Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for the Formation of RNA Duplexes with Single Nucleotide Bulges[†]

Brent M. Znosko, Sara B. Silvestri, Heather Volkman, Bob Boswell, and Martin J. Serra*

Department of Chemistry, Allegheny College, Meadville, Pennsylvania 16335-3902

Received March 6, 2002; Revised Manuscript Received May 22, 2002

ABSTRACT: Thirty-four RNA duplexes containing single nucleotide bulges were optically melted, and the thermodynamic parameters $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$, and T_M for each sequence were determined. Data from this study were combined with data from previous thermodynamic data [Longfellow, C. E., Kierzek, R., and Turner, D. H. (1990) *Biochemistry 29,* 278−85] to develop a model that will more accurately predict the free energy of an RNA duplex containing a single nucleotide bulge. Differences between purine and pyrimidine bulges as well as differences between Group I duplexes, those in which the bulge is not identical to either neighboring nucleotide, and Group II duplexes, those in which the bulge is identical to at least one neighboring nucleotide, were considered. The length of the duplex, non-nearest-neighbor effects, and bulge location were also examined. A model was developed which divides sequences into two groups: those with pyrimidine bulges and those with purine bulges. The proposed model for pyrimidine bulges predicts $\Delta G°_{37,bulge} = 3.9$ kcal/mol $+ 0.10\Delta G°_{37,nn} + \beta$, while the model for purine bulges predicts $\Delta G°_{37,bulge} = 3.3$ kcal/mol $- 0.30\Delta G°_{37,nn} + \beta$, where β has a value of 0.0 and -0.8 kcal/mol for Group I and Group II sequences, respectively, and $\Delta G°_{37,nn}$ is the nearest-neighbor free energy of the base pairs surrounding the bulge. The conformation of bulge loops present in rRNA was examined. Three distinct families of structures were identified. The bulge loop was either extrahelical, intercalated, or in a "side-step" conformation.

A nearest-neighbor model consisting of thermodynamic parameters for various structural motifs is used to predict RNA secondary structure from sequences (*1*−*3*). The resulting secondary structure can then provide insight into structure−function relationship and tertiary structure (*4*, *5*). In order to develop the nearest-neighbor model, the thermodynamics of various RNA motifs (helices, bulges, internal loops, hairpins, dangling ends, multibranch loops, etc.) were studied in short RNA oligonucleotides. Partially on the basis of these thermodynamic parameters, an algorithm was developed to predict secondary structure from sequence, and it currently predicts 73% of the known base pairs (*3*). The accuracy of the prediction is based on the measured thermodynamics of the various motifs. An improvement of the thermodynamic parameters for a particular motif would result in an improvement in the accuracy of the predicted secondary structure.

A bulge occurs when a duplex is interrupted by one or more unpaired nucleotides in one of the strands. Such motifs have been determined to be important in many biological processes, including protein binding (*6*−*13*), intron splicing (*14*, *15*), feedback regulation (*16*), and tertiary folding (*17*).

Although their biological significance has been studied and they are one of the most common motifs in RNA secondary structure (*17*−*22*), relatively few studies have been done on the thermodynamics of single nucleotide bulges in DNA (*23*−*25*) and RNA (*26*−*29*), leading to a lack of thermodynamic characterization of single bulges.

This minimal thermodynamic data on single bulges plays an important, albeit assumption-filled, role in predicting RNA secondary structure from sequence (*1*, *3*−*5*, *29*−*32*). The current algorithm, based on nearest-neighbor parameters, assigns a free energy penalty of 3.8 kcal/mol for each single nucleotide bulge in a duplex (*1*, *3*, *29*). This model, however, has many limitations. For instance, it assumes that stacking interactions between the adjacent base pairs are not disturbed by the presence of the single bulge. Also, the current algorithm does not account for any changes in stability due to the identity of the bulge; a pyrimidine bulge is assigned the same penalty as a purine bulge. In addition, the current model does not consider the nearest neighbors. Two studies, however, have shown that bulges having at least one neighboring nucleotide that is identical to the bulged nucleotide are more stable than single base bulges without an identical neighbor (*33*, *34*). Last, the current model does not account for any interactions that are non-nearest-neighbor. One study, however, has shown that alteration of two base pairs that are not adjacent to the bulge can influence the stability of the bulge by at least 1 kcal/mol (*29*). These limitations suggest that the thermodynamic parameters for single nucleotide bulges may be improved, resulting in more accurately predicted secondary structures. This study more thoroughly investigates single nucleotide bulges and proposes an improved model for predicting their stabilities based on the identity of the bulge and its nearest neighbors.

[†] This publication was supported by NSF Grant MCB-0075962.

* To whom correspondence should be addressed. Phone: (814) 332-5363. Fax: (814) 332-2789. Email: mserra@allegheny.edu.

[1] Abbreviations: CPG, controlled pore glass; C_T, total strand concentration; T_M, melting temperature.

10.1021/bi025781q CCC: $22.00
Published on Web 07/25/2002

MATERIALS AND METHODS

RNA Synthesis and Purification. The oligonucleotides were synthesized on CPG support with an Applied Biosystems 392 DNA/RNA synthesizer using the phosphoramidite method (*35*, *36*). The support and phosphoramidites with 2′-*tert*-butyl dimethyl silyl ether protecting groups were acquired from Glen Research (Baltimore, MD). The protecting groups were removed by treatment with a 3:1 (v/v) ammonia/ethanol solution at 55 °C overnight. A disposable filter column was used to filter the oligonucleotides from the support. Removal of the silyl protecting group was achieved with 1 M triethylamine hydrogen fluoride/pyridine (50 equivalents) at 55 °C for 48 h. The addition and subsequent extraction with diethyl ether were performed to remove organic impurities. The sample was lyophilized and redissolved in 5 mM ammonium bicarbonate pH 7.0. This solution was passed through a Waters Sep-Pak C18 water chromatography column to remove inorganic salts. Purification of the oligonucleotide was done by preparative thin-layer chromatography. A large preparative Whatman TLC plate (20 cm × 20 cm, 500 μm thick) was used with 55:35:10 (v/v/v) 1-propanol/ammonia/water as the solvent. The main product band was identified with UV light and scraped from the plate. Distilled water was used to extract the RNA from the silica. The Sep-Pak procedure was repeated to desalt the sample. The purity of the oligonucleotides was checked by analytical thin-layer chromatography similar to that described above.

