Factors Affecting the Thermodynamic Stability of Small Asymmetric Internal Loops in RNA †

ARTICLE *in* BIOCHEMISTRY · SEPTEMBER 2000

Impact Factor: 3.02 · DOI: 10.1021/bi000229r · Source: PubMed

Factors Affecting the Thermodynamic Stability of Small Asymmetric Internal Loops in RNA†

Susan J. Schroeder and Douglas H. Turner*

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

Received February 2, 2000; Revised Manuscript Received May 10, 2000

ABSTRACT: Optical melting experiments were used to determine the thermodynamic parameters for oligoribonucleotides containing small asymmetric internal loops. The results show a broad range of thermodynamic stabilities, which depend on loop size, asymmetry, sequence, closing base pairs, and length of helix stems. Imino proton NMR experiments provide evidence for possible hydrogen bonding in GA and UU mismatches in some asymmetric loops. The stabilizing effects of GA, GG, and UU mismatches on the thermodynamic stability of internal loops vary depending on the size and asymmetry of the loop. The dependence of loop stability on Watson−Crick closing base pairs may be explained by an account of hydrogen bonds. Models are presented for approximating the free energy increments of 2 × 3 and 1 × 3 internal loops.

RNA secondary structure predictions based on free energy minimization use experimental and extrapolated thermodynamic parameters for RNA motifs to generate a set of possible secondary structures for a given RNA sequence (*1*). Approximately half of the nucleotides in known RNA structures form double helices with canonical Watson−Crick or GU base pairs; the other half of the nucleotides form noncanonical regions (*1, 2*). Asymmetric internal loops are a type of noncanonical region defined by an unequal number of nucleotides on opposing strands between two helices. Asymmetric internal loops are common in known RNA secondary structures and often have a role in RNA function. The notation $m \times n$ loop[1] refers to the number of nucleotides on each side of the loop; for example, a 2 × 3 loop contains two nucleotides opposite three nucleotides.

Asymmetric internal loops play an important role in the structure and function of RNA; thus, understanding the thermodynamic stability of these motifs is important for understanding the structure and function of RNA. For example, in the *Tetrahymena thermophila* group I intron, the asymmetric J5/5a loop acts as a hinge that forms a bend in the helix (*3, 4*), and the asymmetric J6a/6b loop acts as a tetraloop receptor (*4, 5*). Identifying these types of tertiary contacts that determine global folding requires a good secondary structure prediction of internal loops. Understanding how RNA functions also requires identification of asymmetric internal loops. For example, the asymmetric J4/5 loop in the group I intron is part of the docking site for the P1 helix (*6*) and may also help to stabilize the transition state of the splicing reaction (*7*). Asymmetric internal loops in other RNAs bind proteins (*8*−*11*). A good secondary structure prediction of internal loops can also help to identify potential target sites for therapeutics. For example, the aminoacyl-tRNA site in 16S rRNA contains an asymmetric internal loop that binds aminoglycoside antibiotics, which inhibit translocation (*12*−*16*). Thus, the thermodynamic parameters for asymmetric internal loops can help to improve RNA secondary structure prediction, tertiary structure prediction, and rational drug design.

Initial studies of symmetric internal loops focused on loops containing all C nucleotides and showed that the thermodynamic stability depends on the number of nucleotides in the loop (*17*). Later Papanicolau et al. (*18*) proposed a penalty for the degree of asymmetry in internal loops, and Peritz et al. (*19*) determined an asymmetry penalty based on optical melting experiments with internal loops containing all A nucleotides. Further experiments on internal loops containing all C nucleotides showed a correlation between loop asymmetry and a less favorable stacking enthalpy and increased accessibility of the major groove to chemical modification (*20*). Thermodynamic studies of 2 × 2 internal loops have shown a broad range of stabilities that depend on the sequence and closing base pairs of the loop (*21*−*25*). Thermodynamic studies of 1 × 2 internal loops also show that loop stability depends on the sequence and potential mismatches in the loop (*26*). NMR and X-ray crystal structures containing asymmetric internal loops reveal a variety of stacking and hydrogen-bonding interactions that stabilize novel structures of the loops (*4, 5, 9*−*11, 27*−*43*).

This paper presents thermodynamic parameters and structural insights for 1 × 3 and 2 × 3 internal loops as deduced from optical melting experiments and imino proton NMR spectra. Several internal loops with diverse sequences and biological functions were selected as a starting point for studying small asymmetric internal loops (Figure 1); other

† This work was supported by NIH Grant GM22939.

* To whom correspondence should be addressed: Tel (716) 275-3207; fax (716) 473-6889; e-mail: Turner@chem.rochester.edu.

[1] Abbreviations: EDTA, ethylenediaminetetraacetic acid; eu, entropic units, calories per kelvin per mole; HPLC, high-pressure liquid chromatography; MES, 2-(*N*-morpholino)ethanesulfonic acid; $m \times n$ loop, an internal loop containing m nucleotides opposite n nucleotides where m and n are integers; NCR, noncoding region; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; TLC, thin-layer chromatography; T_M, melting temperature in kelvins; T_m, melting temperature in degrees Celsius; UTR, untranslated region.

10.1021/bi000229r CCC: $19.00
Published on Web 07/14/2000

Segment of Natural Sequence	Melted Duplex	Biological Interest
5'UUU ^{GC} GCU AAA $_{AAA}$ CGG	5'CCUCU ^{GC} GGUGA GGAGA$_{AAA}$CCGCU	Potato spindle tuber virus conserved core (*71*)
5'GAC ^{GA} GCA CUG $_{GAA}$ CGU	5'CGAC ^{GA} GCAG GCUG$_{GAA}$CGUC	p57 binding site in loop E of 5' NCR of encephalomyocarditis virus (*72*)
5'GCCU ^{GC} UUU CGGA$_{ACA}$AAA	5'ACCU ^{GC} UUGC UGGA$_{ACA}$AACG	Conserved 5' region of influenza virus, polymerase binding site (*73*)
5'AGG ^{UU} UU UCC $_{UUU}$ AG	5'UGAC ^{UU} CUCA ACUG$_{UUU}$GAGU	Conserved loop in R2 retrotransposable element (*74*)
5'CCCCU ^{GG} ACGA GGGGG $_{AGA}$ UGCU	5'CUGU ^{GG} ACGA GACG$_{AGA}$ UGCU	Protected by S6 and S18 protein binding to 16S rRNA (*63*)
5'GGCUG ^{GA} CUC CCGAC $_{AAA}$ GAG	5'GAGC ^{GA} CGAC CUCG $_{AAA}$ GCUG	rev protein binding aptamer (*30*)
5'CUGG ^{CC} G GUCC $_{AAU}$ C	5'UCAG ^{CC} GUGA AGUC $_{AAU}$ CACU	most commonly occurring 2X3 loop, conserved in small subunit rRNA (*62*)
5'GGGU ^{C} UGGG CCCA $_{UUU}$ ACCC	5'AAGGU ^{C} UGGAA UUCCA $_{UUU}$ ACCUU	Turnip crinkle virus, recombination mechanism (*75*)
5'CGC ^{G} CUC GUG$_{AAA}$GAG	5'CCAC ^{G} CUCC GGUG$_{AAA}$GAGG	3'UTR Coronavirus mouse hepatitis (*76*)
5'CUU ^{A} UGC GAA $_{AAG}$ACG	5'UCCU ^{A} UGCA AGGA $_{AAG}$ ACGU	Potato spindle tuber virus pathogenicity domain (*77*)
5'UGG ^{C} CGGA ACC $_{AAU}$ GCCU	5'UCAG ^{C} GUGA AGUC$_{AAU}$ CACU	Domain V 5'NCR poliovirus (*78*)
5'UC ^{A} G AG $_{GGA}$C	5'UCCG ^{A} CGCA AGGC $_{GGA}$ GCGU	Ni^{+} binding aptamer (*79*)
5'CGCG ^{A} CGG GCGU $_{AAC}$ GCC	5'UCAG ^{A} GUGA AGUC $_{AAC}$ CACU	most commonly occurring 1X3 loop, conserved in small subunit rRNA (*62*), binds protein S16 (*63*)

FIGURE 1: Sequences for the duplexes studied were chosen on the basis of sequences of known internal loops and on flanking helices that gave two-state helix-to-coil transitions without indications of aggregation or formation of other structures.

loops were designed to test specific hypotheses. The thermodynamic stability of asymmetric internal loops depends on the loop size, degree of asymmetry, loop sequence, closing base pairs, and length of stem helix. The stabilizing effect of potential GA, GG, and UU mismatches varies depending on the type of loop. Models are presented for approximating the thermodynamic stabilities of 1×2, 1×3, and 2×3 loops.

MATERIALS AND METHODS

Oligoribonucleotide Synthesis. Oligoribonucleotides were synthesized on an ABI 392 DNA/RNA synthesizer using phosphoramidite chemistry (*44*, *45*). Incubation in 3:1 (v/v) ammonia/ethanol at 55 °C overnight removed the base protecting groups and the CPG support; incubation in triethylaminehydrogen fluoride at 55 °C for 48 h removed the silyl protecting group on the 2′-hydroxyls. Oligomers were desalted on a Sep-Pak C-18 cartridge and purified on a Baker 500Si TLC plate with either 60:35:5 or 55:35:10 1-propanol/ammonia/water solvents. Purity was analyzed by HPLC; all oligomers were greater than 90% pure.

Optical Melting Experiments and Thermodynamics. Concentrations of single-strand oligomers were determined from the absorbance at 280 nm at 80 °C (*46*). All oligomers were melted as single strands and as duplexes in a 1:1 concentration ratio. Standard melt buffer was 1 M NaCl, 10 or 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7. To measure the effects of pH and magnesium ions, melting experiments were also done in 1 M NaCl, 50 mM MES, and 0.5 mM Na_2EDTA buffer, pH 5.5, or 10 mM $MgCl_2$, 0.15 M KCl, 10 mM sodium cacodylate, and 0.5 mM Na_2EDTA buffer, pH 7. Melting curves were measured at 280 nm on a Gilford 250 spectrophotometer with a heating rate of 1 °C/min controlled by a Gilford 2527 thermoprogrammer.

Melting curves were fit to the two-state model with sloping baselines (*47*). Thermodynamic parameters were also obtained by fitting plots of inverse melting temperature, T_M^{-1}, versus ln (C_T/4) to the following equation (*48*):

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

where R is the gas constant and C_T is total strand concentration. Data from the melting experiments were analyzed with the program Meltwin 3.0 (*49*). When one or both single-strand oligomers could form a stable self-complementary duplex, the competition between the self-complementary and the non-self-complementary duplexes was analyzed graphically and/or investigated by imino proton NMR. Graphical analysis included plotting the distribution of single-strand oligomer, self-complementary duplex, and non-self-complementary duplex versus temperature. The computer program uses ΔH°, ΔS°, and concentration from the single-strand melt data and the mixed duplex data and solves the necessary quartic equation using an iterative approach (D. H. Mathews, personal communication).

The free energies of formation of 2×3 internal loops with CG closing base pairs were used to derive thermodynamic parameters for the additional stability conferred by GA, GG, and UU mismatches in 2×3 loops. The linear regression analysis was computed with JMP software, a program developed by SAS Institute Inc.

NMR. Imino proton spectra were recorded with a Varian INOVA spectrometer at 500 MHz and Varian VNMR software on a Sun 4/260 computer. Standard NMR buffer was 80 mM NaCl, 10 mM sodium phosphates, and 0.5 mM Na_2EDTA, pH 6.7, in 90% H_2O and 10% D_2O. Oligonucleotide strand concentrations ranged from 0.2 to 1 mM. A binomial 1:3:3:1 pulse sequence was used to suppress the water signal (*50*). Spectra were collected with a sweep width of 12 kHz and a frequency offset set to maximize the signal-to-noise ratio at about 13 ppm with the first nodes at 21 and 5 ppm. The data were multiplied by a 2−5 Hz exponential line-broadening function. All variable temperature experiments began at 0.5−1 °C and continued in increments of 5 or 10 °C until all resonances disappeared. For self-complementary oligomers with a melting temperature greater than 20 °C, variable temperature spectra were recorded before addition of the non-self-complementary oligomer. One-dimensional NOE spectra were collected with an irradiation of 2−3 s at low decoupler power; experiments with irradiation on- and off-resonance were collected in blocks of 16 and interleaved to correct for instrumental drift.

RESULTS

Thermodynamic Parameters. Table 1 lists the thermodynamic parameters determined for the formation of duplexes with and without internal loops; duplexes are grouped by type of loop and listed in order of decreasing duplex stability. The thermodynamic parameters from T_M^{-1} vs ln (C_T/4) plots and from curve fits of the data agree within 15%, which is consistent with the two-state model assumption (*51*−*53*). Supporting Information contains graphs of typical melting curves and sample T_M^{-1} vs ln (C_T/4) plots. Supporting Information also contains a table of the thermodynamic parameters for single-strand oligomers that show cooperative transitions in the absorbance versus temperature curves and

