 loops with two GC closing pairs (*20*) are 3.3, 2.8, and 2.1 kcal/mol, respectively. The loop free energy parameters published before 1998 were recalculated with updated nearest neighbor parameters for the interrupted base pairs (*19*, *22*). [d] The solution structure and the phosphorus spectrum are reported in ref *42*. The terminal 5′ phosphate resonates at 0.80 ppm. The G6 5′ phosphate and A6 5′ phosphate resonate at −3.09 and 0.27 ppm, respectively, giving a phosphorus dispersion of 3.4 ppm without including the terminal 5′ phosphate. [e] The loop free energy is that reported in refs *17* and *22*. [f] The duplex free energy and one-dimensional ^{1}H spectrum are reported in ref *17*. The loop free energy was calculated with updated nearest neighbor parameters for the interrupted GU pairs (*22*). [g] The solution structure and the phosphorus spectrum are reported in ref *53*. [h] The solution structure and phosphorus chemical shifts are reported in ref *76*. The duplex free energy is reported in ref *37*. The loop free energy is calculated with updated nearest neighbor parameters for the interrupted CG pairs (*19*). [i] The solution structure and phosphorus spectrum are reported in ref *52*. [j] The one-dimensional imino proton spectrum is reported in ref *20*.

have typical phosphorus chemical shift dispersions and no unusually downfield phosphorus resonance. The 2 × 3 loop with a GU closing pair and no GA pairs also has a typical phosphorus chemical shift dispersion (Figure 4 and Table 3). Of the loops with imino GA pairs, only the 5′G**GA**U/3′U**AG**G loop has a slightly larger than typical phosphorus dispersion, but it has no resonance downfield of the large phosphate buffer peak (see the Supporting Information). Otherwise, loops that contain imino GA pairs with either GC or GU closing pairs have a typical phosphorus chemical shift dispersion. Thus the largest dispersions are associated with internal loops having a sheared GA pair adjacent to a GU or AU pair.

Table 3 also compares the phosphorus chemical shift dispersion and type of GA pairing with the loop free energy. Despite having an unusually downfield phosphorus chemical shift, all the loops above the bold line in Table 3 have a thermodynamic stability more favorable than or equal to the average loop stability for all loops with the same size, asymmetry, and closing pairs. Thus, in these cases, the backbone configuration that gives rise to a deshielded phosphorus resonance is not thermodynamically destabilizing overall. Hydrogen bonding to the backbone or changes in backbone angles that allow for thermodynamically favorable base stacking are possible configurations that would be thermodynamically stable and also generate a deshielded phosphorus resonance. The thermodynamically stable and unique conformation of the 5′UG/GA motif could thus be a good drug target.

DISCUSSION

Internal loops with GU closing pairs are a common motif in RNA secondary structure (*1*, *7*) and can be part of important tertiary structure motifs such as the tetraloop receptor and AA platform (*14*, *15*). The free energy increments for three conserved internal loops with GU closing pairs in group I introns, J4/5, J6/6a, and J6a/6b (footnote i of Table 1), provide information about the stabilities of the loops before they form tertiary contacts. The 2 × 3 internal loop in helix 22 in 16S rRNA (nucleotides 662−666 and 740−743 in *E. coli*), which has a GU closing pair, forms part of a binding site for protein S15 and thus plays a role in the folding of the central domain (*48*). Thus, understanding the energetics of internal loops can facilitate modeling of RNA secondary, tertiary, and quaternary structure and thereby also facilitate selecting sites for targeting RNA with therapeutics.