Optical Melting Experiments. Once the oligonucleotides were pure, the concentrations of the RNA solutions were determined by measuring the high-temperature absorbance of the sample on a Beckman DU 640 Spectrophotometer with a Beckman High Performance Temperature Controller. Equal molar amounts of two different strands of RNA were combined to form a duplex with a single nucleotide bulge. Purified oligonucleotides were lyophilized and redissolved in 1 M NaCl, 0.01 M cacodylic acid, and 0.001 M EDTA at pH 7.0. Using a heating rate of 1 °C/min, curves of absorbance at 280 nm versus temperature were obtained. The melting curves were fit to a two-state model, assuming linear sloping baselines and temperature-independent $\Delta H°$ and $\Delta S°$ (*37*). Additionally, the T_M values at different concentrations were used to calculate thermodynamic parameters according to Borer et al. (*38*):

$$T_M^{-1} = (2.303R/\Delta H°)\log(C_T/4) + (\Delta S°/\Delta H°) \quad (1)$$

For transitions that conform to the two-state model, $\Delta H°$ values from the two methods generally agree within 15%. The Gibb's free energy change at 37 °C was calculated as

$$\Delta G°_{37} = \Delta H° - (310.15)\Delta S° \quad (2)$$

Structural Analysis of Bulge Loops. The recent determination of the crystal structures of the small and large ribosomal subunits has provided a wealth of information about the structure of RNA. Atomic coordinates for the bulge loops were extracted from pdb files using MolReader (Serra and Watson, unpublished data) available at http://webpub-.alleg.edu/employee/m/mserra/. The large ribosomal subunit file used was 1FFK (*39*), and the small subunit file used was 1FJF (*40*). The extracted files were visualized using RASMOL. PDB files for all of the bulge loop structures are available at http://webpub.alleg.edu/employee/m/mserra/.

RESULTS

Recently, X-ray crystallography (*39*, *40*, *47*−*54*) and NMR studies (*23*, *55*−*62*) have been able to produce images which allow for the determination of the position of a single nucleotide bulge. For example, the three-dimensional structure of the 23S rRNA of *Haloarcula marismortui* (*39*) and the 16S rRNA of *Thermus thermophilus* (*40*) have been determined by X-ray crystallography. The three-dimensional structure of single nucleotide bulges in rRNA were analyzed to look for trends in bulge location, bulge stacking, and local effects. A total of 30 single bulge loop structures were identified. The observations of these analyses are shown in Table 1.

The bugle loops form three structural families (Table 1; Figure 1), indicated by the geometry of the bulged nucleotide. The extrahelical family has the bulged nucleotide directed away from the helix, the intercalated family has the bulged nucleotide positioned between the adjacent base pairs, and the side-by-side "side-step" family has the bulge loop arranged coplanar to the base either to the 5′ or 3′ side of the bulge. Table 1 shows that over 2/3 of the bulge loops in the rRNA are extrahelical. The degree to which the bulge interacts with the nucleotides around it varies greatly. Approximately half of the bulged nucleotides do interact with other nucleotides. They interact with other bases through π stacking, hydrogen bonding, or base pair formation (Table 1). These interactions with other nucleotides are both local and tertiary interactions (Figure 1). In the remaining cases, the bulged nucleotide appears to sit within a pocket devoid of obvious interactions. It is interesting to note that all except one of the extrahelical pyrimidine bulges either form stacking interactions or base pairs with purines, helping to stabilize the extrahelical conformation. The lone exception is bulge loop 1178 of the small ribosomal subunit where the nearest interaction is with C_{1037}, which is arranged perpendicular to the bulged $U_{1178.}$ The addition of the bulged residue appears to disrupt the stacking or base pairing of neighboring residues in almost every case.

There are only two examples where the bulge is intercalated in the helix, A_{182} and A_{2436} (Table 1). Interestingly, in both cases, the bulged nucleotide is an adenosine followed by a guanine. The A_{182} and G_{183} nucleotides are highly conserved in bacterial RNAs (*63*). The intercalated base causes a bend in the helix axis (Figure 1). We observed a similar geometry at position 1475 in the large subunit rRNA with the sequence GAC/C C_{1465}; however, one of the nearest-neighbor pairs is a CC and is therefore not included in Table 1.

There are seven examples of the "side-step" family of bulge loops (Table 1). This family of bulge loops forms two distinct groups, one where the bulge nucleotide is side-by-side with the nucleotide to the 5′ side of the bulge and the other where the bulge is side-by-side with the nucleotide to the 3′ side of the bulge. Figure 1 shows an example of the bulge forming a side step with the nucleotide to the 3′ side of the bulge. The side-step geometry causes an increase in the twist of the helix and a break in the stacking interactions on the side of the helix with the bulge.

Table 1: Structural Analysis of Single Bulge Loops in Ribosomal RNAs[a]

position	sequence	bulge characteristics	base pair characteristics
		Large Ribosomal Subunit RNA	
		Extrahelical Bulge	
307	GUC C G_{321}	forms WC[b] pair with A_{338}	
374	UGC G G_{273}	bulge stacks on A_{410}	bases U and C on either side of the bulge are unstacked, UG forms wobble geometry
448	GAC C G_{32}		Slight twist of backbone
834	GUG C C_{846}	bulge stacks on U_{2644}	
856	GAU C A_{787}	bulge stacks on A_{873} and forms base pair with U_{1830}	alters base pair geometry and forms out-of-plane pairs
942	UAG G C_{1022}	bulge perpendicular to helix	wobble base pair
1108	GUG C C_{1252}	bulge forms non-WC pair with A	GC pairs non WC and out-of-plane
1136	UGG G C_{1224}	bulge stacks on C_{1284}	
1259	AGA U U_{1094}		AU pairs non WC and out-of-plane
1351	GAC C G_{1301}		GC pairs non WC and out-of-plane
2021	CAG G C_{1825}	bulge H-bonds to G_{829}	
2108	AUG U C_{2476}	bulge forms pair with A_{2466}	slight twist of helix
2290	UAC A G_{2278}		disruption of base stacking
2636	CAG G C_{2625}		slight tilting of base pairs
2895	CAC G G_{2559}	bulge stacks with A_{2756} and forms WC pair with U_{2755}	
		Intercalated Bulge	
182	GAG C C_{152}	π stacking	base pair C_{152}-G_{183} non planar and helix bent
2436	UAG A C_{2453}	π stacking	planar base pairs and helix bent
		Side-by-Side Bulge	
1009	UCC A G_{273}	UC side-step bulge stacks with A_{960}	Slight twist of helix backbone
2492	UCG G U_{2527}	CG side-step U and C stack with G_{2524}	GU wobble
2615	UGG G C_{2541}	GG side-step bulge stacks with U_{2644}	GU wobble pairs

position	sequence	bulge characteristics	base pair characteristics
		Small Ribosomal Subunit RNA	
		Extrahelical Bulge	
32	UGA A U_{536}	bulge stacks with C_{49}	AU pairs have non-WC geometry
364	UUC A G_{2388}	bulge pairs with A_{56}	Non WC
904	GGG U U_{1373}	bulge stacks with G_{1483}	GU wobble pairs
1032	GUG C C_{1190}	bulge stacks with A_{1183}	
1178	CUG G C_{1043}		base pair $G_{1179}-C_{1043}$ out-of-plane
1377	UAC A G_{900}		U and C unstacked
		Side-by-Side Bulge	
56	CAU G A_{352}	different geometry than standard side-by-side bulge, bulged A stacked with U_{57} and A paired with U_{364}	
579	GGC C G_{628}	G beside C, G stacked on G and U_{625} and C_{580} unstacked	Increased twist of helix backbone
732	CCC G G_{641}	bulged C beside C_{733}	Increased twist of helix backbone
1209	CAC G G_{931}	CA side step	both CG base pairs out-of-plane