Table 1: Thermodynamic Parameters of Duplex Formation[a]

duplex	T_M^{-1} vs Ln (C_T/4) PARAMETERS $-\Delta G^o_{37}$ (kcal/mol)	$-\Delta H^o$ (kcal/mol)	$-\Delta S^o$ (eu)	T_m (°C)	CURVE FIT PARAMETERS $-\Delta G^o_{37}$ (kcal/mol)	$-\Delta H^o$ (kcal/mol)	$-\Delta S^o$ (eu)	T_m (°C)
2X3 Loops								
5'CCUCU **GC** GGUGA[b,c] 3'GGAGA**AAA**CCGCU	11.58±0.16	88.50±3.07	248.03±9.39	55.7	11.86±0.58	93.50±10.29	263.22±31.28	55.8
5'GAGC **GA** CGAC[e] 3'CUCG**AAG**GCUG	10.51±0.10	86.70±2.42	245.67±7.50	51.9	10.17±0.07	78.24±1.90	219.47±6.10	52.1
5'CCAC **GG** CUCC 3'GGUG**AGA**GAGG	9.84±0.06	81.92±1.78	232.41±5.545	50.1	9.75±0.11	78.49±3.09	221.63±9.67	50.3
5'GAGC **AA** CGAC 3'CUCG**AAG**GCUG	9.23±0.09	74.63±2.68	210.86±8.38	48.7	9.35±0.14	77.62±3.03	220.12±9.35	48.7
5'CCAC **GG** CUCC[f] 3'GGUG**AAA**GAGG	9.22±0.06	78.21±2.12	222.43±6.65	48.1	9.16±0.09	75.80±1.19	214.86±3.61	48.1
5'GAGC **AG** CGAC[f] 3'CUCG**GAA**GCUG	8.80±0.08	79.23±3.24	227.07±10.18	46.1	8.68±0.08	72.54±3.84	205.93±12.31	46.5
5'GAGC **GA** CGAC[f] 3'CUCG**AAA**GCUG	8.77±0.07	75.07±2.59	213.77±8.12	46.5	8.68±0.11	72.06±2.29	204.35±7.19	46.5
5'GAGC **AA** CGAC[f] 3'CUCG**GAA**GCUG	8.61±0.06	78.04±2.63	223.86±8.30	45.5	8.59±0.09	75.96±2.95	217.23±9.32	45.6
5'GAGC **AG** CGAC 3'CUCG**AAG**GCUG	8.50±0.10	77.40±4.06	222.16±12.83	45.1	8.40±0.13	71.44±3.79	203.23±12.10	45.4
5'CGAC **GA** GCAG[b,e] 3'GCUG**GAA**CGUC	8.43±0.05	74.66±2.13	213.52±6.73	45.1	8.41±0.09	72.04±3.53	205.14±11.18	45.3
5'GAGC **AG** CGAC[f] 3'CUCG**AAA**GCUG	8.14±0.07	77.44±3.65	223.45±11.59	43.6	8.03±0.21	68.80±7.05	195.94±22.29	43.9
5'GAGC **AA** CGAC 3'CUCG**AAA**GCUG	7.99±0.05	70.79±2.61	202.48±8.28	43.5	8.03±0.09	71.47±2.86	204.56±9.09	43.6
5'UGAC **UU** CUCA[e] 3'ACUG**UUU**GAGU	7.36±0.02	74.29±2.06	215.79±6.64	40.5	7.43±0.13	79.43±5.11	232.14±16.14	40.6
5'ACCU **GC** UUGC[b] 3'UGGA**ACA**AACG	7.33±0.02	74.13±2.03	215.38±6.48	40.4	7.42±0.09	77.88±6.40	227.17±20.36	40.6
5'UGAC **UU** CUCA 3'ACUG**CUU**GAGU	6.99±0.01	73.34±1.22	213.93±3.92	39.0	7.02±0.06	69.06±4.73	200.03±15.38	39.2
5'UGAG **AA** GUCA 3'ACUC**CGA**CAGU	6.58±0.03	60.79±2.15	174.78±6.98	37.3	6.69±0.13	56.93±7.28	161.96±23.89	37.9
5'UGAC **UU** CUCA[c,e] 3'ACUG**CCU**GAGU	6.32±0.02	73.70±1.97	217.27±6.40	36.1	6.41±0.17	63.73±6.21	184.80±20.30	36.4
5'CUGU **GG** ACGA[f] 3'GACG**AGA**UGCU	5.45±0.05	59.48±1.83	174.22±6.06	31.4	5.47±0.12	60.89±4.78	178.67±15.73	31.7
5'UCAC **UU** CUGA[f] 3'AGUG **CUC** GACU	5.35±0.06	69.36±2.45	206.41±8.1	31.8	5.54±0.18	62.02±3.58	182.10±12.06	32.1
5'UCAG **CC** GUGA[c] 3'AGUC **AAU**CACU	5.11±0.07 (5.74±0.07)	55.57±2.05 (61.25±3.10)	162.69±6.82 (178.97±10.2)	29.3 (33.1)	5.10±0.15 (5.85±0.14)	58.38±5.91 (58.43±3.70)	171.78±19.51 (169.53±12.3)	29.6 (33.4)
1X3 loops								
5'AAGGC **U** CGGAA[f] 3'UUCCG**UUU**GCCUU	11.60±0.09	88.65±1.63	248.42±4.97	55.8	11.90±0.30	94.17±5.77	265.25±17.66	55.8
5'CCUCU **C** GGUGA[c] 3'GGAGA**AAA**CCGCU	9.86±0.11	90.89±3.33	261.25±10.4	48.8	9.45±0.09	76.90±2.98	217.48±9.72	49.2
5'CCAC **G** CUCC[b,f] 3'GGUG**AAA**GAGG	9.64±0.04	77.24±1.23	217.94±3.82	50.0	9.88±0.38	82.97±9.51	235.66±29.46	50.1
5'AGGC **U** CGGA[e] 3'UCCG**UUU**GCCU	9.25±0.05	77.43±1.83	219.83±5.72	48.3	9.19±0.11	74.28±4.45	209.86±14.02	48.5
5'CCAC **A** CUCC 3'GGUG**AAA**GAGG	9.18±0.04	76.63±1.50	217.48±4.7	48.1	9.26±0.14	78.30±3.39	222.60±10.50	48.2

Table 1 (Continued)[a]

duplex	T_M^{-1} vs Ln (C_T/4) PARAMETERS $-\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m (°C)	CURVE FIT PARAMETERS $-\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m (°C)
5'UCCG **A** CGCA[f] 3'AGGC**AAG**GCGU	9.02±0.04	70.03±1.38	196.70±4.32	48.4	8.98±0.06	68.30±1.75	191.25±5.58	48.6
5'CGAC **G** GCAG[e] 3'GCUG**GAA**CGUC	8.34±0.02	70.84±1.02	201.52±3.23	45.1	8.34±0.07	69.17±3.23	196.14±10.25	45.3
5'UCCG **A** CGCA[f] 3'AGGC**GGA**GCGU	8.13±0.02	62.13±1.20	174.10±3.80	45.2	8.11±0.08	59.25±3.47	164.88±11.10	45.5
5'CGAC **A** GCAG[f] 3'GCUG**GAA**CGUC	8.02±0.03	71.34±1.79	204.18±5.70	43.6	7.97±0.11	68.07±6.40	193.77±20.35	43.7
5'AAGGU **C** UGGAA[b] 3'UUCCA **UUU** ACCUU	7.87±0.01	71.74±0.82	205.93±2.61	42.9	7.91±0.06	73.79±1.65	212.41±5.14	42.9
5'UCCU **G** UGCA 3'AGGA**GAG**ACGA	6.99±0.01	58.47±1.11	165.99±3.57	39.4	7.02±0.04	57.48±2.78	162.69±9.03	39.7
5'UGAC **U** CUCA 3'ACUG**CUU**GAGU	6.86±0.01	59.23±0.86	169.13±2.79	38.7	6.90±0.11	53.81±4.01	151.25±13.23	39.2
5'UCCU **A** UGCA[b] 3'AGGA**GAG**ACGA	6.62±0.03	58.13±2.20	166.06±7.13	37.5	6.66±0.06	57.61±2.71	164.30±8.78	37.7
5'UGAC **A** CUCA[c,f] 3'ACUG**AAC**GAGU	6.60±0.02	57.57±1.67	164.36±5.42	37.3	6.68±0.07	56.08±6.05	159.28±19.70	37.8
5'UCCU **A** UGCA[b,f] 3'AGGA**AAG**ACGU	6.56±0.02	58.53±1.97	167.58±6.36	37.1	6.60±0.07	59.32±6.26	169.97±20.19	37.4
5'UGAG **A** GUCA 3'ACUC**CGA**CAGU	6.55±0.03	60.98±1.90	175.48±6.18	37.1	6.68±0.17	52.54±5.67	147.87±18.59	37.9
5'UGAC **C** CUCA 3'ACUG**CUU**GAGU	6.34±0.02	57.56±1.49	165.15±4.86	36.0	6.45±0.14	55.13±6.78	156.96±22.23	36.5
5'UCCU **G** UGCA[b] 3'AGGA**AAA**ACGU	6.18±0.02	52.83±1.41	150.44±4.58	34.9	6.21±0.05	54.24±3.65	154.86±11.81	35.2
5'UCCU **A** UGCA 3'AGGA**GAA**ACGU	6.17±0.02	50.69±1.37	143.52±4.48	34.8	6.20±0.04	52.93±4.07	150.64±13.20	35.1
5'UCCU **A** UGCA[b] 3'AGGA**AGA**ACGU	6.09±0.04	57.40+2.06	165.43±6.74	34.7	6.18±0.10	55.86±4.56	160.19±14.97	35.0
5'UCCU **A** UGCA 3'AGGA**AAA**ACGU	6.09±0.03	48.88±1.46	137.95±4.76	34.3	6.10±0.03	52.99±4.18	151.17±13.51	34.5
5'GGC **U** CGG[f] 3'CCG**UUU**GCC	6.03±0.03	56.93±1.40	164.11±4.59	34.3	6.07±0.06	59.18±7.02	171.23±22.69	34.6
5'UCAG **C** GUGA[c,f] 3'AGUC **AAU** CACU	5.73±0.04	58.07±1.69	168.75±5.56	32.8	5.90±0.15	50.09±2.13	142.48±7.01	33.1
5'UCAC **U** CUGA 3'AGUG **CUC** GACU	5.30±0.07	55.88±2.21	163.08±7.35	30.3	5.47±0.19	51.62±3.97	148.79±13.40	30.8
5'ACCU **C** UUGC[c] 3'UGGA**ACA**AACG	5.01±0.04	55.83±1.15	163.83±3.81	28.8	5.34±0.24	47.11±4.95	134.68±16.67	29.4
1X2 Loops								
5'AGGC **U** CGGA[f] 3'UCCG**UU**GCCU	10.43±0.08	77.96±1.85	217.73±5.72	53.3	10.87±0.31	87.40±4.25	246.73±12.72	53.2
5'UGAG **C** GUCA[c,d] 3'ACUC**CC**CAGU	6.62±0.01 (7.66±0.04)	54.93±0.84 (66.69±2.53)	155.75±2.72 (190.33±8.05)	37.5 (42.4)	6.67±0.06 (7.65±0.12)	54.96±3.72 (62.37±3.96)	155.68±12.1 (176.42±12.8)	37.8 (42.7)
5'UCAG **A** GUGA[e] 3'AGUC**AG**CACU	6.43±0.03	63.10±2.24	182.70±7.25	36.5	6.50±0.08	61.42±5.28	177.05±17.24	36.9
5'UCAG **U** GUGA[f] 3'AGUC**CU**CACU	5.92±0.03	65.03±1.74	190.61±5.68	34.1	5.99±0.10	62.82±3.85	183.26±12.68	34.3
Stem Duplexes								
5'AAGGCCGGAA 3'UUCCGGCCUU	18.16±0.37 [16.86±0.30][h]	98.34±3.31 [91.2±5.7][h]	258.54±9.50 [239.8±18][h]	78.6 [76.5][h]	17.00±0.32	87.95±2.23	228.76±6.16	78.9
5'CCUCUGGUGA 3'GGAGACCGCU	15.32±0.28	93.46±3.20	251.96±9.47	69.2	15.42±0.16	94.49±1.63	254.94±4.79	69.2

Table 1 (Continued)

duplex	T_M^{-1} vs Ln (C_T/4) PARAMETERS $-\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m (°C)	CURVE FIT PARAMETERS $-\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m (°C)
5'AGGCCGGA[f] 3'UCCGGCCU	15.26±0.56 [15.29±0.29][h]	84.09±5.54 [77.56±5.6][h]	221.91±16.08 [201.8±17][h]	72.9 [74.9][h]	14.95±1.30	81.13±12.02	213.39±34.64	72.9
5'UCCGCGCA 3'AGGCGCGU	14.59±0.23 [14.29±0.29][h]	81.15±2.47 [76.26±5.6][h]	214.61±7.22 [199.8±17][h]	71.2 [72.1][h]	14.62±0.17	81.35±1.67	125.17±4.85	71.2
5'GAGCCGAC[f] 3'CUCGGCUG	13.94±0.14 [13.97±0.28][h]	80.13±1.52 [82.06±5.5][h]	213.99±4.47 [219.4±17][h]	68.5 [68.1][h]	14.34±0.25	84.66±2.42	226.75±7.01	68.5
5'GUCGCAGC 3'CAGCGUCG	13.93±0.21 [13.89±0.28][h]	86.18±2.44 [81.55±5.5][h]	232.95±7.20 [218.0±17][h]	66.1 [68.0][h]	13.80±0.14	84.49±1.86	227.92±5.59	66.2
5'UUCCAACCUU[g] 3'AAGGUUGGAA	12.69±0.17	75.84±2.18	203.62±6.48	64.4	13.06±0.17	80.74±1.52	218.12±4.46	64.3
5'CGACGCAG[f,g] 3'GCUGCGUC	12.32±0.12	70.45±1.55	187.43±4.60	64.8	12.92±.37	77.68±3.07	208.80±8.78	64.8
5'CUCGCACA[g] 3'GAGCGUGU	12.11±0.13	72.61±1.75	195.09±5.24	62.8	12.75±0.22	81.07±1.87	220.30±5.36	62.8
5'UGCAAGGA[g] 3'ACGUUCCU	11.09±0.11	70.27±1.77	190.82±5.37	58.5	11.51±0.28	76.59±3.46	209.85±10.28	58.6
5'UCACCUGA[f] 3'AGUGGACU	10.84±0.12 [11.51±0.29][h]	71.51±2.19 [69.98±5.6][h]	195.62±6.66 [188.60±17][h]	56.9 [60.6][h]	11.12±0.06	75.37±8.52	207.17±25.53	57.1
5'ACCUUUGC[g] 3'UGGAAACG	10.64±0.17	77.42±3.39	215.3±10.39	54.4	10.75±0.33	78.08±6.46	217.08±19.79	54.7

[a] Melting experiments were done in 1 M NaCl, 10 or 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA buffer solution, pH 7. Sequences are listed in order of type of loop and then decreasing duplex stability. Estimated errors from all sources are ±10%, ±10%, ±2%, and ±1 °C for ΔH°, ΔS°, ΔG°_{37}, and T_m, respectively. Significant figures beyond error estimates are given to allow accurate calculation of T_M and other parameters. T_m listed is for total strand concentration of 1.0×10^{-4} M. [b] These duplexes were also melted in 10 mM $MgCl_2$, 0.15 M KCl, 10 mM cacodylate, and 0.5 mM Na_2EDTA buffer, pH 7.0, at one or more concentrations. No significant change in duplex stability relative to 1 M NaCl buffer was observed. [c] These duplexes were also melted in 1 M NaCl, 50 mM MES, and 0.5 mM Na_2EDTA buffer, pH 5.5, at one or more concentrations. Where the change in pH caused a change in duplex stability, the thermodynamic parameters from pH 5.5 melts are listed in parentheses. [d] Reference *26*. [e] These duplexes were studied by variable temperature imino proton NMR (500 MHz) and one-dimensional NOE NMR in 80 mM NaCl, 10 mM phosphates, and 0.5 mM Na_2EDTA buffer, pH 6.7, and 10% D_2O. [f] These duplexes were studied by variable temperature imino proton NMR (500 MHz) in 80 mM NaCl, 10 mM phosphates, and 0.5 mM Na_2EDTA buffer, pH 6.7, and 10% D_2O. [g] The thermodynamic parameters for these stem duplexes were included in an analysis of nearest-neighbor parameters (*54*). [h] These duplexes were not included in the analysis of nearest-neighbor parameters by Xia et al. (*54*). Listed in brackets below the experimental values are the predicted thermodynamic parameters for duplex formation, calculated from the nearest-neighbor parameters (*54*).

a sample graph showing the distribution of non-self-complementary duplexes, self-complementary duplexes, and single strands versus temperature.