In Small Asymmetric Internal Loops, the Free Energy Increment for Changing a GC Closing Pair to a GU Closing Pair Is Consistent with a Hydrogen Bonding Model. In eq 2, the AU/GU penalty is 0.2 kcal/mol in addition to a 0.45 kcal/mol penalty for terminating a helix with an AU pair (*18*, *22*). A loop with one or two GU closing pairs thus is predicted to be 0.65 or 1.3 kcal/mol, respectively, less stable than a loop with one or two GC closing pairs. This prediction is within experimental error for 2 × 3 loops and one 1 × 3 loop with one or two GU closing pairs (see the Supporting Information). GC pairs have three hydrogen bonds, and wobble GU pairs have two hydrogen bonds. Half of the free energy contribution from the hydrogen bonds in the closing pairs is counted in the free energy of the internal loop, and half is counted in the nearest neighbor term in the helix adjacent to the loop. Thus, the free energy of a loop with one or two GU pairs, respectively, has contributions from one-half or one fewer hydrogen bond than a loop with two GC pairs. This hydrogen bonding model is consistent with previous measurements of the free energy of a hydrogen bond in RNA (*29*, *49*). Similar hydrogen bond accounting explains

the differences in thermodynamic stabilities for 1 × 3, 2 × 2, and 2 × 3 loops with AU or GC closing pairs and helices terminating with AU, GC, or isoG/isoC pairs (*17*−*20*, *22*, *50*). Asymmetric internal loops may have a flexible structure that accommodates the nonisosteric GU closing pairs without any further thermodynamic penalties.

Free Energy Increments for 2 × 2 Loops with GU Closing Pairs Have a Wide Range of Stabilities That Are Not Explained by a Hydrogen Bonding Model. The measured free energy increments for 2 × 2 loops with GU closing pairs range from 0.1 to 3.6 kcal/mol (Tables 1 and 2). The differences in stabilities between loops with GC and GU closing pairs range from 0.8 to 4.4 kcal/mol (Table 2 and the Supporting Information), which is much different than the value of 1.3 kcal/mol predicted by a hydrogen bonding model. When the closing pairs change from AU to GU and no change in the number of hydrogen bonds is expected, the changes in loop free energy range from 0.6 to 2.5 kcal/mol (Table 2). For example, because the closing pairs have the same number of hydrogen bonds, similar stabilities were previously predicted for 5′A**AG**U/3′U**GA**A, 5′A**GA**U/3′U**AG**A, and 5′U**AG**G/3′G**GA**U loops. The 5′A**AG**U/3′U**GA**A loop, however, is 1.4 kcal/mol less stable than the 5′A**GA**U/3′U**AG**A loop and 1.7 kcal/mol more stable than the 5′U**AG**G/3′G**GA**U loop (Tables 1 and 2). Stacking interactions between the closing pair and an adjacent GA pair, either sheared or imino, have a particularly dramatic effect on the thermodynamic stabilities of 2 × 2 loops (Table 2). Although a simple hydrogen bonding model does not explain the thermodynamic stabilities of symmetric 2 × 2 loops with GA pairs, the more complete list of experimental loop free energies in Table 2 should improve predictions of loop stabilities and RNA secondary structures.

Stacking Interactions of the Closing GU Pair with the Adjacent Helix Can Stabilize 2 × 2 Loops. Studies of secondary structures of large and small subunit ribosomal RNA reveal a preference for a loop to occur on the 5′ side of a G in a GU pair at a loop−helix junction; the 3′ side of the G is considered the "stacking" side (*1*, *51*). This pattern is consistent with the loop free energies for 2 × 2 loops containing GU closing pairs and only A nucleotides. For example, the 5′U**AA**C/3′G**AA**G loop, in which the 3′ side of the G in the GU pair stacks on the stem helix, is 0.8 kcal/mol more stable than the 5′C**AA**U/3′G**AA**G loop, in which the 3′ side of the G in the GU pair stacks on the A nucleotides in the loop (Table 1). The 5′U**AA**G/3′G**AA**U loop (Table 1), in which the 3′ side of both of the G nucleotides in the GU pairs stacks on the stem helices, is 1.5 kcal/mol more stable than the 5′U**AA**U/3′G**AA**G loop, in which the 3′ side of only one G nucleotide in a GU pair stacks on the stem helix. Using the known nonsymmetric loop values for loops with GU closing pairs and all A nucleotides, the 5′G**AA**U/3′U**AA**G loop is predicted to have a stability increment of 4.1 kcal/mol. Thus, the 5′G**AA**U/3′U**AA**G loop, in which the stacking sides of both G's in the GU pairs stack on the loop, is predicted to be approximately 2 kcal/mol less stable than the 5′U**AA**G/3′G**AA**U loop, in which the stacking sides of the G's stack on the stems (Table 2). Evidently, the stacking interactions of GU closing pairs on the loop and on the helix stems affect the thermodynamic stability of the loop.