[a] Ban et al. (39) and Wimberly et al. (40) [b] Watson−Crick

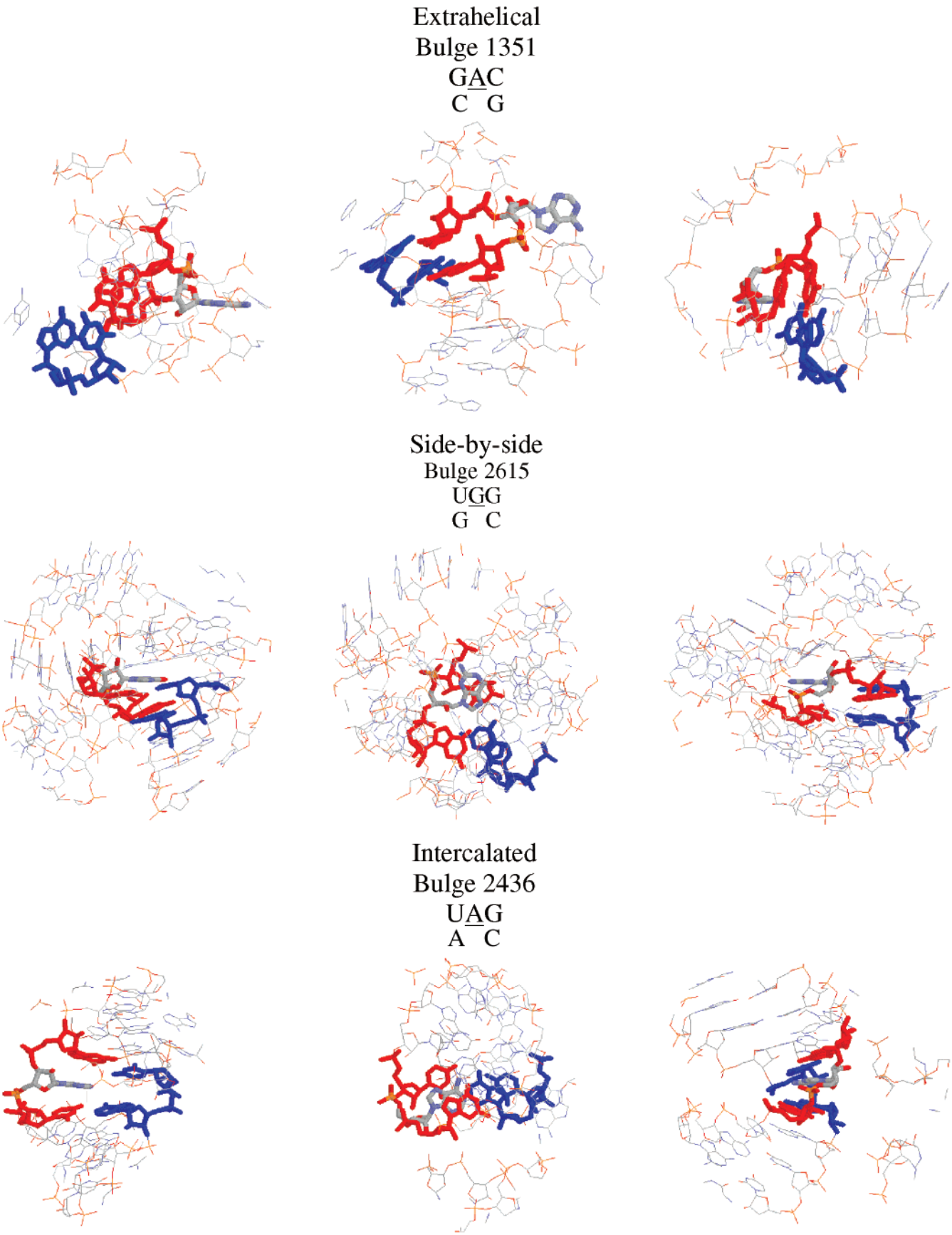

FIGURE 1: Orthogonal views of representative bulge loop structures observed in the large subunit rRNA. The bulged nucleotide is shown in CPK colors, and nearest-neighbor base pairs are shown in red (same strand as bulge) and blue (opposite strand). Surrounding atoms are shown in CPK colors in thin wire frame.

Initial sequences chosen for this study arose from the known tertiary structures (Figure 2) of the 23S rRNA of *H. marismortui* and the 16S rRNA of *T. thermophilus* (*39*−*42*). Additional sequences were chosen due to their highly

Segment of Natural Sequence	Melted Duplex	Location and Structure
-GU**A**C- -CA G-	CCAUU**A**CUACC GGUAA GAUGG	H 2290 Extrahelical and H 377 Extrahelical
-CC**A**GAGC- -GG CUCG-	GCAC**A**GAGG CGUG CUCC	H 2021 Extrahelical and H 2636 Extrahelical
-GAC**A**UGC- -UUG ACG-	GACC**A**UGUC CUGG ACAG	T 56 Side step
-CCC**A**CGAG- -GGG GCUC-	UGAC**A**CUCA ACUG GAGU	H 2895 Extrahelical and T 1209 Side step
-GGG**A**G- -CCC C-	GCG**A**GCG CGC CGC	H 182 Intercalated
-GGGU**G**AACG- -CCCA UUGC-	CAUGU**G**ACUAC GUACA UGAUG	T 32 Extrahelical
-**A**GACCU- -U UGGA-	GACA**G**AGUC CUGU UCAG	H 1259 Extrahelical
-CGUC**U**GGU- -GCUG CCG-	GCAC**U**GAGG CGUG CUCC	T 1178 Extrahelical
-GG**U**GCUG- -CC CGGU-	ACUG**U**GAGU UGAC CUCA	H 834 Extrahelical, and H 1108 Extrahelical, and T 1032 Extrahelical
-UGG**GG**CG- -ACUC GC-	AGCU**GG**CAG UCGAC GUC	T 579 Side step
-CCA**CCC**GU- -GGUG GCA-	UGA**CCC**UCA ACUG GAGU	T 732 Side step
-CU**CC**- -GA G-	UCCU**CC**AAC AGGA GUUG	H 1009Side step

FIGURE 2: Sequences for the oligomers studied here were chosen on the basis of the sequences of known bulge loops and on flanking helices that gave two-state helix-to-coil transitions. Nucleotide positions are represented by T (*T. thermophilus* SSU rRNA) or H (*H. marismortui* LSU rRNA) followed by the nucleotide number.

conserved nature over multiple organisms or to provide a balanced set of bulge loops and base pairs.