As noted in Table 1, 14 duplexes were measured in pH 5.5 or magnesium buffers. No significant changes in the thermodynamic parameters were noted with magnesium buffers. Only two duplexes showed a modest increase of stability in pH 5.5 buffer. The duplex 5′UCAG**CC**GUGA/3′AGUC**AAU**CACU showed an increase of 4 °C in T_m and a 0.6 kcal/mol more favorable ΔG°_{37} at pH 5.5, which is similar to a previous report of increased stability for symmetric 2 × 2 loops containing AC mismatches (*22*). The duplex 5′UGAC**C**CUCA/3′ACUG**CC**GAGU containing a 1 × 2 loop of all C nucleotides showed an increase of 5 °C in T_m and 1 kcal/mol more favorable ΔG°_{37} at pH 5.5 as compared to values previously reported for this loop at pH 7 (*26*). This increase in stability is slightly less than half the stability increase for a 2 × 2 loop containing all C nucleotides, in which case an increase of 12 °C in T_m and 2.5 kcal/mol more favorable ΔG°_{37} were observed at pH 5.5 (*22*).

Table 2 lists the thermodynamic parameters for formation of the internal loops. The thermodynamic parameters for formation of an internal loop are calculated according to the following equation (*17*):

$$\Delta G^\circ_{\text{loop}} = \Delta G^\circ_{\text{duplex with loop}} - (\Delta G^\circ_{\text{duplex without loop}} - \Delta G^\circ_{\text{interrupted base pair}}) \quad (2a)$$

For example

$$\Delta G^\circ \begin{smallmatrix} 5'C\,^{GA}C \\ 3'G_{AAG}G \end{smallmatrix} =$$

$$\Delta G^\circ \begin{smallmatrix} 5'GAGC\,^{GA}CGAC \\ 3'CUCG_{AAG}GCUG \end{smallmatrix} - \left(\Delta G^\circ \begin{smallmatrix} 5'GAGCCGAC \\ 3'CUCGGCUG \end{smallmatrix} - \Delta G^\circ \begin{smallmatrix} 5'CC \\ 3'GG \end{smallmatrix} \right) \quad (2b)$$

The values for $\Delta H^\circ_{\text{loop}}$ and $\Delta S^\circ_{\text{loop}}$ are similarly derived. The values for $\Delta G^\circ_{\text{duplex with loop}}$ and the corresponding ΔH° and ΔS° are derived from the $1/T_M$ versus ln (C_T/4) plots (Table 1). The last term in the equation accounts for the nearest neighbor stacking interaction present in the uninterrupted helix but absent in the helix containing an internal loop. These Watson−Crick nearest neighbor parameters were taken from Xia et al. (*54*). The values for $\Delta G^\circ_{\text{duplex without loop}}$ were either experimentally measured or predicted from the nearest-neighbor approximation (*54*). For duplexes not

Table 2: Thermodynamic Parameters for Internal Loop Formation[a]

	$\Delta G^\circ_{loop,37}$ (kcal/mol)	ΔH°_{loop} (kcal/mol)	ΔS°_{loop} (eu)
2X3 loops			
5'GAGC **GA** CGAC 3'CUCG**AAG**GCUG	0.17±0.18 (0.38)	-19.78±3.12	-64.38±9.52
5'CCAC **GG** CUCC 3'GGUG**AGA**GAGG	1.37±0.30 (1.69)	-13.98±8.42	-49.51±22.58
5'GAGC **AA** CGAC 3'CUCG**AAG**GCUG	1.45±0.17 (1.24)	-7.71±3.32	-29.57±10.23
5'CGAC **GA** GCAG 3'GCUG**GAA**CGUC	1.53±0.16 (1.86)	-14.85±3.11	-52.79±9.56
5'CCUCU **GC** GGUGA 3'GGAGA**AAA**CCGCU	1.63±0.33 (2.3)	-5.48±4.63	-22.97±13.90
5'UGAC **UU** CUCA 3'ACUG**UUU**GAGU	1.72±0.07 (1.91)	-11.59±7.97	-42.99±21.88
5'GAGC **AG** CGAC 3'CUCG**GAA**GCUG	1.88±0.17 (1.98)	-12.31±3.78	-45.78±11.75
5'GAGC **GA** CGAC 3'CUCG**AAA**GCUG	1.91±0.17 (1.69)	-8.15±3.25	-32.48±10.02
5'CCAC **GG** CUCC 3'GGUG**AAA**GAGG	1.99±0.30 (1.69)	-10.27±8.50	-39.53±22.88
5'GAGC **AA** CGAC 3'CUCG**GAA**GCUG	2.07±0.16 (1.98)	-11.12±3.28	-42.57±10.16
5'UGAC **UU** CUCA 3'ACUG**CUU**GAGU	2.09±0.07 (2.23)	-10.64±7.80	-41.13±21.21
5'GAGC **AG** CGAC 3'CUCG**AAG**GCUG	2.18±0.18 (1.86)	-10.48±4.51	-40.87±14.11
5'UCAC **UU** CUGA 3'AGUG**CUC**GACU	2.23±0.15 (2.55)	-11.24±3.51	-43.49±11.14
5'ACCU **GC** UUGC 3'UGGA**ACA**AACG	2.38±0.18 (2.95)	-3.53±4.0	-19.08±12.49
5'UCAG **CC** GUGA 3'AGUC**AAU**CACU	2.47±0.16 (2.55)	2.55±3.24	0.23±10.26
5'UGAG **AA** GUCA[b] 3'ACUC**AAA**CAGU	2.49±0.21 (2.55)	2.25±9.79	-0.85±27.14
5'UGAG **AA** GUCA 3'ACUC**CGA**CAGU	2.50±0.08 (2.55)	1.91±8.00	-1.98±21.99
5'GAGC **AG** CGAC 3'CUCG**AAA**GCUG	2.54±0.17 (2.55)	-10.52±4.15	-42.16±12.99
5'CUGU **GG** ACGA 3'GACG**AGA**UGCU	2.68±0.58 (2.05)	-3.08±7.02	-18.62±22.47
5'GAGC **AA** CGAC 3'CUCG**AAA**GCUG	2.69±0.16 (2.55)	-3.87±3.26	-21.19±10.15
5'UGAC **UU** CUCA 3'ACUG**CCU**GAGU	2.76±0.07 (2.23)	-11.0±7.95	-44.47±21.81
1X3 loops			
5'CCAC **G** CUCC 3'GGUG**AAA**GAGG	1.57±0.30 (2.36)	-9.3±8.32	-35.04±22.22
5'CGAC **G** GCAG 3'GCUG**GAA**CGUC	1.62±0.15 (2.36)	-11.03±2.48	-40.79±7.52
5'UCAG **C** GUGA 3'AGUC**AAU**CACU	1.85±0.15 (2.36)	0.05±3.03	-5.83±9.47
5'CGAC **A** GCAG 3'GCUG**GAA**CGUC	1.94±0.15 (2.36)	-11.53±2.89	-43.45±8.87
5'CCAC **A** CUCC 3'GGUG**AAA**GAGG	2.03±0.30 (2.36)	-8.69±8.36	-34.58±22.39
5'UCCG **A** CGCA 3'AGGC**AAG**GCGU	2.15±0.25 (2.36)	-3.76±3.24	-18.99±9.74
5'GGC **U** CGG 3'CCG**UUU**GCC	2.18±0.29 (2.36)	-8.24±5.50	-33.61±17.00
1X3 loops cont.			
5'UGAC **U** CUCA 3'ACUG**CUU**GAGU	2.22±0.07 (2.36)	3.47±7.75	3.67±21.03
5'UCAC **U** CUGA 3'AGUG**CUC**GACU	2.28±0.16 (2.36)	2.24±3.35	-0.16±10.62
5'UGAC **A** CUCA 3'ACUG**AAC**GAGU	2.48±0.07 (2.36)	5.13±7.88	8.44±21.54
5'UGAG **A** GUCA 3'ACUC**CGA**CAGU	2.53±0.08 (2.36)	1.72±7.93	-2.68±21.75
5'UGAC **A** CUCA[b] 3'ACUG**AAA**GAGU	2.53±0.20 (2.36)	5.93±9.57	10.88±26.40
5'UGAC **C** CUCA 3'ACUG**CUU**GAGU	2.74±0.07 (2.36)	5.14±7.84	7.65±21.41
5'AGGC **U** CGGA 3'UCCG**UUU**GCCU	2.75±0.56 (2.36)	-6.73±5.96	-30.62±17.49
5'UGAG **A** GUCA[b] 3'ACUC**AAA**CAGU	2.92±0.20 (2.36)	12.93±9.17	32.2±25.15
5'UCCG **A** CGCA 3'AGGC**GGA**GCGU	3.04±0.25 (2.36)	4.14±3.17	3.61±9.51
5'AAGGC **U** CGGAA 3'UUCCG**UUU**GCCUU	3.30±0.38 (2.36)	-3.7±3.89	-22.58±11.37
5'C **X** C 3'G**XXX**G (average for 17 loops)	2.36±0.49		
5'UCCU **G** UGCA 3'AGGA**GAG**ACGA	3.17±0.11 (3.89)	4.98±2.24	5.83±6.92
5'UCCU **A** UGCA 3'AGGA**GAG**ACGA	3.54±0.12 (3.89)	5.32±2.94	5.76±9.27
5'UCCU **A** UGCA 3'AGGA**AAG**ACGU	3.6±0.11 (3.89)	4.92±2.76	4.24±8.70
5'AAGGU **C** UGGAA 3'UUCCA**UUU**ACCUU	3.89±0.17 (3.89)	-2.72±2.46	-21.31±7.42
5'UCCU **G** UGCA 3'AGGA**AAA**ACGU	3.98±0.11 (3.89)	10.62±2.40	21.38±7.49
5'UCCU **A** UGCA 3'AGGA**GAA**ACGU	3.99±0.11 (3.89)	12.76±2.39	28.3±7.43
5'UCCU **A** UGCA 3'AGGA**AGA**ACGU	4.07±0.12 (3.89)	6.05±2.83	6.39±8.98
5'UCCU **A** UGCA 3'AGGA**AAA**ACGU	4.07±0.12 (3.89)	14.57±2.43	33.87±7.60
5'ACCU **C** UUGC 3'UGGA**ACA**AACG	4.7±0.18 (3.89)	14.77±3.66	32.47±11.34
5'U **X** U 3'A**XXX**A (average for 9 loops)	3.89±0.39		
5'CCUCU **C** GGUGA 3'GGAGA**AAA**CCGCU	3.35±0.31 (3.12)	-7.87±4.81	-36.19±14.60
1X2 loops			
5'UCAG **A** GUGA 3'AGUC**AG**CACU	1.15±0.14	-4.98±3.36	-19.78±10.55
5'AGGC **U** CGGA 3'UCCG**UU**GCCU	1.57±0.57	-7.26±5.97	-28.52±17.48
5'UCAC **U** CUGA 3'AGUG**CU**GACU	1.66±0.14	-6.91±3.06	-27.69±9.54

[a] Melting experiments were done in 1 M NaCl, 10 or 20 mM sodium cacodylate, and 0.5 mM EDTA buffer solutions, pH 7. Sequences are listed in order of internal loop free energy for each type of loop. The experimental loop free energy is calculated according to the following equation (*17*): $\Delta G^\circ_{loop} = \Delta G^\circ_{duplex\ with\ loop} - (\Delta G^\circ_{duplex\ without\ loop} - \Delta G^\circ_{interrupted\ base\ pair})$. Thermodynamic values used in this equation are from T_M^{-1} vs $\ln(C_T/4)$ plots. Estimated errors for the thermodynamic parameters of the interrupted base pair are ±10%, ±10%, and ±2% for ΔH°, ΔS°, and ΔG°_{37}, respectively. The predicted loop free energies, calculated according to the rules presented in this paper and summarized in Table 3, are listed in parentheses below the experimental loop free energies. [b] Reference *19*.

Table 3: Free Energy Parameters[a] for Predicting Stabilities of 2 × 3 and 1 × 3 Internal Loops at 37 °C

parameter	2 × 3 loop	1 × 3 loop
$\Delta G^\circ_{\text{loop initiation}} + \Delta G^\circ_{\text{asymmetry penalty}}$	2.55 ± 0.11[b]	2.36 ± 0.49
$\Delta G^\circ_{\text{AU/GU penalty}}$[c]	0.65[d]	0.77
$\Delta G^\circ_{\text{GA/GG/UU bonus}}$		
5′GG 3′CA	−1.31 ± 0.23[b]	0
5′CG 3′GA	−0.87 ± 0.18[b]	0
5′WG 3′SG	−0.70 ± 0.24[b]	0
5′CA 3′GG	−0.58 ± 0.24[b]	0
5′WU 3′SU	−0.32 ± 0.14[b]	0

[a] The parameters in this table are used with the following equation (*1*): $\Delta G^\circ_{\text{loop (predicted)}} = \Delta G^\circ_{\text{loop initiation}} + \Delta G^\circ_{\text{asymmetry penalty}} + \Delta G^\circ_{\text{AU/GU penalty}} + \Delta G^\circ_{\text{GA/GG/UU bonus}}$. Values are given in kilocalories per mole. [b] The values for $\Delta G^\circ_{\text{GA/GG/UU bonus}}$ were derived from a linear regression analysis. The r^2 was 0.842. The analysis has 18 data points, 12 degrees of freedom, and a confidence level of 0.05 for the six parameters. The parameter for 5′GA/3′CG was not statistically significant ($P > 0.05$) and was approximately zero. Thus, this motif is not given any bonus. The intercept of the linear equation represents the sum of the parameters for loop initiation and asymmetry. WS represents GC base pairs. [c] The AU/GU penalty includes the 0.45 ± 0.04 kcal/mol penalty applied to each AU or GU pair that terminates a helix (*54*). When a GA mismatch occurs adjacent to an AU or GU base pair closing the loop, the $\Delta G^\circ_{\text{GA bonus}}$ is approximated as −0.9 kcal/mol, the average of the three parameters for $\Delta G^\circ_{\text{GA bonus}}$ when the closing base pairs are CG. [d] References *1* and *54*.

included in a recent analysis of nearest-neighbor parameters (*54*), the predicted thermodynamic parameters for duplex formation were calculated from the new nearest-neighbor parameters and are listed in brackets below the experimental values. Although the 10-mer duplex is more stable than predicted, the predictions for the 8-mer duplexes are all within experimental error.