Changing the Closing Pairs from GC to GU Apparently Does Not Change the Conformation of the GA Pair. Sheared and imino GA pairs can be distinguished on the basis of chemical shifts in a one-dimensional imino proton NMR spectrum. The spectra of (5′GC**UGAG**GC)$_2$ (*17*) and (5′GG**CGAG**CC)$_2$ (*52*) duplexes both show a resonance near 10 ppm, indicating a sheared GA pair. The spectra of the (5′GC**GGAU**GC)$_2$ (*17*), (5′GC**GGAC**GC)$_2$ (*17*, *53*), (5′GC**GAGU**GC)$_2$ (spectrum iv of Figure 2), and (5′GC**GAGC**GC)$_2$ (*17*) duplexes show a G imino resonance near 11.5 ppm, indicating an imino hydrogen-bonded GA pair. Evidently, changing the closing pairs of the loop from GC to GU does not change the conformation of the adjacent GA pair in these cases. By analogy, the GA pairs of the (5′CUG**UAGG**CAG)$_2$ duplex are expected to be imino because the 5′C**AG**G/3′G**GA**C loop has tandem imino GA pairs (*53*). The imino proton spectrum of the 5′U**AG**G/3′G**GA**U loop, however, does not show a resonance that indicates the conformation of the GA pairs (Figure 3). The 5′U**AG**G/3′G**GA**U loop is very unstable thermodynamically (Table 2), suggesting weak stacking and hydrogen bonding in a flexible loop that allows exchange between the G6 imino proton and water.

Sheared GA Pairs Are More Stable Than Imino GA Pairs in Symmetric 2 × 2 Loops with GU Closing Pairs. When a symmetric 2 × 2 loop has GU closing pairs, the loop with sheared GA pairs (0.1 kcal/mol) is more stable than the loops with imino GA pairs (1.8 and 2.6 kcal/mol) (Table 3). Thus, changing the closing pairs from $^{5'G}_{3'U}$ to $^{5'U}_{3'G}$ for a loop with the sequence $^{5'GA}_{3'AG}$ causes the GA pairs to switch from an imino to a sheared conformation, and the loop stability to become more favorable by 1.7 kcal/mol (*17*, *52*, *53*). In contrast, changing the closing pairs from $^{5'G}_{3'C}$ to $^{5'C}_{3'G}$ for a loop with the sequence $^{5'GA}_{3'AG}$ also causes the GA pairs to switch from an imino to a sheared conformation, but the loop stability to become less favorable by 1.9 kcal/mol (*17*, *52*, *53*).

Sheared GA Pairs Adjacent to GU Closing Pairs Have an Unusual Phosphorus Resonance. The 2 × 3 and 2 × 2 loops with a sheared GA pair adjacent to a GU closing pair have a phosphorus chemical shift dispersion wider than 1.5 ppm and an unusually downfield resonance (Table 3 and Figure 4). These loops are also thermodynamically more stable than the average for loops with the same loop size, symmetry, and closing pair (Table 3). The 5′UG/3′GA motif may have interactions that stabilize its apparent non-A-form conformation. Sheared GA pairs present amino and imino groups in positions where they can form hydrogen bonds to the ribose−phosphate backbone (*52*, *54*−*56*). The type of closing pair also affects the type of sugar pucker in the G of the GA pair (*57*). The 5′UG/3′GA motif is very common phylogenetically and places the "stacking side" of the GU pair on the stem helix (*1*). Favorable electrostatic overlap of the bases may stabilize a particular orientation of GA and GU pairs more than other configurations (*43*, *45*, *53*). Evidently, the configuration of adjacent GU and sheared GA pairs generates a downfield-shifted phosphorus resonance but is not overall thermodynamically destabilizing.