Thermodynamic Parameters. Thermodynamic parameters for all oligonucleotides studied are listed in Table 2. The bulges are arranged in order of decreasing free energy for each respective duplex. The parameters are the average values derived from fits of melting curves and from T_M^{-1} versus $\log(C_T/4)$ plots. The parameters from the two methods agree within 15% for all duplexes, suggesting that the two-state model is a reasonable approximation for these transitions (*43*). Thermodynamic data from Longfellow et al. (*29*) was added to expand the data set. Residues in bold are the bulge residues. For Group II sequences, all possible bulge residues are shown in bold.

The sequences were first categorized as Group I or Group II sequences on the basis of the identity of the bulged nucleotide and neighboring nucleotides. Group I sequences were those sequences with a single nucleotide bulge that was not identical to either of the neighboring nucleotides. Group II sequences contained bulges that were identical to at least one of the neighboring nucleotides. Sequences were further divided into two groups: duplexes with purine bulges and duplexes with pyrimidine bulges. The free energy of each bulged nucleotide was calculated from the experimental data and the nearest-neighbor model (*2*) according to the following equation:

$$\Delta G^\circ_{37,\text{bulge}} = \Delta G^\circ_{37,\text{measured}} - \Delta G^\circ_{37,\text{duplex}} \qquad (3)$$

where $\Delta G^\circ_{37,\text{measured}}$ is the value determined experimentally from the melt (eqs 1 and 2) and $\Delta G^\circ_{37,\text{duplex}}$ is calculated from the nearest-neighbor model (*2*) for the duplex as if it did not contain the bulge. Bulge parameters for purine bulges Group I sequences, purine bulges Group II sequences, pyrimidine bulges Group I sequences, and pyrimidine bulges Group II sequences are listed in Table 3. From these data, comparisons can be made concerning the stability of various groups of bulges. In general, pyrimidine bulges are more stable than purine bulges, and Group II sequences are more stable than Group I sequences. Table 4 shows a summary of average values calculated from these bulge parameters.

DISCUSSION

Thermodynamic parameters derived from melting curves of short oligoribonucleotides have been the primary source of energetic characterization of various RNA motifs. Watson−Crick pairs, G·U pairs, dangling ends, terminal mismatches, hairpin loops, tandem mismatches, single mismatches, internal loops, multibranch loops, and bulge loops have all been thermodynamically studied by melting curves (*3*). The energetics of single nucleotide bulges, one of the most prevalent motifs, are based on a rather limited and assumption-filled study of only seven sequences (*29*). To improve the prediction of secondary structure, we optically melted 34 oligoribonucleotides containing single nucleotide bulges and proposed a model to predict their energetics.

Sequence Dependence of Free Energy Increments for Bulges. Similar to previous experiments (*26*, *28*, *29*), in all cases here, the introduction of a bulged base into a duplex results in a destabilization of the duplex. $\Delta G^\circ_{37,\text{bulge}}$ values range from 2.2 to 4.8 kcal/mol, depending on the identity of the bulge and the duplex sequence (Table 3). Three aspects of sequence dependence were studied: the identity of the bulge, the identity of the nucleotides adjacent to the bulge, and the sequence not adjacent to the bulge. In addition, the effect of helix length on the thermodynamics of single bulges was also examined. Each of these is further discussed below.

Purine versus Pyrimidine Bulges. Sequences containing purine bulges were compared to sequences containing pyrimidine bulges within the same group to examine the differences in stability between the two types of bases. The free energies of the bulges ($\Delta G^\circ_{37,\text{bulge}}$) of the two classes of nucleotides in Group I sequences and Group II sequences are shown in Table 3.

Previous studies have reported that purine bulges destabilize the duplex more and, therefore, have a more positive free energy than pyrimidine bulges (*24*, *28*). The results of this experiment concur with the previous reports. On average, pyrimidine single bugles are more stable than purine single bulges by 0.4 kcal/mol. More specifically, purine bulges are on average 0.20 and 0.80 kcal/mol more destabilizing than pyrimidine bulges for Group I and Group II sequences, respectively (Table 4).

To further examine the differences in the thermodynamic effects of the two classes of bulged nucleotides, we plotted the free energy of nearest-neighbor base pairs ($\Delta G^\circ_{37,\text{nn}}$) versus the free energy of the bulge in Figure 3. Linear regression of the data points results in the following lines:

$$\Delta G^\circ_{37,\text{bulge}} = 0.10\Delta G^\circ_{37,\text{nn}} + 3.9 \text{ kcal/mol}$$
for pyrimidine bulges (4)

Table 2: Thermodynamic Parameters for Duplex Formation[a]

	T_M^{-1} vs log C_T/4 plots				average of curve fits			
oligomers[b]	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal/mol)	T_m[c] (°C)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal/mol)	T_m[c] (°C)
			Group I Purine					
CCAUU**A**CUACC GGUAA GAUGG	81.5	231.3	9.8	49.8	85.2	242.7	9.9	49.7
GCAC**A**GAGG CGUG CUCC	64.4	78.4	9.1	49.9	75.7	213.2	9.6	49.9
GACC**A**UGUC CUGG ACAG	62.1	173.9	8.1	45.1	66.8	188.9	8.2	45.1
UGAG**A**GUCA ACUC CAGU	60.7	170.5	7.8	43.8	60.5	169.8	7.9	44.0
UGAC**A**CUCA ACUG GAGU	57.1	161.0	7.2	40.4	56.6	159.4	7.2	40.6
GCG**A**GCG[d] CGC CGC	57.0	162	6.9	38.8	46.8	129	6.9	39.6
CGC**A**CGC[d] GCG GCG	66.4	192	6.9	38.7	45.8	126	6.8	38.9
GACU**A**UGUC CUGA ACAG	53.3	152.3	6.0	34.2	48.7	136.5	6.4	36.1
CAUGU**G**ACUAC GUACA UGAUG	82.8	238.3	8.9	46.0	82.6	237.8	8.9	46.0
UGAC**G**CUCA ACUG GAGU	65.1	185.5	7.6	42.2	57.2	160.2	7.5	42.2
GACU**G**UGUC CUGA ACAG	44.8	120.6	7.4	42.9	41.1	109.2	7.2	42.2
GACA**G**AGUC CUGU UCAG	61.5	176.9	6.6	37.4	54.0	152.6	6.7	38.0
			Group I Pyrimidine					
GCAC**U**GAGG CGUG CUCC	68.9	191.1	9.6	51.4	81.9	231.4	9.6	51.3
ACUG**U**GAGU UGAC CUCA	59.3	165.2	8.0	45.1	74.0	212.2	8.2	44.1
UGAC**U**CUCA ACUG GAGU	54.1	149.3	7.8	44.5	60.6	169.7	7.9	44.4
GCG**U**GCG[d] CGC CGC	55.2	157	6.7	38.0	40.5	108	6.9	40.2
GACA**U**AGUC CUGU UCAG	64.7	187.7	6.5	37.0	71.6	209.8	6.5	36.8
CAUGA**C**GCUAC GUACU CGAUG	80.7	226.1	10.6	53.3	82.5	231.6	10.7	53.5
CAUGU**C**ACUAC GUACA UGAUG	74.3	211.8	8.6	46.1	82.2	236.3	8.9	46.2
UGAG**C**GUCA ACUC CAGU	72.8	207.6	8.5	45.4	76.1	218.1	8.5	45.1
GACU**C**UGUC CUGA ACAG	48.2	131.9	7.3	41.8	56.0	157.0	7.3	41.2
UCCU**C**GAAC AGGA CUUG	58.7	167.7	6.7	37.8	64.0	184.9	6.7	37.6
GACA**C**AGUC CUGU UCAG	73.7	218.5	5.9	34.4	77.3	230.0	6.0	35.0
			Group II Purine					
GACC**AA**GUC CUGGU CAG	68.2	189.6	9.4	50.7	62.1	170.3	9.2	51.1
GACG**AA**GUC CUGCU CAG	65.0	183.0	8.2	45.4	64.1	179.8	8.3	45.8
GAC**AAA**GUC CUGUU CAG	61.7	176.8	6.9	38.8	61.0	174.2	6.9	39.1
AGCU**GG**CAG UCGAC GUC	57.2	153.6	9.6	54.6	56.0	149.5	9.6	55.1
UGA**GGG**UCA ACUCC AGU	55.4	149.5	9.0	51.6	52.6	141.0	8.9	51.8
AGCA**GG**CAG UCGUC GUC	52.2	140.6	8.6	49.6	57.0	155.8	8.7	49.2
AGAC**GG**CAG UCUGC GUC	63.8	179.7	8.1	44.8	64.0	180.5	8.1	44.6
			Group II Pyrimidine					
GACC**UU**GUC CUGGA CAG	71.5	200.6	9.3	49.4	66.9	186.3	9.1	49.5
GAC**UUU**GUC CUGAA CAG	62.3	177.5	7.2	40.5	63.5	181.0	7.3	40.9
UGA**CCC**UCA ACUGG AGU	67.3	186.9	9.3	50.5	68.2	189.7	9.4	50.6
UCCU**CC**AAC AGGAG UUG	54.5	148.6	8.4	48.1	57.0	156.3	8.5	48.1
UCCA**CC**AAC AGGUG UUG	53.7	146.8	8.2	47.1	59.8	165.9	8.4	46.9
			Closed by G−U Pair					
CGGU**A**GUCU GCCG CAGA	59.0	163.5	8.3	46.5	63.2	176.6	8.4	46.5
UGAU**A**CUCA ACUG GAGU	57.0	165.1	5.8	33.2	55.1	158.6	5.9	33.3