Models for Predicting Loop Stability. Listed in parentheses in Table 2 are the free energies of loop formation predicted by the models described below. The thermodynamic stability of an internal loop is predicted with the following equation (*1*):

$$\Delta G^\circ_{\text{loop (predicted)}} = \Delta G^\circ_{\text{loop initiation}} + \Delta G^\circ_{\text{asymmetry penalty}} + \Delta G^\circ_{\text{AU/GU penalty}} + \Delta G^\circ_{\text{GA/GG/UU bonus}} \quad (3)$$

The loop initiation term accounts for the entropy penalty of forming a loop closed by two GC pairs and depends on loop size. The asymmetry penalty increases with an increasing difference in the number of nucleotides on each side of the loop (*19*). The AU/GU penalty accounts for the different number of hydrogen bonds in AU or GU base pairs closing loops compared with GC base pairs (*54*). On the basis of studies on symmetric 2 × 2 internal loops (*1, 23*), the AU/GU penalty is 0.2 kcal/mol in addition to a 0.45 ± 0.04 kcal/mol penalty for terminating a helix with an AU base pair. The GA/GG/UU bonus accounts for the potential stabilizing effects of GA, GG, and UU mismatches when these mismatches begin an internal loop. The value of this bonus depends on sequence, as described below and in Table 3.

The parameters for GA, GG, and UU bonuses in 2 × 3 loops are derived from a linear regression analysis of the free energy for formation of 2 × 3 loops with CG closing base pairs (Table 3). The r^2 for the linear regression is 0.842. The intercept of the linear equation represents the sum of the parameters for loop initiation and asymmetry; the value of 2.55 ± 0.11 kcal/mol resulting from the regression analysis is 0.3 kcal/mol less stable than previously estimated (*1*). The data set for the regression analysis includes only 18 points and each bonus parameter depends on only a few data points. For example, only two loops contain possible GG mismatches and three loops contain possible UU mismatches. Thus, the bonus parameters for UU and GG mismatches make no distinction between CG and GC closing base pairs (Table 3); parameters that depend on whether the closing base pair is CG or GC were not statistically significant ($P > 0.05$) for UU and GG mismatches. In contrast, 11 loops contain possible GA mismatches; a statistical analysis reveals that the additional stability of a GA mismatch depends on the stacking interaction between the mismatch and the closing base pairs of the loop (Table 3). Although 5′GA/3′CG occurs four times in the data set, the parameter for 5′GA/3′CG was approximately 0 and not statistically significant ($P > 0.05$). Thus, this motif is not given any bonus. Possible AC and UC mismatches did not increase the stability of the loop compared to a loop with all A nucleotides.

The experimental free energy of internal loop formation for 1 × 3 loops depends mainly on the closing base pairs of the loop rather than the sequence within the loop (Table 2). For example, the loop 5′G**A**C/3′G**AAG**C is 1.5 kcal/mol more stable than 5′U**A**U/3′A**AAG**A but has similar stability to 5′ C**A**G/3′G**GAA**C, 5′G**A**G/3′C**CGA**C, and 5′C**A**C/3′G**AAA**G. The stabilities of loops 5′U**A**U/3′A**GAA**A, 5′U**A**U/3′A**AGA**A, and 5′U**A**U/3′A**AAA**A are similar to that of 5′U**A**U/3′A**AAG**A. The average free energies for the formation of 1 × 3 internal loops are 2.4 ± 0.5 kcal/mol and 3.9 ± 0.4 kcal/mol when both closing base pairs are CG or AU, respectively. These averages are consistent with eq 3 if the GA and UU mismatch bonuses are 0 for 1 × 3 loops. The previous value for $\Delta G^\circ_{\text{loop initiation}} + \Delta G^\circ_{\text{asymmetry penalty}}$ for 1 × 3 loops closed by GC pairs was 2.7 ± 0.5 kcal/mol (*1*). According to eq 3, the previously predicted $\Delta G^\circ_{\text{loop (predicted)}}$ for a 1 × 3 loop closed by two AU closing base pairs was 4.0 ± 0.5 kcal/mol.

Effects of Stem Length. Table 4 shows the thermodynamic parameters for loop formation in duplexes that have the same loop sequence and closing base pairs but stem helices of different lengths. For a 1 × 3 loop with U nucleotides, adding a base pair to each end of the helix destablizes the internal loop by an average of 0.6 kcal/mol. This is consistent with experiments on a 2 × 2 loop containing GA mismatches (*56*). The sequence of the stem helix may also affect the thermodynamic parameters for loop formation. The values for three internal loops, 5′C**A**C/3′G**AAA**G, 5′C**U**C/5′G**UU**G, and 5′C**A**G/3′G**A**C, have been measured in duplexes that have the same helix length but different sequence (Table 2; *19, 26, 56*); the average difference in $\Delta G^\circ_{\text{loop}}$ is 0.4 kcal/mol. Thus, comparisons between free energy increments for loops with different stem helices are valid within about 0.5 kcal/mol, which is close to the expected experimental error.

Imino Proton NMR. Figures 2 and 3 show imino proton spectra of several 2 × 3 internal loops containing possible GA or UU mismatches, respectively. Imino protons in AU base pairs typically resonate in the region between 13.5 and

Table 4: Free Energy of Loop Formation at 37 °C for Helices with Different Lengths[a]

duplex	$\Delta G^{\circ}_{loop,\ 37}$ (kcal/mol)	Average $\Delta\Delta G^{\circ}_{loop}$ (kcal/mol)
5'GGC **U** CGG 3'CCG**UUU**GCC	2.2	
5'AGGC **U** CGGA 3'UCCG**UUU**GCCU	2.8	0.6
5'AAGGC **U** CGGAA 3'UUCCG**UUU**GCCUU	3.3	
5'GC**GA**GC[b] 3'CG**AG**CG	-1.3	
5' GC**GA**GCG[b] 3'GCG**AG**CG	-0.9	0.5
5'GGC**GA**GCC[b] 3'CCG**AG**CGG	-0.3	
5' GC**AA**GCG[b] 3'GCG**AA**CG	1.1	0.2
5'CGC**AA**GCG[b] 3'GCG**AA**CGC	1.3	

[a] The free energy of loop formation is calculated according to the following equation (*17*): $\Delta G^{\circ}_{loop} = \Delta G^{\circ}_{duplex\ with\ loop} - (\Delta G^{\circ}_{duplex\ without\ loop} - \Delta G^{\circ}_{interrupted\ base\ pair})$. The thermodynamic parameters required for this equation were derived from T_M^{-1} vs ln (C_T/4) or ln C_T plots based on optical melting experiments in 1 M NaCl, 10 or 20 mM sodium cacodylate, and 0.5 mM EDTA, buffer, pH 7. [b] Reference *21*.

15 ppm; imino protons in CG base pairs typically resonate in the region between 11 and 13.5 ppm (*58*). Imino protons in AU base pairs at the end of a helix are usually not observed (*19*). GC base pairs at the end of a helix also may not be observed due to exchange with water (*59*), which may be the case in Figure 2 spectra ii and viii. Resonances for imino protons at the end of a helix also broaden at a lower temperature than for protons at internal positions. The number of peaks and the chemical shifts of the peaks in the spectra are consistent with the formation of the expected duplexes. Where noted in Table 1, the assignments of the imino protons were confirmed by one-dimensional NOE difference spectra.

Imino protons in UU mismatches in internal loops resonate near 11.3 and 10.4 ppm (*22*, *60*). The imino protons in imino hydrogen-bonded GA mismatches in internal loops resonate near 11.5 ppm, and imino protons in sheared GA mismatches in internal loops resonate near 10 ppm (*24*). GG mismatches would be difficult to distinguish from GA mismatches in an imino proton spectrum because the resonances for protons in GG mismatches may occur in the same frequency range. The small peaks upfield in the spectra shown in Figure 2 spectra i, ii, iv, and v and Figure 3 spectra i−iii are assigned to nucleotides within the internal loop. The imino protons may form a hydrogen bond with a base in the loop or the phosphate backbone, or the imino protons may be positioned into the loop in such a way that they are protected from exchange with water. The observation of these peaks supports the hypothesis that 2 × 3 loops allow formation of ordered mismatches at the helix−loop junction; this provides a structural rationale for the GA/GG/UU bonus in eq 3.

In one-dimensional NOE difference experiments, most of the small peaks assigned to imino protons within the loop did not show any NOEs to imino protons in the adjacent base pairs or to other imino protons within the loop. In the duplex 5′UGAC**UU**CUCA/3′ACUG**UUU**GAGU, however, the imino protons between 10 and 11.5 ppm do produce NOEs, as shown in Figure 4 spectra viii and ix. The asterisks indicate NOEs upon saturation of the resonance that is marked by an arrow. The question marks in spectra vii and viii may be NOEs between a UU mismatch or spillover effects from saturation of the nearby frequency. Spillover effects are also seen in spectra iii−vi. The pattern of NOEs in spectra i−vi is consistent with formation of the expected helices surrounding the loop. The imino protons in the terminal AU base pairs are not observed. Irradiating the imino proton of each internal AU base pair gives an NOE to an AH2 proton in the 7−9 ppm region. Thus, the NOE results confirm the formation of the expected duplex and suggest that this 2 × 3 loop with all U nucleotides has some degree of structure.

Several duplexes containing 1 × 2 and 1 × 3 loops were also analyzed by NMR, and imino proton spectra confirm the formation of the expected duplex and internal loop. Most spectra of 1 × 2 and 1 × 3 loops, however, showed no resonances assignable to the internal loop. Imino proton spectra for 1 × 2 and 1 × 3 loops are available in Supporting Information.

Imino proton spectra for three duplexes previously thought to contain 1 × 2 internal loops and for which the thermodynamic parameters were previously reported (*26*) did not show resonances consistent with the formation of the expected duplex and internal loop. Each of these duplexes shares a similar sequence motif that forms an unexpectedly stable alternative structure with consecutive terminal mismatches. Two of these loops, 5′G**A**G/3′C**AG**C and 5′C**U**C/5′G**CU**G, were studied with different stem helices and imino proton spectra confirmed the formation of the expected duplex. The thermodynamic parameters for the formation of these internal loops (Table 2) differ from those previously reported but are consistent with the conclusion that GA and UU mismatches can stabilize 1 × 2 internal loops. Several model duplexes designed to contain the loop 5′G**A**G/3′C**GG**C did not form the expected duplex without competition from other structures. Therefore, the loop free energy for 5′G**A**G/3′C**GG**C is estimated as 1.62 kcal/mol, which is an average of 5′G**A**G/3′C**AG**C and 5′G**A**G/3′C**GA**C. The revised parameters for the formation of these internal loops are included in the current *mfold* program (*1*). Imino proton spectra of four other duplexes containing the following 1 × 2 loops, 5′G**A**G/3′C**GA**C, 5′C**G**C/3′G**GA**G, 5′C**G**C/3′G**AA**G, and 5′C**C**C/3′G**CC**G (*26*), confirmed the formation of the expected duplexes.

Figure 5 shows imino proton spectra of 1 × 2, 1 × 3, and 2 × 3 internal loops containing all U nucleotides. The resonances between 10 and 11.2 ppm can be assigned to the U nucleotides in the internal loop. Each type of loop shows a different number of peaks for U nucleotides in the loop, although these peaks are broader than peaks for UU mismatches in 2 × 2 loops (*22*). These peaks suggest each loop has some structure, whether from UU mismatches or U nucleotides stacked into the loop and protected from exchange with water. On the basis of line widths, however,

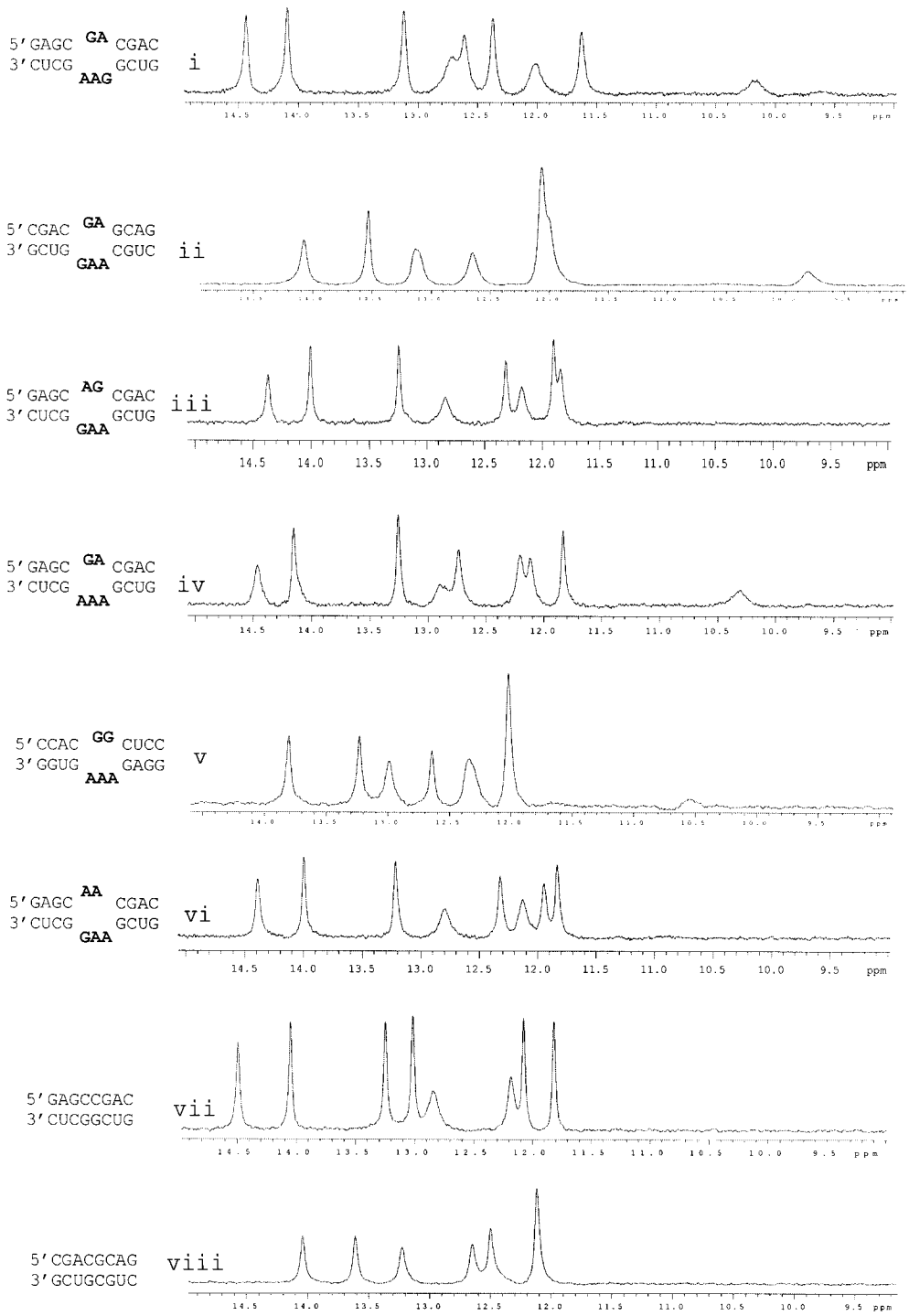

FIGURE 2: Imino proton NMR spectra (9−15 ppm) in 80 mM NaCl, 10 mM phosphate, and 0.5 mM Na_2EDTA buffer, pH 6.7. Spectra are shown in order of increasing loop free energy. The single spectrum shown from a variable-temperature study was selected for the best dispersion. (i) 5′GAGC**GA**CGAC/3′CUCG**AAG**GCUG, 0.4 mM, 5 °C, $\Delta G°_{loop\ 37C} = 0.2$ kcal/mol. (ii) 5′CGAC**GA**GCAG/3′GCUG-**GAA**CGUC, 0.7 mM, 0.5 °C, $\Delta G°_{loop\ 37C} = 1.5$ kcal/mol. (iii) 5′ GAGC**AG**CGAC/3′CUCG**GAA**GCUG, 0.5 mM, 10 °C, $\Delta G°_{loop\ 37C} = 1.9$ kcal/mol. (iv) 5′GAGC**GA**CGAC/3′CUCG**AAA**GCUG, 0.5 mM, 20 °C, $\Delta G°_{loop\ 37C} = 1.9$ kcal/mol. (v) 5′CCAC**GG**CUCC/3′GGUG**AAA**GAGG, 0.2 mM, 1 °C, $\Delta G°_{loop\ 37C} = 2.0$ kcal/mol. (vi) 5′GAGC**AA**CGAC/3′CUCG**GAA**GCUG, 0.5 mM, 10 °C, $\Delta G°_{loop\ 37C} = 2.1$ kcal/mol. (vii) 5′5′GAGCCGAC/3′CUCGGCUG, 0.4 mM, 10 °C. (viii) 5′CGACGCAG/3′GCUGCGUC, 0.4 mM, 1 °C.