Non-Nearest Neighbor Behavior May Affect the Stability of 2 × 2 Loops with GU Closing Pairs. The changes in phosphorus backbone angles may also cause some non-nearest neighbor effects. The thermodynamic stabilities of

Table 4: Comparison of Free Energies for the J4/5 Loop and for a Tertiary Interaction in Group I Introns

Group I Introns

Organism	J4/5 Loop Model	$\Delta G^{\circ}_{loop,37}$ (kcal/mol)[a]	Docking $\Delta\Delta G^{\circ}_{GC->GU}$ (kcal/mol)[b]
T. thermophila	5'GAG**U AA C**GAC 3'CUC**GAAAG**CUG	3.2±0.5	-2.6[c]
C. albicans	5'GAG**UAAU**GAC 3'CUC**GAAG**CUG	3.6±1.0	-2.2[d]
P. carinii	5'GAG**U GA U**GAC 3'CUC**GAAGG**CUG	1.9±1.0	-4.3[e]

[a] See footnote a of Table 2. [b] The free energy difference for the loss of tertiary interactions when the conserved GU pair in the P1 helix is changed to a GC pair is calculated according to the methods in refs *8*, *26*, and *27*. [c] From ref *8*. [d] From ref *27*. [e] From ref *26*.

tandem GU pairs display non-nearest neighbor effects in some contexts (*22*, *32*). Structural studies of GU pairs do not agree about the extent to which GU pairs, which are not isosteric with Watson−Crick pairs, perturb the helix beyond the adjacent pairs (*44*, *58*−*68*). Thus, the nearest neighbor $\Delta G^{\circ}_{\text{interrupted base pair}}$ term in eq 1a may be too simple an approximation (*17*). Other possible conformations, for example, GU closing pairs with one hydrogen bond (*44*), GU closing pairs with longer than normal hydrogen bonds as observed for AU closing pairs in a tandem GA loop (*42*), or GU closing pairs that are part of a base triple as observed in two loops in 23S rRNA (nucleotides 2090−2094 and 2652−2654, and 2492−2496 and 2526−2531, in *Haloarcula marismortui*) (*69*), may affect the loop stability in some contexts.

Thermodynamic Stabilities of Internal Loops That Model J4/5 Loops in Different Group I Introns Correlate with Assessments of Tertiary Interactions at This Site. Understanding the thermodynamic stabilities of internal loops can contribute to understanding and predicting the stabilities of tertiary structures that involve internal loops. For example, the J4/5 loop in the *T. thermophila* group I intron is the docking site for the P1 helix (*8*). The GU pair in the P1 helix makes specific contacts to the tandem sheared AA pairs in the J4/5 loop (*8*, *9*, *70*). Experiments in which the GU pair in the P1 helix is mutated to a GC pair provide a measure of the contribution of the G exocyclic amine group to the free energy of the docking interaction (*8*, *71*, *72*). The structure and stability of the unbound J4/5 loop will affect the net free energy of such docking interactions.