[a] Measurements were made in 1.0 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7.0. [b] Nucleotides in bold are the bulge residue. If the bulge residue could be more than one residue, all possibilities are in bold. [c] Calculated at 10^{-4} M oligomer concentration. [d] Longfellow et al. (29)

Table 3: Bulge Parameters

sequence number	oligomer	bulge	nearest neighbors[a]	length[b]	$\Delta G^\circ_{37,bulge}$ (kcal/mol)
		Group I Purine			
1	CCAUU**A**CUACC GGUAA GAUGG	A	Pyr−Pyr	10	4.8
2	GCAC**A**GAGG CGUG CUCC	A	Pyr−Pur	8	4.4
3	GACC**A**UGUC CUGG ACAG	A	Pyr−Pyr	8	4.4
4	UGAG**A**GUCA ACUC CAGU	A	Pur−Pur	8	3.7
5	UGAC**A**CUCA ACUG GAGU	A	Pyr−Pyr	8	4.3
6	GCG**A**GCG[c] CGC CGC	A	Pur−Pur	6	3.9
7	CGC**A**CGC[c] GCG GCG	A	Pyr−Pyr	6	3.9
8	GACU**A**UGUC CUGA ACAG	A	Pyr−Pyr	8	4.0
9	CAUGU**G**ACUAC GUACA UGAUG	G	Pyr−Pur	10	3.8
10	UGAC**G**CUCA ACUG GAGU	G	Pyr−Pyr	8	4.0
11	GACU**G**UGUC CUGA ACAG	G	Pyr−Pyr	8	2.7
12	GACA**G**AGUC CUGU UCAG	G	Pur−Pur	8	3.6
		Group I Pyrimidine			
13	GCAC**U**GAGG CGUG CUCC	U	Pyr−Pur	8	3.9
14	ACUG**U**GAGU UGAC CUCA	U	Pur−Pur	8	3.3
15	UGAC**U**CUCA ACUG GAGU	U	Pyr−Pyr	8	3.6
16	GCG**U**GCG[c] CGC CGC	U	Pur−Pur	6	4.0
17	GACA**U**AGUC CUGU UCAG	U	Pur−Pur	8	3.7
18	CAUGA**C**GCUAC GUACU CGAUG	C	Pur−Pur	10	4.1
19	CAUGU**C**ACUAC GUACA UGAUG	C	Pyr−Pur	10	3.9
20	UGAG**C**GUCA ACUC CAGU	C	Pur−Pur	8	3.0
21	GACU**C**UGUC CUGA ACAG	C	Pyr−Pyr	8	2.9
22	UCCU**C**GAAC AGGA CUUG	C	Pyr−Pur	8	4.1
23	GACA**C**AGUC CUGU UCAG	C	Pur−Pur	8	4.3
		Group II Purine			
24	GACC**AA**GUC CUGGU CAG	A	−[d]	−	3.2
25	GACG**AA**GUC CUGCU CAG	A	−	−	3.6
26	GAC**AAA**GUC CUGUU CAG	A	−	−	3.3
27	AGCU**GG**CAG UCGAC GUC	G	−	−	3.2
28	UGA**GGG**UCA ACUCC AGU	G	−	−	2.5
29	AGCA**GG**CAG UCGUC GUC	G	−	−	4.2
30	AGAC**GG**CAG UCUGC GUC	G	−	−	4.0
		Group II Pyrimidine			
31	GACC**UU**GUC CUGGA CAG	U	−	−	3.3
32	GAC**UUU**GUC CUGAA CAG	U	−	−	3.0
33	UGA**CCC**UCA ACUGG AGU	C	−	−	2.2
34	UCCU**CC**AAC AGGAG UUG	C	−	−	2.4
35	UCCA**CC**AAC AGGUG UUG	C	−	−	2.4
		Closed by G−U Pair[e]			
36	CGGU**A**GUCU GCCG CAGA	A	Pyr−Pur	8	3.3
37	UGAU**A**CUCA ACUG GAGU	A	Pyr−Pyr	8	3.1

[a] Identity of bases located 5′ and 3′ of the bulged nucleotide. Pyr and Pur are the abbreviations for pyrimidine and purine, respectively. [b] Length (residues) of helix excluding the bulge. Since bulges are always located in the middle of the duplex, bulges are 3−5 residues from the helix end. [c] Longfellow et al. (26) [d] Omitted due to positional degeneracy. [e] Shown for comparison. Not included in most calculations.