these three asymmetric loops are less rigidly structured than a 2 × 2 loop with four U nucleotides. Evidently, the degree of asymmetry affects the amount of flexibility in the loop and the ability of the loop to form stabilizing mismatches. Different degrees of structure and flexibility of the loop may explain the differences in the $\Delta G°_{loop}$ observed in Table 5 for internal loops containing only A or U nucleotides.

Figure 6 shows imino proton spectra for three duplexes with the same 1 × 3 loop of all U nucleotides but different helix lengths adjacent to the loop. The imino protons in the terminal AU base pairs are not detected in spectra ii and iv; the imino protons from the AU penultimate base pairs detected in spectrum iii are seen as one broad peak just below 14 ppm. The peaks between 10 and 11.5 ppm can be assigned

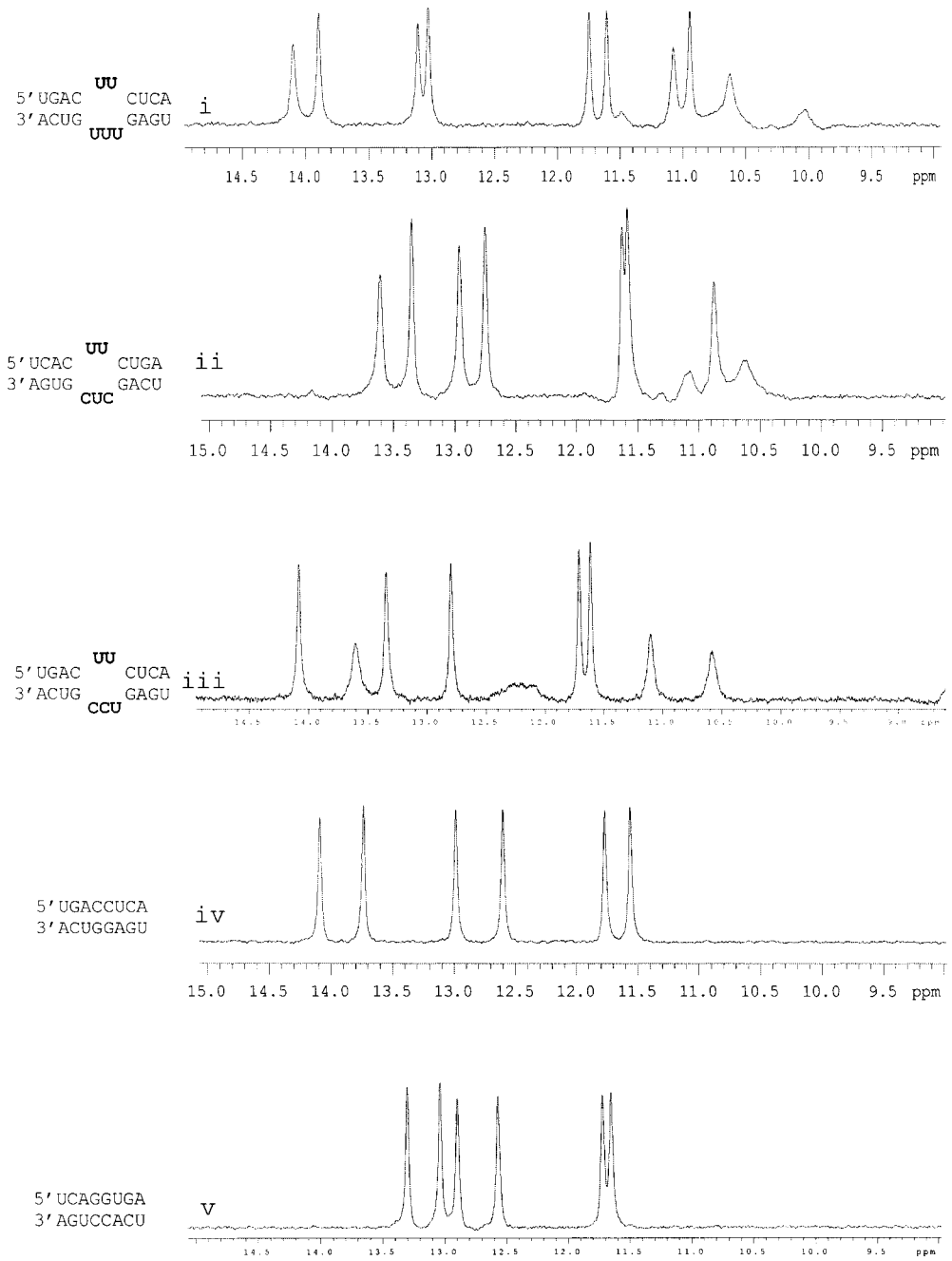

FIGURE 3: Imino proton NMR spectra (9−15 ppm) in 80 mM NaCl, 10 mM phosphate, 0.5 mM Na_2EDTA buffer, pH 6.7. Spectra are shown in order of increasing loop free energy. The single spectrum shown from a variable-temperature study was selected for the best dispersion. (i) 5′UGAC**UU**CUCA/3′ACUG**UUU**GAGU, 0.5 mM, 5 °C, $\Delta G°_{loop\ 37C} = 1.7$ kcal/mol. (ii) 5′UCAC**UU**CUGA/3′AGUG**CUC**-GACU, 0.5 mM, 1 °C, $\Delta G°_{loop\ 37C} = 2.2$ kcal/mol. (iii) 5′UGAC**UU**CUCA/3′ACUG**CCU**GAGU, 0.7 mM, 5 °C, $\Delta G°_{loop\ 37C} = 2.8$ kcal/mol. (iv) 5′UGACCUCA/3′ACUGGAGU, 0.7 mM, 5 °C. (v) 5′UCAGGUGA/3′AGUCCACU, 0.6 mM, 5 °C.

to the U residues within the loops. The three peaks between 10 and 11.5 ppm in spectrum ii may be the result of the interaction of bulged U nucleotides with the phosphate backbone or helix ends in a way that is different from the other two helices. Alternatively, increased loop flexibility may cause two of these peaks to broaden and not be observed in spectra i and iii. The dispersion of the peaks assigned to the stem helices also changes with helix length. These spectra suggest that structural changes may underlie the non-nearest-neighbor effects in $\Delta G°_{loop}$ shown in Table 4.

Natural Occurrence. Figure 1 lists the internal loops of biological interest and the small duplex models studied by optical melting. The thermodynamic parameters for loop formation do not distinguish these loops from loops with an arbitrary sequence and no known occurrence or biological role. Table 6 lists the most common 1 × 3 and 2 × 3 internal loops occurring in a database of 101 small subunit rRNA, 218 large subunit rRNA, and 79 group I introns (*61*, *62*, *87*). A total of 743 2 × 3 loops and 689 1 × 3 loops were counted. Loops that comprise less than 1% of a possible class are not listed. Only 17% and 50% of the possible loop sequences for 2 × 3 and 1 × 3 loops, respectively, occur in this database. A high occurrence of a particular loop reflects conservation of that loop during evolution but does not correlate with the thermodynamics of loop formation. For example, the most common 2 × 3 loop, **CC**(1403)/

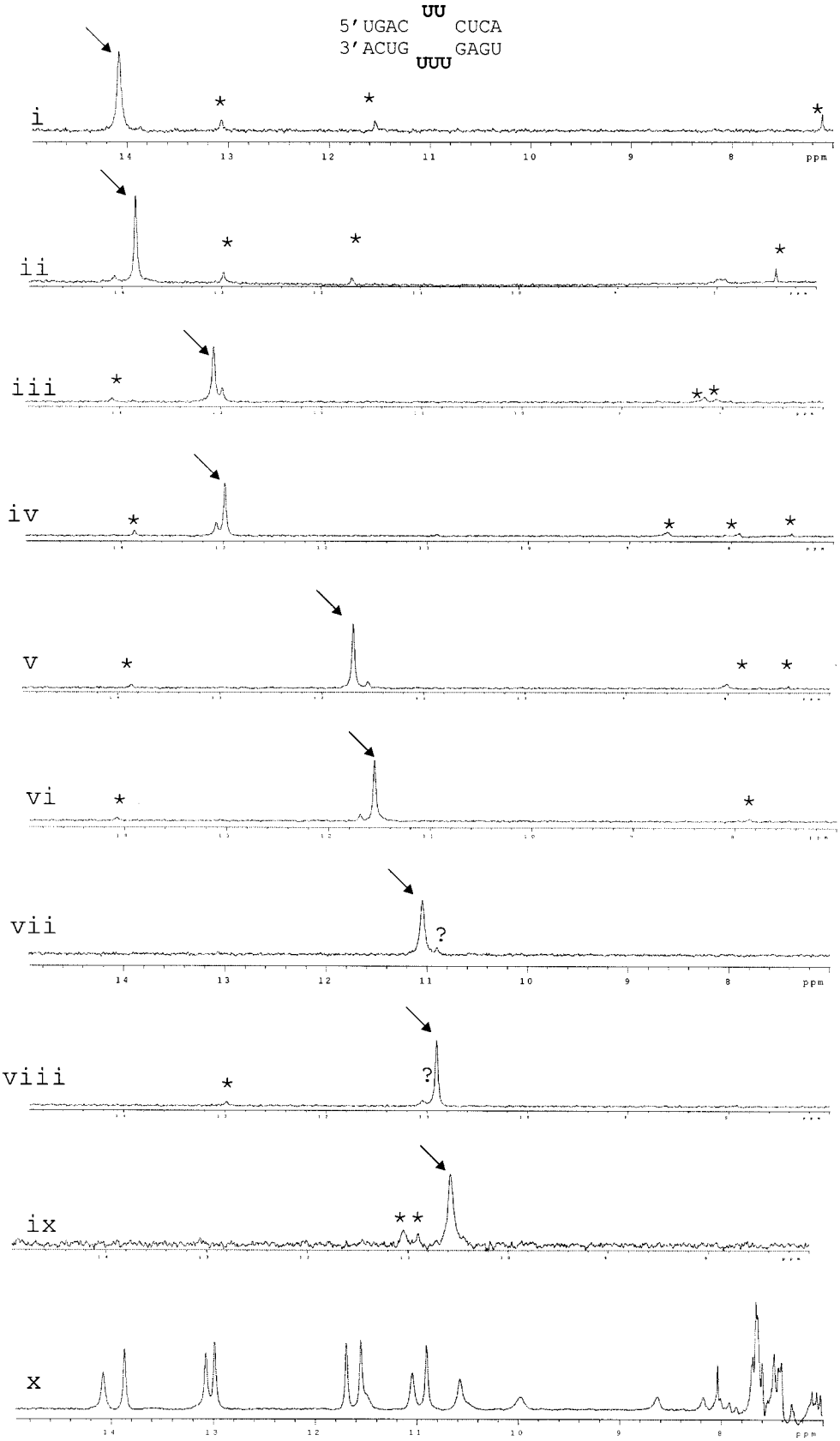

FIGURE 4: One-dimensional NOE difference NMR spectra (7−15 ppm) of 0.5 mM 5′UGAC**UU**CUCA/3′ACUG**UUU**GAGU at 0.5 °C in 80 mM NaCl, 10 mM phosphate, and 0.5 mM Na_2EDTA buffer, pH 6.7. Asterisks mark observed NOE peaks when the resonance marked with an arrow is saturated; question marks indicate small peaks that are possible spillover effects or NOEs. Spectra were acquired with 3 s saturation at the following frequencies: 14.07 ppm (i), 13.86 ppm (ii), 13.07 ppm (iii), 12.98 ppm (iv), 11.69 ppm (v), 11.55 ppm (vi), 11.04 ppm (vii), 10.90 ppm (viii), 10.57 ppm (ix), and off-resonance (x).