As shown in Tables 1 and 4, the internal loops that model the J4/5 loops in the *T. thermophila* and *Candida albicans* introns, 5′U**AA**C/3′G**AAA**G and 5′U**AA**U/3′G**AA**G, have similar thermodynamic stabilities, 3.2 and 3.6 kcal/mol, respectively. In contrast, the loop that models the J4/5 loop in *P. carinii*, 5′U**GA**U/3′G**AAG**G, is more stable by 1.5 kcal/mol on average. When the GU pair in the P1 helix is changed to a GC pair, the docking interactions in the *T. thermophila*, *C. albicans*, and *P. carinii* ribozymes are less favorable by 2.6, 2.2, and 4.4 kcal/mol, respectively (*8*, *26*, *27*). The difference in loop stability for the *P. carinii* J4/5 loop compared to the other two loops is similar to the difference in the free energy contribution to docking of the exocyclic amine of the GU pair. These data suggest that the J4/5 loop in *P. carinii* is more preorganized for the docking interaction, and therefore, the tertiary interactions involving the J4/5 loop make a more favorable net contribution to docking. Thus, elimination of the interactions involving the exocyclic amine of the G in the GU pair is more destabilizing. These correlations suggest that understanding the sequence dependence of internal loop stabilities and structures will provide a foundation for predicting both sites and stabilities of tertiary interactions in RNA.

ACKNOWLEDGMENT

We thank Mathew D. Disney for helpful discussions.

SUPPORTING INFORMATION AVAILABLE

Detailed methods, sample melting curves and T_M^{-1} versus $\ln(C_T/n)$ plots for duplexes with internal loops that model the J4/5 loops in *T. thermophila*, *P. carinii*, and *C. albicans*, a table of thermodynamic parameters for self-complementary single-strand oligomers, a table comparing the thermodynamic parameters derived from T_M^{-1} versus $\ln(C_T/n)$ plots and from curve fits, a table of internal loop enthalpies and entropies, a table comparing the free energies for internal loops with GU and GC closing pairs, an alternate format for Table 3, imino proton spectra for duplexes with 2 × 3 and 3 × 3 loops, and additional one-dimensional phosphorus spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Gautheret, D., Konings, D., and Gutell, R. R. (1995) *RNA 1*, 807−814.
2. Wimberly, B. T., Broderson, D. E., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan (2000) *Nature 407*, 327−339.
3. Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., and Yonath, A. (2000) *Cell 102*, 615−623.
4. Cate, J. H., Yusupov, M. M., Yusupova, G. Zh., Earnest, T. N., and Noller, H. F. (1999) *Science 285*, 2095−2104.
5. Gutell, R. R., Subashchandran, S., Schnare, M., Du, Y., Lin, N., Madabusi, L., Muller, K., Pande, N., Yu, N., Shang, Z., Date, S., Konings, D., Schweiker, U., Weiser, B., and Cannone, J. Comparative Sequence Analysis and the Prediction of RNA Structure and the Web. http://www.rna.icmb.utexas.edu (accessed Oct 2000).
6. Carter, A., Clemons, W., Broderson, D., Morgan-Warren, R., Hartsch, T., Wimberly, B., and Ramakrishnan, V. (2001) *Science 291*, 498−501.
7. Damberger, S. H., and Gutell, R. R. (1994) *Nucleic Acids Res. 22*, 3508−3510.
8. Strobel, S. A., and Cech, T. R. (1995) *Science 267*, 675−679.
9. Strobel, S. A., and Ortoleva-Donnely, L. (1999) *Chem. Biol. 5*, 153−165.
10. Bevilacqua, P. C., Kierzek, R., Johnson, K. A., and Turner, D. H. (1992) *Science 258*, 1355−1358.
11. Herschlag, D. (1992) *Biochemistry 31*, 1386−1399.
12. Pyle, A. M., Murphy, F. L., and Cech, T. R. (1992) *Nature 358*, 123−128.
13. Michel, F., and Westhof, E. (1994) *Nat. Struct. Biol. 1*, 5−7.
14. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1696−1699.
15. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1678−1685.

16. Basu, S., Rambo, R. P., Strauss-Soukup, J., Cate, J. H., Ferre-D'Amare, A. R., Strobel, S. A., and Doudna, J. A. (1998) *Nat. Struct. Biol. 5*, 986−992.