Table 4: Comparison of Bulge Parameters

oligomer	average $\Delta G^\circ_{37,bulge}$ (kcal/mol)
all duplexes	3.6 ± 0.7
all pyrimidine bulges	3.4 ± 0.7
all purine bulges	3.8 ± 0.6
all Group I sequences	3.8 ± 0.5
all Group II sequences	3.1 ± 0.7
Group I sequences, pyrimidine bulges	3.7 ± 0.5
Group II sequences, pyrimidine bulges	2.6 ± 0.5
Group I sequences, purine bulges	3.9 ± 0.5
Group II sequences, purine bulges	3.4 ± 0.6
Group I sequences closed by purine[a]	3.7 ± 0.4
Group I sequences closed by pyrimidine[b]	3.8 ± 0.7
Group I sequences closed by purine-pyrimidine[c]	4.1 ± 0.2
Group I 6-mer	3.9 ± 0.1
Group I 8-mer	3.7 ± 0.5
Group I 10-mer	4.1 ± 0.4

[a] There is a purine on the 5′ and 3′ sides of the bulge. [b] There is a pyrimidine on the 5′ and 3′ sides of the bulge. [c] There is one purine and one pyrimidine on the 5′ and 3′ sides of the bulge.

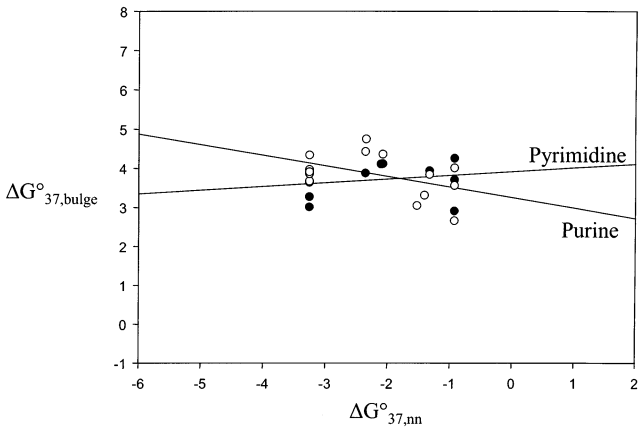

FIGURE 3: Plot of the free energy change for bulge loop formation, $\Delta G^\circ_{37,bulge}$, vs the free energy increment, $\Delta G^\circ_{37,nn}$, for the nearest-neighbor interaction at the site of the bulge. Energy differences between Group I purine (○) and Group I pyrimidine (●) bulges. The solid line is a linear regression of the data points.

The first term is the free energy contribution of the nearest

$$\Delta G^\circ_{37,bulge} = -0.30\Delta G^\circ_{37,nn} + 3.3 \text{ kcal/mol} \quad \text{for purine bulges} \quad (5)$$

neighbors. It is important to note that we have assumed no interruption in the stacking of the nearest neighbors adjacent to the single bulge when calculating $\Delta G^\circ_{37,bulge}$ (as shown in eq 3). The stacking terms included in eqs 4 and 5 are in addition to the normal stacking already assumed. For pyrimidine bulges, it appears as if the stacking of the nearest neighbors is not disrupted; in fact, there is an added stability of 10% of the value of the stacking interaction. On the other hand, for purine bulges, the bulge destabilizes nearest-neighbor stacking by 30% of the value of the stacking interaction. The second term in both equations, 3.9 and 3.3 kcal/mol, is the positive free energy due to inserting a single nucleotide bulge into the duplex.

Nearest Neighbors. Sequences containing purine closing bases adjacent to the bulge were compared to sequences containing pyrimidine closing bases adjacent to the bulge nucleotide (within Group I) to examine the differences in stability between bulges with different closing base pair

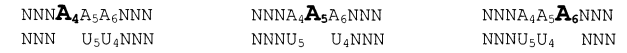

FIGURE 4: Possible conformations of a Group II sequence.

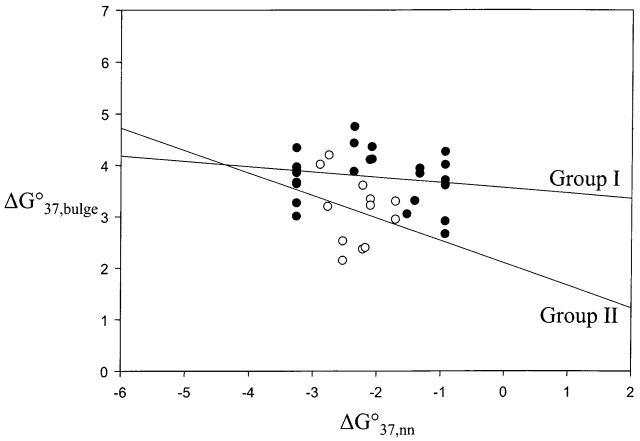

FIGURE 5: Plot of the free energy change for bulge loop formation, $\Delta G^\circ_{37,bulge}$, vs the free energy increment, $\Delta G^\circ_{37,nn}$, for the nearest-neighbor interaction at the site of the bulge. Energy differences between Group I (●) and Group II sequences (○). The solid line is a linear regression of the data points.

types. Previous studies have suggested that single nucleotide bulges located between two purines are more stable than bulges located between two pyrimidines. After investigating tRNA and 5S rRNA structures, Papanicolaou et al. (*31*) suggested a stabilization factor of 1.6 kcal/mol for single base bulges located between two purines than those between two pyrimidines. Similarly, LeBlanc and Morden (*24*) observed a 0.1−0.5 kcal/mol stabilization of bulges between two purines than between two pyrimidines. This study, however, in agreement with Longfellow et al. (*29*), did not observe this effect. Bulges between two purines, between two pyrimidines, and between one purine and one pyrimidine had $\Delta G^\circ_{37,bulge}$ values of 3.7 ± 0.4, 3.8 ± 0.7, and 4.1 ± 0.2 kcal/mol, respectively (Table 4), all within experimental error.

Group I sequences were those sequences with a single nucleotide bulge that was not identical to either of the neighboring nucleotides. Group II sequences contained bulges that were identical to at least one of the neighboring nucleotides. In Group I sequences, the identity of the bulged nucleotide is known. In Group II sequences, however, this is not the case. If the bulged nucleotide is identical to at least one of the neighboring nucleotides, the bulge could theoretically be any of those identical nucleotides. For example, Figure 4 shows the possible location of single nucleotide bulges for a general Group II sequence. As a result, the bulged base may not be one particular nucleotide, but the bulge may change positions with one of its identical neighbors. This may result in conformational freedom, which may increase the entropy, decrease the free energy, and stabilize Group II sequences relative to Group I sequences. Since these sequences are inherently different from the Group I sequences, the two groups of sequences were compared to identify any distinguishing trends.

Table 3 lists $\Delta G^\circ_{37,bulge}$ values for Group I and Group II sequences categorized by the type of bulge. Previous studies have reported an enhanced stabilization for Group II sequences over Group I sequences by 0.3 kcal/mol (*33*, *34*).