AAU(1502)**,** where the numbering refers to *Escherichia coli*, occurs adjacent to the stem−loop structure that forms the A site in rRNA (*62*). The most common 1 × 3 loop, **A**(389)/**AAC**(372), occurs in a stem−loop region that is

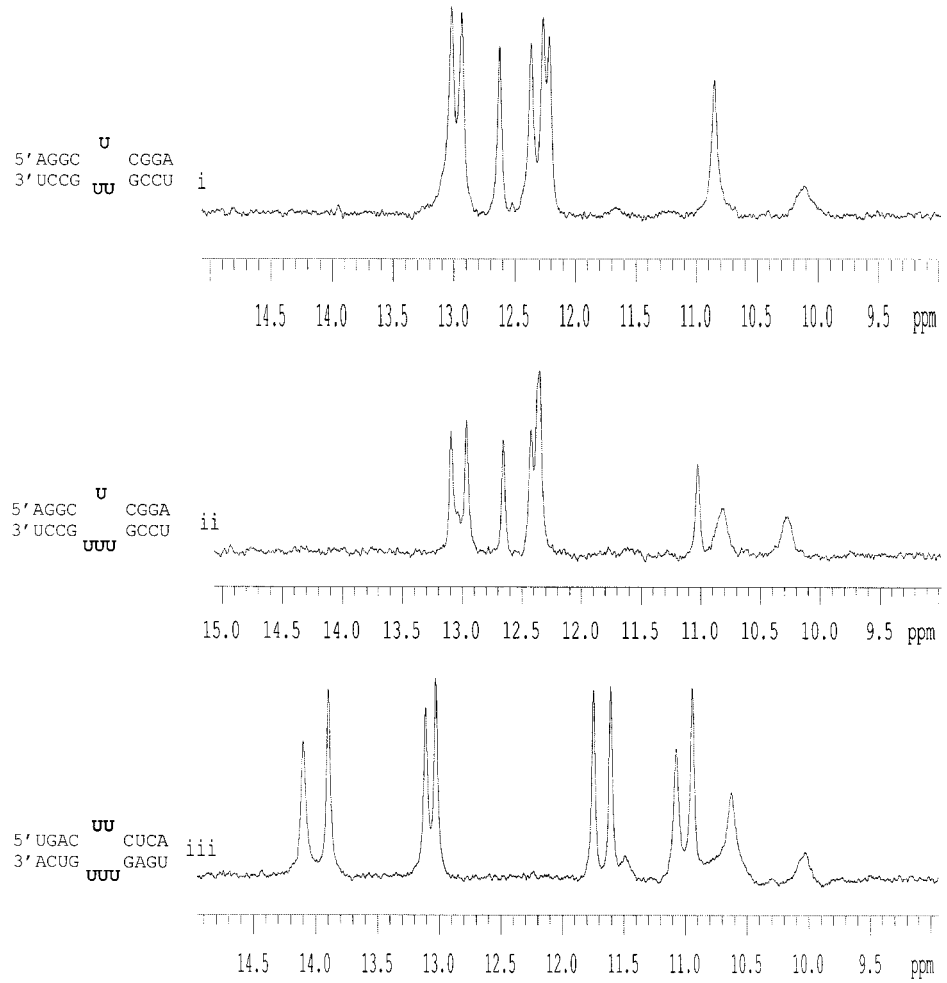

FIGURE 5: Imino proton NMR spectra (9−15 ppm) in 80 mM NaCl, 10 mM phosphate and 0.5 mM Na_2EDTA buffer, pH 6.7. (i) 5′AGGC**U**CGGA/3′UCCG**UU**GCCU, 0.5 mM, 5 °C, $\Delta G°_{loop\ 37C} = 1.6$ kcal/mol. (ii) 5′AGGC**U**CGGA/3′UCCG**UUU**GCCU, 0.3 mM, 5 °C, $\Delta G°_{loop\ 37C} = 2.8$ kcal/mol. (iii) 5′UGAC**UU**CUCA/3′ACUG**UUU**GAGU, 0.5 mM, 5 °C, $\Delta G°_{loop\ 37C} = 1.7$ kcal/mol.

protected when S16 protein binds in footprinting experiments (*63*). Both these loops are highly conserved but show no unusual thermodynamic stability. The proposed methods for predicting the thermodynamic stability of 1 × 3 and 2 × 3 internal loops work well for loops that have a known biological function or are conserved in phylogenetic databases and for loops that have no known biological role or occurrence.

DISCUSSION

Thermodynamic parameters for asymmetric internal loops are important for predicting RNA secondary structure, understanding RNA structure and function, and developing therapeutics that target RNA. The approximations developed here for predicting stabilities of 2 × 3 and 1 × *n* loops can be incorporated into programs for RNA secondary structure prediction (*1*, *64*, *65*). Secondary structure prediction methods such as *mfold* (*1*) can contribute to the interpretation of the rapidly increasing database of sequence information. Accurate prediction of secondary structure and accurate thermodynamic parameters for internal loops are also a step toward tertiary structure prediction and identification of loops that may form tertiary contacts. For example, the tetraloop receptor in *T. thermophila* group I introns is an asymmetric internal loop (*4*, *5*); identifying the asymmetric internal loops correctly in the secondary structure can contribute to identifying tetraloop−receptor interactions. The thermodynamic stability of internal loops may also be useful for identifying drug target sites. For example, a rational approach to drug design may prefer a preorganized site, such as 5′G**GA**C/3′C**AG**G, a very stable 2 × 2 loop with unique tandem imino GA mismatches (*55*). In contrast, aptamers selected for tight binding of small molecules and high specificity often contain highly asymmetric, flexible internal loops (*28*). Thus, the experimental identification of aptamer motifs could be coupled with secondary structure prediction to identify possible RNA targets for small molecules. Thermodynamic parameters for small loops and accurate secondary structure prediction of internal loops can thereby facilitate many investigations of RNA. The results presented here provide a basis for simple models that provide reasonable predictions of the stabilities for 2 × 3 and 1 × 3 internal loops (Table 3).

The GA Bonus for 2 × 3 Loops Depends on the Possible Stacking Arrangement of the Mismatch and Closing Base Pair. In 2 × 3 loops, the additional stability of a single GA mismatch depends on the orientation of the GA mismatch and its adjacent GC pair. The additional stability of a single GA mismatch in 2 × 3 loops relative to a loop with no stabilizing mismatches is −0.58, −0.87, and −1.31 kcal/mol for 5′CA/3′GG, 5′CG/3′GA, and 5′GG/3′CA, respectively. The possible mismatch 5′GA/3′CG does not provide

Table 5: Free Energy Increments of Loop Formation at 37 °C

2 × 3 loops	2 × 2 loops	1 × 3 loops	1 × 2 loops	1 × 1 loops	terminal mismatches
AA	AA	A	A	A	A
AAA	AA	AAA	AA	A	A
2.7	1.3[b]	2.3[b]	2.5[b]	(0.8)* [c]	−1.5[d]
UU	UU	U	U	U	U
UUU	UU	UUU	UU	U	U
1.8	−0.5* [b]	2.8	1.5[b]	0.0* [c]	−1.2[d]
GA	GA	G	G	G	G
AAA	AA	AAA	AA	A	A
1.9	−0.2* [b]	1.6	1.7b	−0.1* [c]	−1.4[d]
AA	AA	A	A	A	A
AAG	AG	AAG	AG	G	G
1.5	−0.3* [b]	(2.15)	0.8[b]	(−0.1)* [c]	−1.4[d]
GG	GG				
AAA	AA				
2.0	−2.0* [b]				
GA	GA				
AAG	AG				
0.2	−1.7* [b]				

[a] The free energy of loop formation is calculated by the following equation (*17*): $\Delta G^\circ_{loop} = \Delta G^\circ_{duplex\ with\ loop} - (\Delta G^\circ_{duplex\ without\ loop} - \Delta G^\circ_{interrupted\ base\ pair})$. The thermodynamic parameters required for this equation were derived from optical melting experiments in 1 M NaCl, 10 or 20 mM sodium cacodylate, and 0.5 mM EDTA buffer, pH 7. All the loops have two CG closing base pairs with orientations like 5′C**A**C/3′G**AA**G, except for values in parentheses, where the orientations are different. An asterisk indicates that the loop was bounded by helices of three base pairs; all other loops were bounded by helices of four base pairs. When multiple measurements of the same loop were available, the value listed is the average of the measurements. For example, the value 1.3 kcal/mol for the loop 5′CAAC/3′GAAG is the average of 0.5 and 2.04 kcal/mol (*1*, *19*, *25*). [b] Reference *1*. [c] Reference *57*. [d] Reference *70*.

any apparent additional stability. Interestingly, in symmetric 2 × 2 loops with GA mismatches, the most stable motif also contains 5′GG/3′CA; the second most stable motif contains 5′GA/3′CG, however (*1, 24*).

The peaks for GA mismatches in imino proton spectra of 2 × 3 loops correlate with the thermodynamic parameters for GA mismatches in 2 × 3 loops. Some spectra in Figure 2 show small peaks between 9.8 and 10.5 ppm, which is a typical resonance frequency for sheared GA mismatches (*24*); these loops also contain the two most stable GA stacking arrangements, 5′CG/3′GA and 5′GG/3′CA. Spectra iii and vi in Figure 2 do not show extra peaks for imino protons within the loop. The absence of a peak in the imino proton spectrum does not necessarily imply that a hydrogen-bonded mismatch does not form; the imino protons in the mismatch can exchange with water if the structure is partly flexible. The loops for which no imino proton resonances are detected, however, have the least stabilizing GA stacking arrangements, 5′CA/3′GG and 5′GA/3′CG.

UU and GG Mismatches Also Stabilize 2 × 3 Loops. The average additional stabilities of UU and GG mismatches relative to a loop with no stabilizing mismatches in three 2 × 3 loops are −0.32 ± 0.14 and −0.70 ± 0.24 kcal/mol, respectively. Given the few molecules studied and the small range of stabilities, an average is the best approximation for the UU and GG mismatch bonuses at this time. For UU mismatches, the spectra in Figure 5 show small peaks between 10 and 11.5 ppm, suggesting some structure in all the loops. The NOEs in Figure 4 spectra viii and ix suggest that 5′C**UU**C/3′G**UUU**G has stable UU stacking on a helix end and interactions between the U nucleotides in the loop.

AU/GU Penalty Derived from 2 × 2 Loops Applies to 2 × 3 Loops. On the basis of stabilities of symmetric 2 × 2 loops, the penalty for AU or GU closing base pairs is 0.2 kcal/mol in addition to the 0.45 kcal/mol penalty for terminating a helix with an AU base pair (*1, 54*). With this penalty, stabilities for the loops 5′U**GC**G/3′A**AAA**C, 5′U**GC**U/3′A**ACA**A, and 5′U**GG**A/3′G**AGA**U are all predicted within experimental error by eq 3 (Table 2). When a possible GA mismatch could form adjacent to the AU closing base pairs, the prediction uses an average value of the three stabilizing GA bonuses for 2 × 3 loops, −0.9 kcal/mol. The success of predictions for three loops with AU closing base pairs suggests that the AU/GU penalty derived from studies of 2 × 2 loops is applicable to 2 × 3 loops.

Structures of 2 × 3 Loops May Be Partly Flexible and May Involve the Nucleotide Not Adjacent to the Closing Base Pairs. In the imino proton spectra of 2 × 3 internal loops in Figure 2, the peaks assigned to nucleotides within the loop are small and broad, which suggests some flexibility. In many NMR studies of asymmetric internal loops, the loop was flexible until bound to its substrate (*8, 10, 11, 13, 34, 36*). Although it is possible that a flexible loop may form transient mismatches not adjacent to the helix ends, this does not seem to add significant thermodynamic stability to the loop. For example, the loops 5′G**AA**G/3′C**CGA**C and 5′C**UU**C/3′G**CUC**G, which could possibly form GA or UU mismatches involving the middle nucleotide of CGA and CUC, respectively, have similar ΔG°_{loop} values as the loop 5′G**AA**G/3′C**AAA**C. This may be explained by the disruption of stacking between the possible GA or UU mismatch and the helix. The imino proton spectrum of 5′C**UU**C/3′G**CUC**G in Figure 3 (ii), however, does suggest that the loop has structure. Also, the loop 5′C**GG**C/G**AGA**G is 0.6 kcal/mol more stable than 5′C**GG**C/G**AAA**G. More detailed structural studies may reveal the subtle role of the middle nucleotide in the loop structure and stability.

The structures of 2 × 3 internal loops solved by NMR and X-ray crystallography are consistent with the observation that certain mismatches are thermodynamically stabilizing. In the NMR structures of the bound and free Rev-responsive element 2 × 3 internal loop in HIV-1 RNA, GA and GG mismatches form adjacent to the helix ends (*10, 11*), which is consistent with the model used for eq 3. The NMR structure of an aptamer that binds the Rev peptide also contains a 2 × 3 loop with an imino hydrogen-bonded GA mismatch (*30*). Also, in the NMR structure of the J6/6a tetraloop receptor of the *T. thermophila* group I intron, a UU mismatch forms adjacent to the helix ends in a 2 × 3 loop with a GU closing base pair, which is also consistent with the model of eq 3 (*34*). When the J6/6a loop is bound to a hairpin in the X-ray crystal structure, however, the loop has a different structure with an AA platform and a reverse Hoogsteen AU pair in a base triple (*4, 5*). The J4/5 loop in the *T. thermophila* group I intron, a 2 × 3 loop that is bound to a helix in the minor groove, contains all A nucleotides and forms two stacked purine−purine mismatches (*4, 5*). Further investigation on the structure and stability of 2 × 3 internal loops is needed to understand the perturbations of loops forming tertiary contacts.