17. Walter, A. E., Wu, M., and Turner, D. H. (1994) *Biochemistry 33*, 11349−11354.

18. Wu, M., McDowell, J. A., and Turner, D. H. (1995) *Biochemistry 34*, 3204−3211.

19. Xia, T., SantaLucia, J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) *Biochemistry 37*, 14719−14735.
20. Schroeder, S. J., and Turner, D. H. (2000) *Biochemistry 39*, 9257−9274.
21. Gorenstein, D. G. (1984) *Phosphorus-31 NMR: Principles and Applications*, Academic Press, Orlando, FL.
22. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) *J. Mol. Biol. 288*, 911−940.
23. Rivas, E., and Eddy, S. E. (1999) *J. Mol. Biol. 285*, 2053−2068.
24. Gultyaev, A., van Batenburg, F., and Pleij, C. (1995) *J. Mol. Biol. 250*, 37−51.
25. Luck, R., Steger, G., and Reisner, D. (1996) *J. Mol. Biol. 258*, 813−826.
26. Disney, M. D., Gryaznov, S. M., and Turner, D. H. (2000) *Biochemistry 39*, 14269−14278.
27. Disney, M., Haidaris, C., and Turner, D. H. (2001) *Biochemistry 40*, 6507−6519.
28. Freier, S. M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M. H., and Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.
29. Freier, S. M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M. H., and Turner, D. H. (1986) *Biochemistry 25*, 3214−3219.
30. Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.
31. Gralla, J., and Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301−319.
32. He, L., Kierzek, R., SantaLucia, J., Jr., Walter, A. E., and Turner, D. H. (1991) *Biochemistry 30*, 11124−11132.
33. Schroeder, S. J., Burkard, M. E., and Turner, D. H. (2001) *Biopolymers 52*, 157−167.
34. Peritz, A. E., Kierzek, R., Sugimoto, N., and Turner, D. H. (1991) *Biochemistry 30*, 6428−6436.
35. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1991) *Biochemistry 30*, 8242−8251.
36. Burkard, M. E., Xia, T., and Turner, D. H. (2001) *Biochemistry 40*, 2478−2483.
37. Xia, T., McDowell, J. A., and Turner, D. H. (1997) *Biochemistry 36*, 12486−12497.
38. Xia, T., Mathews, D., and Turner, D. H. (1999) in *Prebiotic Chemistry, Molecular Fossils, Nucleosides, and RNA* (Soll, D., Nishimura, S., and Moore, P., Eds.) pp 21−48, Elsevier Press, Oxford, England.
39. Wijmenga, S. S., Heus, H. A., Werten, B., van de Marel, G. A., van Boom, J. H., and Hilbers, C. W. (1994) *J. Magn. Reson., Ser. B 103*, 134−141.
40. Johnston, P. D., and Redfield, A. G. (1978) *Nucleic Acids Res. 5*, 3913−3927.
41. Hilbers, C. W., and Walters, J. A. L. I. (1990) in *Landolt-Bornstein Numerical Data and Functional Relationships in Science and Technology* (Madelung, O., Ed.) pp 191−200, Springer-Verlag, Heidelberg, Germany.
42. Heus, H., Wijmenga, S., Hoppe, H., and Hilbers, C. (1997) *J. Mol. Biol. 271*, 147−158.
43. McDowell, J. A., and Turner, D. H. (1996) *Biochemistry 35*, 14077−14089.
44. Chen, X., McDowell, J. A., Kierzek, R., Krugh, T. R., and Turner, D. H. (2000) *Biochemistry 39*, 8970−8982.
45. McDowell, J. A., He, L., Chen, X., and Turner, D. H. (1997) *Biochemistry 36*, 8030−8038.