Table 5: Comparison of Penalty and Proposed Models

sequence number	oligomer	$\Delta G^\circ_{37,bulge}$ (kcal/mol)	$\Delta G^\circ_{37,}$ penalty model (kcal/mol)	$\Delta\Delta G^\circ_{37,bulge\ and}$ penalty model (kcal/mol)	$\Delta G^\circ_{37,proposed}$ model (kcal/mol)	$\Delta\Delta G^\circ_{37,bulge}$ proposed model) (kcal/mol)
			Group I Purine			
1	CCAUU**A**CUACC GGUAA GAUGG	4.75	3.8	−0.95	4.01	−0.75
2	GCAC**A**GAGG CGUG CUCC	4.43	3.8	−0.63	4.01	−0.42
3	GACC**A**UGUC CUGG ACAG	4.36	3.8	−0.56	3.92	−0.44
4	UGAG**A**GUCA ACUC CAGU	3.67	3.8	0.13	4.28	0.61
5	UGAC**A**CUCA ACUG GAGU	4.34	3.8	−0.54	4.28	−0.06
6	GCG**A**GCG[a] CGC CGC	3.86	3.8	−0.06	4.28	0.42
7	CGC**A**CGC[a] GCG GCG	3.91	3.8	−0.11	4.28	0.37
8	GACU**A**UGUC CUGA ACAG	4.01	3.8	−0.21	3.58	−0.43
9	CAUGU**G**ACUAC GUACA UGAUG	3.84	3.8	−0.04	3.70	−0.14
10	UGAC**G**CUCA ACUG GAGU	3.97	3.8	−0.17	4.28	0.31
11	GACU**G**UGUC CUGA ACAG	2.66	3.8	1.14	3.58	0.92
12	GACA**G**AGUC CUGU UCAG	3.56	3.8	0.24	4.28	0.72
			Group I Pyrimdine			
13	GCAC**U**GAGG CGUG CUCC	3.88	3.8	−0.08	3.66	−0.22
14	ACUG**U**GAGU UGAC CUCA	3.27	3.8	0.53	3.57	0.30
15	UGAC**U**CUCA ACUG GAGU	3.64	3.8	0.16	3.57	−0.07
16	GCG**U**GCG[a] CGC CGC	3.96	3.8	−0.16	3.57	−0.39
17	GACA**U**AGUC CUGU UCAG	3.71	3.8	0.09	3.81	0.10
18	CAUGA**C**GCUAC GUACU CGAUG	4.12	3.8	−0.32	3.69	−0.42
19	CAUGU**C**ACUAC GUACA UGAUG	3.94	3.8	−0.14	3.77	−0.17
20	UGAG**C**GUCA ACUC CAGU	3.01	3.8	0.79	3.57	0.56
21	GACU**C**UGUC CUGA ACAG	2.91	3.8	0.89	3.81	0.90
22	UCCU**C**GAAC AGGA CUUG	4.11	3.8	−0.31	3.69	−0.42
23	CUGA**C**ACAG GACU UGUC	4.26	3.8	−0.46	3.81	−0.45
			Group II Purine			
24	GACC**AA**GUC CUGGU CAG	3.22	3.8	0.58	3.13	−0.09
25	GACG**AA**GUC CUGCU CAG	3.61	3.8	0.19	3.17	−0.44
26	GAC**AAA**GUC CUGUU CAG	3.30	3.8	0.50	3.01	−0.29
27	AGCU**GG**CAG UCGAC GUC	3.20	3.8	0.60	3.33	0.13
28	UGA**GGG**UCA ACUCC AGU	2.53	3.8	1.27	3.26	0.73
29	AGCA**GG**CAG UCGUC GUC	4.20	3.8	−0.40	3.33	−0.88
30	AGAC**GG**CAG UCUGC GUC	4.02	3.8	−0.22	3.37	−0.65
			Group II Pyrimidine			
31	GACC**UU**GUC CUGGA CAG	3.34	3.8	0.46	2.89	−0.45
32	GAC**UUU**GUC CUGAA CAG	2.95	3.8	0.85	2.93	−0.02
33	UGA**CCC**UCA ACUGG AGU	2.15	3.8	1.65	2.85	0.70
34	UCCU**CC**AAC AGGAG UUG	2.37	3.8	1.43	2.88	0.51
35	UCCA**CC**AAC AGGUG UUG	2.40	3.8	1.40	2.88	0.48
			Closed by G−U Pair			
36	CGGU**A**GUCU GCCG CAGA	3.31	3.8	0.49	3.93	0.62
37	UGAU**A**CUCA ACUG GAGU	3.05	3.8	0.75	4.01	0.96
	average	3.56	3.80	0.24	3.62	0.06
	deviation	0.64	0.00	0.64	0.45	0.52

[a] Longfellow et al. (26)

This study also observed an enhanced stabilization of Group II bulge sequences; Group II bulge sequences on average were 0.7 kcal/mol more stable than Group I bulge sequences.

To further examine the differences in the thermodynamic effects of the two Groups of sequences, we constructed Figure 5 similar to Figure 3 and plotted the free energy of nearest-neighbor base pairs ($\Delta G^\circ_{37,nn}$) versus the free energy of the bulge. When considering Group II sequences, it is apparent that $\Delta G^\circ_{37,nn}$ is dependent on knowing which of the possible nucleotides is acting as the actual bulged nucleotide. Without doing structural studies on each of the sequences, it is impossible to know which nucleotide is the bulge. We could identify the bulged nucleotide based on a variety of simple assumptions. For example, we could assume the bulge will be situated between the pair of nearest neighbors that is thermodynamically least stabilizing. Or we could assume that the bulge will be situated closest to the helix end. Both of these assumptions, however, could still result in multiple possibilities for the bulge location. So for simplicity and lack of structural data, we will assume that the bulge has an equal probability of being located in each of the possible locations and will use the average of all possible $\Delta G^\circ_{37,nn}$. Linear regression of the data points results in the following lines:

$$\Delta G^\circ_{37,bulge} = -0.10\Delta G^\circ_{37,nn} + 3.6 \text{ kcal/mol} \quad \text{for Group I sequences} \quad (6)$$

To compare Group I sequences to Group II sequences,

$$\Delta G^\circ_{37,bulge} = -0.40\Delta G^\circ_{37,nn} + 2.1 \text{ kcal/mol} \quad \text{for Group II sequences} \quad (7)$$

we must look at the range of possible $\Delta G^\circ_{37,nn}$ values, −0.93 to −3.42 kcal/mol (*2*). Due to the different slopes of the two lines, the absolute difference between the two types of sequences can be minimized by setting the two equations equal to each other at the middle of the $\Delta G^\circ_{37,nn}$ range, −2.18 kcal/mol. This can only be done by adding a stabilization of −0.8 kcal/mol to Group II sequences. This value is more negative than the −0.3 kcal/mol stabilization of Group II sequences reported by Zhu and Wartell (*34*) for RNA determined by temperature gradient gel electrophoresis. It appears as if pyrimidine bulges are more stabilized by being located in a Group II sequence. Moving a pyrimidine from a Group I to a Group II sequence stabilizes the bulge by, on average, 1.1 kcal/mol. Doing the same for a purine bulge only stabilizes the bulge, on average, 0.5 kcal/mol. The difference, 0.6 kcal/mol, is the difference between purine and pyrimidine bulge I sequences. The incorporation of the additional stabilization term for Group II sequences into programs such as mfold (*5*) must be done with care since the output is a single structure while the thermodynamic stability represents an ensemble of structures as shown in Figure 4.