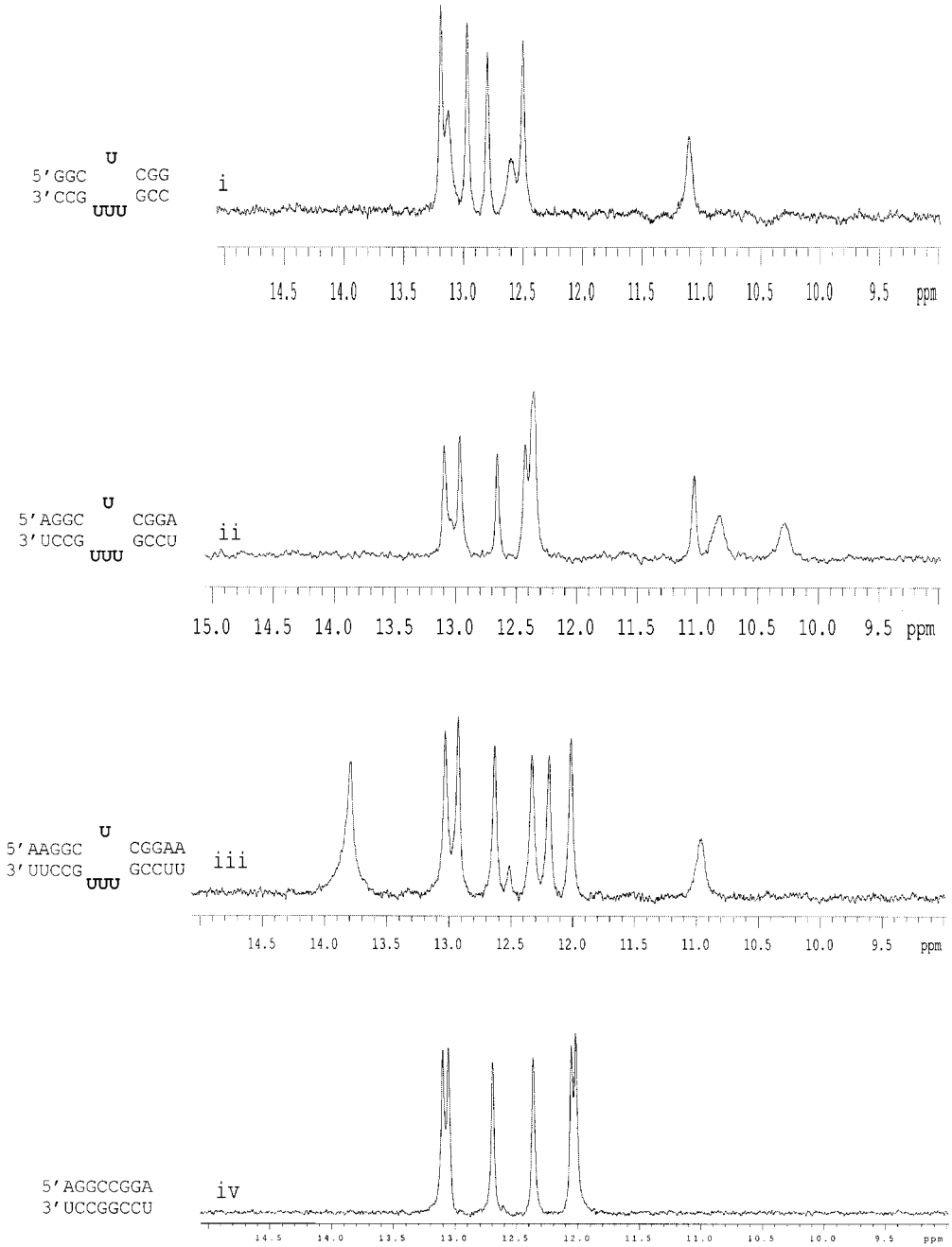

FIGURE 6: Imino proton NMR spectra (9−15 ppm) in 80 mM NaCl, 10 mM phosphate, and 0.5 mM Na_2EDTA buffer, pH 6.7. Spectra are listed in order of increasing loop free energy. The single spectrum shown from a variable-temperature study was selected for the best dispersion. (i) 5′GGC**U**CGG/3′CCG**UUU**GCC, 0.2 mM, 5 °C, $\Delta G^\circ_{\text{loop 37C}} = 2.2$ kcal/mol. (ii) 5′AGGC**U**CGGA/3′UCCG**UUU**GCCU, 0.3 mM, 5 °C, $\Delta G^\circ_{\text{loop 37C}} = 2.8$ kcal/mol. (iii) 5′AAGGC**U**CGGAA/3′UUCCG**UUU**GCCUU, 0.3 mM, 5 °C, $\Delta G^\circ_{\text{loop 37C}} = 3.3$ kcal/mol. (iv) 5′AGGCCGGA/3′UCCGGCCU, 0.5 mM, 10 °C.

Average Values That Depend on the Closing Base Pairs Predict the Thermodynamic Stabilities of 1 × 3 Internal Loops. The average $\Delta G^\circ_{\text{loop}}$ for 18 1 × 3 loops with two CG closing base pairs is 2.36 ± 0.49 kcal/mol; the average $\Delta G^\circ_{\text{loop}}$ for nine 1 × 3 loops with two AU closing base pairs is 3.89 ± 0.39 kcal/mol. The most stable 1 × 3 loop with CG closing base pairs is 0.79 kcal/mol, or 0.2 kcal/mol per nucleotide, more stable than the average. Similarly, the least stable 1 × 3 loop with two AU closing base pairs is 0.9 kcal/mol, or 0.2 kcal/mol per nucleotide, less stable than the average. Although the averages do not take into account subtleties in the sequence dependence, the number of possible 1 × 3 internal loops prohibits a comprehensive survey. Thus, a reasonable approximation for all 1 × 3 internal loops is to use $\Delta G^\circ_{\text{loop}}$ of 2.36, 3.12, and 3.89 kcal/mol for loops closed with two, one, or zero GC pairs, respectively.

The imino proton spectra of 1 × 3 loops containing possible GA mismatches are consistent with a thermodynamic model that neglects possible mismatches in the loop because the spectra show no resonances that can be assigned to a GA mismatch. The imino proton spectrum for a 1 × 3 loop containing all U nucleotides, however, shows resonances with chemical shifts typical of UU mismatches (Figure 5 spectrum ii). Nevertheless, no NOEs were observed between these resonances in the one-dimensional NOE difference

Table 6: Natural Occurrence of 2 × 3 and 1 × 3 Internal Loops in 23S rRNA, 16S rRNA, and Group I Introns[a]

2X3 Loops		1X3 Loops	
*CC AAU	89	*A AAC	91
AA CAC	35	U UGA	81
CA AUA	27	*A AAG	38
UA UAG	25	A AUG	26
AU AAU	23	A CAC	21
AU CAU	21	U AAA	18
AU CGU	21	C AAC	17
AC GAU	17	G AAC	17
AU GGU	15	C CUC	17
CC UAA	12	*A CGA	17
*GC AAA	11	A GUG	15
GA AUA	10	G AAG	14
CC GAU	9	A UUA	12
AG CGU	9	A CUC	12
GA ACG	7	C UGU	9
AA GAA	7	A UGU	8
GU AAU	7	C CCC	8
CA AAG	7	C CGC	8
		G CUC	7
		U CUG	7

[a] Occurrence in 218 large subunit rRNA, 101 small subunit rRNA, and 79 group I introns (*61, 62, 87*). A total of 743 2 × 3 loops and 689 1 × 3 loops were counted. Loops that occur less than 1% are not listed. Only 17% and 50% of the possible loop sequences for 2 × 3 and 1 × 3 loops, respectively, occur in this database. The sequences of 2 × 3 and 1 × 3 loops are listed as 5′XX3′/3′XXX5′ and 5′X3′/3′XXX5′, respectively. Sequences marked with an asterisk are included in the thermodynamic results reported in Table 2.

spectra; such NOEs are expected for a hydrogen-bonded UU mismatch (*22*). The degree of asymmetry in 1 × 3 loops and the resulting backbone distortion may perturb the stacking interactions at the helix ends and add an unfavorable free energy component, thus diminishing the net stabilizing effects of forming a mismatch. The flexibility of 1 × 3 internal loops also suggests that an average is a good approximation for predicting the thermodynamic stability.

Four structures of 1 × 3 internal loops solved by NMR and X-ray crystallography, however, do show a mismatch adjacent to the helix ends (*29, 39, 40*). The citrulline and arginine aptamers each contain a 1 × 3 loop that forms a GA mismatch adjacent to the helix. Both crystal structures of the 23S rRNA with bound L11 protein contain a 1 × 3 loop that forms either a UU or non-Watson−Crick UA mismatch adjacent to a helix end. In each of these cases, however, the loop is part of a complex tertiary structure. It is possible that 1 × 3 loops are inherently flexible but can form a fixed structure when bound to a substrate or tertiary contact.

An Account of Hydrogen Bonds May Explain the Difference in Stabilities between 1 × 3 Loops with CG or AU-Closing Base Pairs. The difference between the average values for 1 × 3 loops with CG and AU closing base pairs, 1.5 kcal/mol, is consistent with previous studies of the effects of different closing base pairs in 2 × 2 loops (*23*) and terminal AU base pairs in helices (*54*). The penalty for AU closing base pairs in 2 × 2 internal loops is 0.2 kcal/mol in addition to the 0.45 kcal/mol penalty for terminal AU base pairs (*1*). The total penalty for changing two CG closing base pairs to two AU base pairs is then 1.3 kcal/mol, which is within experimental error of the 1.5 kcal/mol average difference observed with 1 × 3 loops.

The penalty for terminating a helix with an AU base pair accounts for the different number of hydrogen bonds in CG and AU base pairs (*54*). In the nearest-neighbor model, half of the free energy contribution from hydrogen bonds in a closing base pair is included in the nearest-neighbor term for the closing base pair stacking on its preceding base pair. The other half of the hydrogen-bonding free energy contribution is included in the ΔG°_{loop} term. Each AU base pair has one less hydrogen bond than a GC base pair. Thus, when the free energy contribution from hydrogen bonds in the closing base pairs is tallied, the ΔG°_{loop} term for a loop with two AU closing base pairs includes one (or two halves) less hydrogen bond than the ΔG°_{loop} term for a loop with two GC closing base pairs. The difference in the average values for 1 × 3 loops with two CG or two AU closing base pairs, 1.5 kcal/mol, is reasonably close to previous estimates of the value of a hydrogen bond in a GC pair (*52, 68*).

Length of the Surrounding Helices Affects Loop Stability. Three measurements of the thermodynamics of forming a 1 × 3 loop containing only U nucleotides show that the loop becomes less stable by about 0.6 kcal/mol at 37 °C when a base pair is added to each end of the helix (Table 4). This is consistent with earlier experiments on 2 × 2 loops containing GA and AA mismatches (*21, 22*). The stability of 1 × 1 loops, or single mismatches, can also depend strongly on the mismatch position in the helix; the closer to the end of the helix, the more stable is an AA or UU mismatch (*57*). Similar non-nearest-neighbor effects have been noted in previous thermodynamic studies of bulges (*69*). For example, thermodynamic studies of bulge loops in RNA showed non-nearest-neighbor effects greater than 1 kcal/mol that depend on nonadjacent base pairs in the helix and 3′ dangling ends (*69*). Because various types of loops are consistently less stable when the flanking helices are lengthened, there are evidently non-nearest-neighbor effects on loop stability. In Figure 6, the imino proton spectra of the three duplexes containing 1 × 3 loops of U nucleotides show differences in the number of peaks assigned to U nucleotides in the loop and the chemical shifts of peaks assigned to the stem helices. These spectra suggest that structural changes may contribute to the changes in ΔG°_{loop}.

The calculation of ΔG°_{loop} uses a nearest neighbor approximation for the helices surrounding a loop and assumes that the internal loop interrupts only the stacking interaction of one base pair (eq 2). Internal loops can affect the structure of the surrounding helices beyond the closing base pair of the loop, however. For example, asymmetric internal loops show an increase in the accessibility to chemical modification of the major groove in the surrounding helix (*20*); the increased major groove accessibility extends four and two nucleotides in the 3′ and 5′ directions, respectively. Bulged nucleotides may also interact with the phosphate backbone or nearby helix. For example, in an RNA aptamer that binds HIV-1 Rev protein, a U nucleotide in a bulge forms a base triple with an AU base pair in the helix (*30*). The ΔG°_{loop} term for loops with the same length helix but different helix sequences varies by an average of 0.4 kcal/mol. These non-nearest-neighbor interactions depend on the type of loop and the sequence of the surrounding helix. The effects of these non-nearest-neighbor interactions, however, are difficult to predict and to quantitate and are probably small compared to the interruption of the base pair stacking. Nevertheless, comparisons of the ΔG°_{loop} for loops with different helices are valid within approximately 0.5 kcal/mol.

Stabilizing Effects of GA and UU Mismatches Vary with the Type of Loop. Table 5 compares the ΔG°_{loop} at 37 °C of internal loops containing possible AA, UU, and GA mismatches in different loops. The last two rows include examples of loops with two possible GA mismatches; the columns for $1 \times n$ loops are left empty in these rows because only one GA mismatch can be formed in a $1 \times n$ loop. The stabilizing effect of UU and GA mismatches relative to AA mismatches varies depending on the size and asymmetry of the loop. For example, an internal loop of four U nucleotides as a symmetric 2×2 loop has a ΔG°_{loop} of -0.5 kcal/mol, which is 1.8 kcal/mol more stable than a 2×2 loop of all A nucleotides. In contrast, a loop of four U nucleotides as an asymmetric 1×3 loop has a ΔG°_{loop} of 2.8 kcal/mol, which is approximately the same as the value of 2.3 kcal/mol for an 1×3 internal loop containing all A residues. The difference between 1×3 and 2×2 loops with four U nucleotides is 3.3 kcal/mol, while for loops with four A nucleotides, the difference is only 1.0 kcal/mol. In Figures 5 and 6, the imino proton spectra of loops containing U nucleotides show small peaks that can be assigned to the U imino protons in the loop, suggesting that the loops have some structure. Although the loops also may be partly flexible, changes in asymmetry of the loop may affect the degree of backbone distortion and the ability of the duplex to accommodate the mismatched nucleotides in a stable structure. Thus, the thermodynamic stability of internal loops depends on a balance of many factors, including loop size, asymmetry, loop sequence, possible stabilizing mismatches, and the surrounding helices.

Method for Predicting the Thermodynamic Stabilities of Internal Loops Depends on the Type of Loop. The rules for internal loop stability are clearly idiosyncratic. The 2×2 loops show the largest range of thermodynamic stability, 4.8 kcal/mol or greater than 1000-fold in a binding constant, for loops with CG closing base pairs (*1, 25, 23*). In contrast, 1×3 loops with CG closing base pairs and helices of the same length have an overall range of only 1.5 kcal/mol, approximately 10-fold in a binding constant (Table 2). The 1×2 and 2×3 loops have a range of stabilities of 1.7 and 2.6 kcal/mol, respectively. Comparison of the stability of terminal mismatches (*53*) versus internal mismatches (*57*) shows that the effect of constraining a mismatch on both sides by a helix has a destabilizing effect greater than 1 kcal/mol (Table 5). The differences in the ranges of stabilities for different loops may result from the effect of backbone distortion, different degrees of constraint on the structure of the loop, and the ability of the loops to accommodate a stabilizing mismatch. Thus, different approximations are necessary for predicting the stabilities of different types of internal loops.

The rules for approximating stabilities of symmetric 2×2 loops are summarized in a periodic table of symmetric tandem mismatches (*23*). The penalty for AU closing base pairs, 0.2 kcal/mol in addition to the 0.45 kcal/mol penalty for AU base pairs terminating a helix, was derived from studies of symmetric 2×2 loops and also applies to other loops. The following equation has provided reasonable approximations for nonsymmetric tandem mismatches:

$$\Delta G^\circ_{loop}(5'\mathrm{P}\mathbf{XY}\mathrm{S}/3'\mathrm{Q}\mathbf{WZ}\mathrm{T}) = {}^1/_2[\Delta G^\circ_{loop}(5'\mathrm{P}\mathbf{XW}\mathrm{Q}/3'\mathrm{Q}\mathbf{WX}\mathrm{P}) + \Delta G^\circ_{loop}(5'\mathrm{T}\mathbf{ZY}\mathrm{S}/3'\mathrm{S}\mathbf{YZ}\mathrm{T})] + \Delta_P \quad (4)$$

where PQ and ST are Watson−Crick base pairs and where Δ_P is a penalty for mismatches of different sizes (*1, 25*).