46. Kettani, A., Gorin, A., Majumdar, A., Hermann, T., Skripkin, E., Zhao, H., Jones, R., and Patel, D. J. (2000) *J. Mol. Biol. 297*, 627−644.
47. Legault, P., and Pardi, A. (1994) *J. Magn. Reson., Ser. B 103*, 82−86.
48. Agalarov, S., Prasad, G., Funke, P., Stout, C., and Williamson, J. (2000) *Science 288*, 107−112.
49. Turner, D. H., Sugimoto, N., Kierzek, R., and Dreiker, S. D. (1987) *J. Am. Chem. Soc. 109*, 3783−7385.
50. Chen, X., Kierzek, R., and Turner, D. H. (2001) *J. Am. Chem. Soc. 123*, 1267−1274.
51. Mizuno, H., and Sundaralingam, M. (1978) *Nucleic Acids Res. 5*, 4451−4461.
52. SantaLucia, J., Jr., and Turner, D. H. (1993) *Biochemistry 32*, 12612−12623.
53. Wu, M., and Turner, D. H. (1996) *Biochemistry 35*, 9677−9689.
54. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1992) *Science 256*, 217−219.
55. Jucker, F., Heus, H., Yip, P., Moors, E., and Pardi, A. (1996) *J. Mol. Biol. 264*, 968−980.
56. Heus, H., and Pardi, A. (1991) *Science 253*, 191−194.
57. Michiels, P., Schouten, C., Hilbers, C., and Heus, H. (2000) *RNA 6*, 1821−1832.
58. Masquida, B., and Westhof, E. (2000) *RNA 6*, 9−15.
59. Limmer, S., Reif, B., Ott, G., Arnold, L., and Sprinzl, M. (1996) *FEBS Lett. 385*, 15−20.
60. Ramos, A., and Varani, G. (1997) *Nucleic Acids Res. 25*, 2083−2090.
61. Betzel, C., Lorenz, S., Furste, J. P., Bald, R., Zhang, M., Schnieder, T. R., Wilson, K. S., and Erdmann, V. A. (1994) *FEBS Lett. 351*, 159−164.
62. Masquida, B., Sauter, C., and Westhof, E. (1999) *RNA 5*, 1384−1395.
63. White, S. A., Nilges, M., Huang, A., Brunger, A. T., and Moore, P. B. (1992) *Biochemistry 31*, 1610−1621.
64. Shi, K., Wahl, M., and Sundaralingam, M. (1999) *Nucleic Acids Res. 27*, 2196−2201.
65. Mueller, U., Schubel, H., Sprinzl, M., and Heinemann, U. (1999) *RNA 5*, 670−677.
66. Trikha, J., Filman, D. J., and Hogle, J. M. (1999) *Nucleic Acids Res. 27*, 1728−1739.
67. Biswas, R., Wahl, M. C., Ban, C., and Sundaralingam, M. (1997) *J. Mol. Biol. 267*, 1149−1156.
68. Biswas, R., and Sundaralingam, M. (1997) *J. Mol. Biol. 270*, 1149−1156.
69. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) *Science 289*, 905−919.
70. Strobel, S. A., and Cech, T. R. (1996) *Biochemistry 35*, 1201−1211.
71. Pyle, A. M., Moran, S., Strobel, S. A., Chapman, T., Turner, D. H., and Cech, T. R. (1994) *Biochemistry 33*, 13856−13863.
72. Knitt, D. S., Narlikar, G. J., and Herschlag, D. (1994) *Biochemistry 33*, 13864−13879.
73. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843−853.
74. Testa, S. M., Haidaris, C. G., Gigliotti, F., and Turner, D. H. (1997) *Biochemistry 36*, 15303−15314.
75. Cech, T., Tanner, K., Tinoco, I., Jr., Weir, B., Zuker, M., and Perlman, P. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3903−3907.
76. Wu, M., SantaLucia, J., Jr., and Turner, D. H. (1997) *Biochemistry 36*, 4449−4460.
77. Burkard, M. E., Turner, D. H., and Tinoco, I., Jr. (1999) in *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) pp 675−680, Cold Spring Harbor Laboratory Press, Plainview, NY.
78. Leontis, N. B., and Westhof, E. (1998) *Q. Rev. Biophys. 31*, 399−455.

BI010489O