Equation 3 and the parameters in Table 3 predict well the thermodynamic stabilities of 2×3 loops. The RNA folding algorithm *mfold* also uses eq 3 to predict larger $m \times n$ internal loops when both m and n are greater than 1 (*1*). Both the $\Delta G^\circ_{loop\ initiation}$ and the $\Delta G^\circ_{asymmetry}$ terms are derived from data for a series of loops of different sizes and asymmetry. The penalty for AU and GU closing base pairs was derived from studies of symmetric 2×2 loops (*23*) and is similar to that measured for 1×3 loops (Tables 2 and 3). On the basis of results in Table 2, the bonus for GA and UU mismatches, however, should depend on the loop size and asymmetry. The loop asymmetry may cause distortions in the phosphate backbone that restrain the structure of the loop and prevent certain stacking interactions. Loops with more nucleotides may have more possible ways to accommodate both backbone distortion and stabilizing mismatches. Thus, the stabilizing effects of GA and UU mismatches depends on the loop size and asymmetry. Comparisons of the free energies for small loops in Table 5 demonstrates this dependence. Although the current bonus of -1.1 and -0.7 kcal/mol for GA and UU mismatches, respectively, may be an overestimate of loop stability, further investigation of larger loops is necessary to approximate better the stabilizing effects of GA and UU mismatches in larger loops. The results in Tables 2 and 3 also suggest that bonuses should be applied to some loops that begin with a GG mismatch.

NMR structures of larger internal loops show that at least one mismatch forms adjacent to the helix end in most loops, although the mismatch may not always be predicted by eq 3. For example, two GA mismatches form adjacent to the helix ends in the FMN aptamer (*29*), and the structure is consistent with the model of eq 3. In the leadzyme NMR structure of the loop 5′G**GA**G/3′C**GAGC**C, however, a

protonated CA mismatch forms adjacent to the helix end but not a possible GG mismatch (*35*). Because it remains difficult to predict which mismatches will occur within a loop, counting the possible stabilizing mismatches adjacent to the helix ends remains a reasonable first approximation of thermodynamic stability, although this may not always reflect the structure of the loops.

The method of predicting thermodynamic stability for 1 × *n* loops depends on *n*. A recent study of 1 × 1 loops, or single mismatches, has shown that an average value of 0.3 kcal/mol for all mismatches other than the unusually stable GG (−2.2 kcal/mol) and 5′G**U**C/3′C**U**G (−0.6 kcal/mol) motifs predicts well the stability of 1 × 1 loops closed by CG base pairs (*57*). A penalty of 0.8 kcal/mol is added per AU closing base pair for 1 × 1 loops. A table listing the experimental and estimated values for all 1 × 2 loops with CG closing base pairs, including the corrections reported in this paper, is used to predict the stability of 1 × 2 loops (*1*, *26*). The most stable 1 × 2 loops have potential GA and UU mismatches. In contrast, 1 × 3 loops do not consistently show an increase in stability when the loop contains a potential GA or UU mismatch. The stability of 1 × 3 loops depends most strongly on the closing base pairs.

The average approximation approach for 1 × 3 loops has been extended to all 1 × *n* loops where *n* is greater than 2 because no data are available for these 1 × *n* loops with possible GA, GG, or UU mismatches. The values from the study of asymmetric internal loops containing all A nucleotides are used to approximate the stability of larger 1 × *n* loops (*1*, *19*). This approach is equivalent to using eq 3 with no GA, GG, or UU bonuses for 1 × *n* loops. This approximation assumes that loops with large asymmetry are probably flexible and that loop stabilities therefore do not depend strongly on the sequence in the loop, as observed for 1 × 3 loops. The NMR structures of 1 × *n* loops in several aptamers and in the U1A protein binding site of a 3′ untranslated region are all flexible until bound to a substrate, which supports this hypothesis (*8*, *27*, *28*). Further investigation of the stability of asymmetric internal loops may clarify the thermodynamics of internal loops and how the stability changes when the loop binds a substrate or forms a tertiary contact.

ACKNOWLEDGMENT

We thank D. H. Mathews for writing the computer program to analyze competition between self-complementary and non-self-complementary duplexes and assistance with the linear regression analysis. The authors also thank Professor Christopher Cox for advice and assistance with the linear regression analysis. We also thank M. E. Burkard for writing the program to search phylogenetic databases.

SUPPORTING INFORMATION AVAILABLE

Tables of thermodynamic parameters for self-complementary single-strand oligomers, the linear regression analysis, a sample graph of the analysis of competition between self-complementary and non-self-complementary duplexes, sample imino proton spectra for the analysis of competition between non-self-complementary and self-complementary duplexes, imino proton spectra for duplexes containing 1 × 2 and 1 × 3 loops, sample melting curves, and sample $1/T_M$ vs ln (C_T/4) plots. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) *J. Mol. Biol. 288*, 911−940.
2. Xia, T., Mathews, D. H., and Turner, D. H. (1999) in *Preobiotic Chemistry, Molecular Fossils, Nucleosides, and RNA* (Soll, D. G., Nishimura, S., and Moore, P. B., Eds.) pp 21−48, Elsevier Press, Oxford, England.
3. Szewczak, A. A., and Cech, T. R. (1997) *RNA 3*, 838−849.
4. 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.
5. 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.
6. Strobel, S. A., Ortoleva-Donnelly, L., Ryder, S. P., Cate, J. H., and Moncoeur, E. (1998) *Nat. Struct. Biol. 5*, 60−66.
7. Strobel, S. A., and Ortoleva-Donnely, L. (1999) *Chem. Biol. 5*, 153−165.
8. Gubser, C. C., and Varani, G. (1996) *Biochemistry 35*, 2253−2267.
9. Kalurachchi, K., and Nikonowicz, E. P. (1998) *J. Mol. Biol. 280*, 639−654.
10. Battiste, J. L., Mao, H., Rao, N. S., Tan, R., Munandiram, D. R., Kay, L. E., Frankel, A. D., and Williamson, J. R. (1996) *Science 273*, 1547−1551.
11. Peterson, R. D., and Feigon, J. (1996) *J. Mol. Biol. 264*, 863−877.
12. Fourmy, D., Recht, M. I., and Puglisi, J. D. (1998) *J. Mol. Biol. 277*, 347−362.
13. Fourmy, D., Yoshizawa, S., and Puglisi, J. D. (1998) *J. Mol. Biol. 277*, 333−345.
14. Woodcock, J., Moazed, D., Cannon, M., Davies, J., and Noller, H. F. (1991) *EMBO J. 10*, 3099−3103.
15. Moazed, D., and Noller, H. F. (1987) *Nature 327*, 389−394.
16. Davies, J., and Davis, B. D. (1968) *J. Biol. Chem. 243*, 3312−3316.
17. Gralla, J., and Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301−319.
18. Papanicolaou, C., Guoy, M., and Ninio, J. (1984) *Nucleic Acids Res. 12*, 31−43.
19. Peritz, A. E., Kierzek, R., Sugimoto, N., and Turner, D. H. (1991) *Biochemistry 30*, 6428−6436.
20. Weeks, K. M., and Crothers, D. M. (1993) *Science 261*, 1574−1577.
21. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1991) *J. Am. Chem. Soc. 113*, 4313−4322.
22. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1991) *Biochemistry 30*, 8242−8251.
23. Wu, M., McDowell, J. A., and Turner, D. H. (1995) *Biochemistry 34*, 3204−3211.
24. Walter, A. E., Wu, M., and Turner, D. H. (1994) *Biochemistry 33*, 11349−11354.
25. Xia, T., McDowell, J. A., and Turner, D. H. (1997) *Biochemistry 36*, 12486−12497.
26. Schroeder, S., Kim, J., and Turner, D. H. (1996) *Biochemistry 35*, 16105−16109.
27. Zimmerman, G. R., Jenison, R. D., Wick, C. L., Simorre, J.-P., and Pardi, A. (1997) *Nat. Struct. Biol. 4*, 644−649.
28. Feigon, J., Dieckmann, T., and Smith, F. W. (1996) *Chem. Biol. 3*, 611−617.
29. Yang, Y., Kochoyan, M., Burgstaller, P., Westhof, E., and Famulok, M. (1996) *Science 272*, 1343−1346.
30. Ye, X., Gorin, A., Ellington, A. D., and Patel, D. J. (1996) *Nat. Struct. Biol. 3*, 1026−1033.
31. Szewczak, A. A., Moore, P. B., Chan, Y.-L., and Wool, I. G. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 9581−9585.
32. Wimberly, B., Varani, G., and Tinoco, I., Jr. (1993) *Biochemistry 32*, 1078−1087.
33. Butcher, S. E., Allain, F. H.-T., and Feigon, J. (1999) *Nat. Struct. Biol. 6*, 212−216.

34. Butcher, S. E., Dieckmann, T., and Feigon, J. (1997) *EMBO J. 16*, 7490−7499.
35. Legault, P., Hoogstraten, C. G., Metlitzky, E., and Pardi, A. (1998) *J. Mol. Biol. 284*, 325−335.
36. Fourmy, D., Recht, M. I., Blanchard, S. C., and Puglisi, J. D. (1996) *Science 274*, 1367−1371.
37. Gnadiec, Z., Sierzputowska-Gracz, H., and Theil, E. C. (1998) *Biochemistry 37*, 1505−1512.
38. Glemarec, C., Kufel, J., Foldesi, A., Maltseva, T., Sandstrom, A., Kirsebom, L. A., and Chattopadhyaya, J. (1996) *Nucleic Acids Res. 24*, 2022−2035.
39. Wimberly, B. T., Guymon, R., McCutcheon, J. P., White, S. W., and Ramakrishnan, V. (1999) *Cell 97*, 491−502.
40. Conn, G. L., Draper, D. E., Lattman, E. E., and Gittis, A. G. (1999) *Science 284*, 1171−1174.
41. Wedekind, J. E., and McKay, D. B. (1999) *Nat. Struct. Biol. 6*, 261−268.
42. Correll, C. C., Freeborn, B., Moore, P. B., and Steitz, T. A. (1997) *Cell 91*, 705−712.
43. Correll, C. C., Mushikin, A., Chan, Y., Zhong, R., Wool, I. G., and Steitz, T. A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 13436−13441.
44. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. J. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.
45. Scaringe, S. A., Franklyn, C., and Usman, N. (1990) *Nucleic Acids Res. 18*, 5433−5441.
46. Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology*: *Nucleic acids* (Fasman, G. D., Ed.) p 5897, CRC Press, Cleveland, OH.
47. Petersheim, M., and Turner, D. H. (1983) *Biochemistry 22*, 256−263.
48. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843−853.
49. McDowell, J. A., and Turner, D. H. (1996) *Biochemistry 35*, 14077−14089.
50. Hore, P. J. (1983) *J. Magn. Reson. 55*, 283−300.
51. Freier, S. M., Kierzek, R., Jaeger, J. A., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.
52. 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.
53. Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.
54. 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.
55. Wu, M., and Turner, D. H. (1996) *Biochemistry 35*, 9677−9689.
56. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1990) *Biochemistry 29*, 8813−8819.
57. Kierzek, R., Burkard, M. E., and Turner, D. H. (1999) *Biochemistry 38*, 14214−14223.
58. Wijmenga, S. S., Heus, H. A., Werten, B., van der Marel, G. A., van Boom, J. H., and Hilbers, C. W. (1994) *J. Magn. Reson., Ser. B 103*, 134−141.
59. Tinoco, I., Jr., and Bustamante, C. (1999) *J. Mol. Biol. 293*, 271−281.
60. Nikonowicz, E., and Pardi, A. (1992) *Nature 327*, 389−394.
61. Gutell, R. R., Gray, M. W., and Schnare, M. N. (1993) *Nucleic Acids Res. 21*, 3055−3074.
62. Gutell, R. R. (1994) *Nucleic Acids Res. 22*, 3502−3507.
63. Powers, T., and Noller, H. F. (1995) *RNA 1*, 194−209.
64. Gultyaev, A., van Batenburg, F., and Pleij, C. (1995) *J. Mol. Biol. 250*, 37−51.
65. Lück, R., Steger, G., and Reisner, D. (1996) *J. Mol. Biol. 258*, 813−826.
66. SantaLucia, J., Jr., and Turner, D. H. (1993) *Biochemistry 32*, 12612−12623.
67. Gautheret, D., Konings, D., and Gutell, R. R. (1994) *J. Mol. Biol. 242*, 1−8.
68. Turner, D. H., Sugimoto, N., Kierzek, R., and Dreiker, S. D. (1987) *J. Am. Chem. Soc. 109*, 3783−7385.
69. Longfellow, C. E., Kierzek, R., and Turner, D. H. (1990) *Biochemistry 29*, 278−285.
70. Serra, M. J., and Turner, D. H. (1995) *Methods Enzymol. 259*, 242−261.
71. Baumstark, T., and Reisner, D. (1995) *Nucleic Acids Res. 23*, 4246−4254.
72. Jang, S. K., and Wimmer, E. (1990) *Genes Dev. 4*, 1560−1572.
73. Fodor, E., Pritlove, D. C., and Brownlee, G. G. (1995) *J. Virol. 69*, 4012−4019.
74. Mathews, D. H., Banerjee, A. R., Luan, D. D., Eickbush, T. H., and Turner, D. H. (1997) *RNA 3*, 1−16.
75. Cascone, P. J., Haydar, T. F., and Simon, A. E. (1993) *Science 260*, 801−804.
76. Hsue, B., and Masters, P. S. (1997) *J. Virol. 71*, 7567−7578.
77. Gora-Sochacka, A., Kierzek, A., Candresse, T., and Zagorski, W. (1997) *RNA 3*, 68−74.
78. Gutierrez, A. L., Denova-Ocampo, M., Racaniello, V. R., and del Angel, R. M. (1997) *J. Virol. 71*, 3826−3833.
79. Hofmann, H., Limmer, S., Hornung, V., and Sprinzl, M. (1997) *RNA 3*, 1289−1300.
80. McCammon, J. A., and Harvey, S. C. (1987) *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge, England.
81. Shen, L. X., Cai, Z., and Tinoco, I., Jr. (1995) *Fed. Am. Soc. Exp. Biol. 9*, 1023−1033.
82. Szewczak, A. A., Ortoleva-Donnelly, L., Ryder, S. P., Moncoeur, E., and Strobel, S. A. (1998) *Nat. Struct. Biol. 5*, 1037−1042.
83. Wijmenga, S. S., Mooren, M. M., and Hilbers, C. W. (1993) In *NMR of Macromolecules, A Practical Approach* (Roberts, G. C. K., Ed.) pp 217−288, Oxford University Press, New York.
84. Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A., and Weiner, A. M. (1987) *Molecular Biology of the Gene*, 4th ed., Benjamin/Cummings Publishing Co., Menlo Park, CA.
85. Wyatt, J. R., and Tinoco, I., Jr. (1993) in *The RNA World* (Gesteland, R. F., and Atkins, J. F., Eds.) pp 465−496, Cold Spring Harbor Press, Plainview, NY.
86. 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.) 2nd edition, Appendix I, pp 675−680, Cold Spring Harbor Press, Plainview, NY.
87. Damberger, S. H., and Gutell, R. R. (1994) *Nucleic Acids Res. 22*, 3508−3510.

BI000229R