Non-Nearest-Neighbor Effects. Longfellow et al. (*29*) suggested that predicting the free energy of RNA based on the length of the bulge and the nearest neighbors was an oversimplification. They reported a destabilization of approximately 2 kcal/mol for changing two base pairs not adjacent to the bulge in one of their sequences. The effects of the presence of bulges on ethidium intercalation in RNA also suggests a non-nearest-neighbor effect of bulges (*44*). In this experiment, three pairs of sequences had the same bulge and nearest-neighbors but different nonnearest neighbors, duplexes 4 and 6, 5 and 7, and 14 and 16. The average deviation of $\Delta G^\circ_{37,bulge}$ for these sets of duplexes is 0.3 kcal/mol. This value falls within the standard deviation of 0.7 kcal/mol for all of the duplexes studied. It is not evident, therefore, that non-nearest neighbors affected the values calculated for $\Delta G^\circ_{37,bulge}$.

Length of Helix. An analysis of rRNA secondary structures suggests that some single mismatches are preferentially located nearer to the helix ends (*41*, *42*). A thermodynamic study of single mismatches demonstrated that moving the mismatch closer to the helix end resulted in an enhanced stability of the mismatch (*45*). It has been suggested that the same may be true of single nucleotide bulges (*46*), and an analysis of rRNA secondary structures also suggests that some single bulges are preferentially located nearer to the helix ends (Znosko, unpublished data). Since all bulges in Group I sequences are located in the center of the helix, a comparison between helices of different lengths is also a comparison between bulges which are located at different distances from the helix ends. In addition to the 6-mers (length of strand not including the bulge) studied by Longfellow et al. (*29*), this experiment measured the thermodynamics of 8- and 10-mers. So the bulge is located 3, 4, or 5 nucleotides from the end of the helix. All Group I 6, 8, and 10-mers were compared to determine any trends in $\Delta G^\circ_{37,bulge}$ on the basis of duplex length and had average $\Delta G^\circ_{37,bulge}$ values of 3.9 ± 0.1, 3.7 ± 0.5, and 4.1 ± 0.4 kcal/mol, respectively (Table 4). It is not apparent that a pattern exists between the distance from the helix end and $\Delta G^\circ_{37,bulge}$.

Penalty Model. The current model (*1*, *3*, *29*) used to predict the free energy of a bulge loop is the penalty model. This model assumes that the bulged nucleotide does not interfere with the base pairing or stacking interactions between the bases adjacent to the bulge. The bulged nucleotide is given a free energy value of 3.8 kcal/mol. This value is constant for all single nucleotide bulge loops. The free energy of a duplex containing a single nucleotide bulge can be determined using a nearest-neighbor calculation with the addition of the penalty for a single nucleotide bulge loop (*1*−*3*, *29*):

$$\Delta G^\circ_{37,duplexwithbulge} = \Delta G^\circ_{37,refduplex} + \Delta G^\circ_{37,penalty} \quad (8)$$

where $\Delta G^\circ_{37,ref\ duplex}$ is the ΔG°_{37} of the duplex with the bulge removed and $\Delta G^\circ_{37,penalty}$ is 3.8 kcal/mol for single bulges, regardless of the identity of the bulge or the identity of the nearest neighbors. In this study, however, the average $\Delta G^\circ_{37,bulge}$ is 3.6 kcal/mol and differences between the identity of the bulge and the identity of the nearest neighbors were observed. By incorporating these data, a less simplified and more accurate model can be devised to calculate the free energy of duplexes containing single nucleotide bugles.

Development of Proposed Model. This study thermodynamically characterized duplexes containing single nucleotide bulges. The data suggests that purine and pyrimidine bulges and Group I and Group II sequences differ significantly in their thermodynamics. From these data, a model was developed that will predict the free energy of a duplex containing a single bulge. The identity of the bulge was considered first and then the type of sequence was accounted for.

By combining eqs 4 and 5 with the data derived from eqs 6 and 7, a new model was devised to account for differences in the identity of the bulge and differences in the nearest neighbors:

$$\Delta G^\circ_{37,\text{bulge}} = 3.9\text{kcal/mol} + 0.10\Delta G^\circ_{37,\text{nn}} + \beta \quad \text{for pyrimidine bulges} \quad (9)$$

where $\Delta G^\circ_{37,\text{nn}}$ is the ΔG°_{37} for the neighboring base pairs

$$\Delta G^\circ_{37,\text{bulge}} = 3.3\text{kcal/mol} - 0.30\Delta G^\circ_{37,\text{nn}} + \beta \quad \text{for purine bulges} \quad (10)$$

for Group I sequences and the average ΔG°_{37} of all possible nearest-neighbor combinations for Group II sequences, and β is 0.0 and −0.8 kcal/mol for Group I and Group II sequences, respectively. In Table 5, the sequences in this study were used to compare this proposed model to the penalty model (*1*, *3*). Not only has the average value of $\Delta\Delta G^\circ_{37}$ decreased from 0.24 to 0.06 kcal/mol, but the standard deviation of $\Delta\Delta G^\circ_{37}$ has decreased from 0.64 to 0.52 kcal/mol by using the new model. In addition, two types of sequences were greatly improved by the use of the proposed model, pyrimidine bulges and Group II sequences. The $\Delta\Delta G^\circ_{37}$ for Group II sequences was improved from −0.69 to −0.02 kcal/mol, and the $\Delta\Delta G^\circ_{37}$ for pyrimidines was improved from 0.42 to 0.06 kcal/mol.

Using Thermodynamics to Predict the Conformation of the Bulge. Due to the variety of interactions in which a bulged base may participate, relationships between thermodynamics, the identity of the bulge and the nearest neighbors, and conformation of the bulge cannot, at present, be determined. For example, the conformation of an adenosine bulge in identical sequences determined by NMR (*55*, *64*) and X-ray crystallography (*47*, *65*) has been determined to have different geometries. The NMR data indicated that the bulge was intercalated within the helix while crystallography indicated that the bulge was extrahelical. Likewise, the bulge sequence C**A**C/G G may adopt either an extrahelical or sidestep geometry (Figure 1). Since regions with the same bulging nucleotide and nearest neighbors can vary greatly in their three-dimensional structure, it is clear that thermodynamics, identity of the bulge, and identity of the nearest neighbors cannot predict bulge loop geometry.

REFERENCES

1. Serra, M. J., and Turner, D. H. (1995) *Methods Enzymol. 259*, 242−261.
2. 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.
3. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) *J. Mol. Biol. 288*, 911−940.
4. Turner, D. H., and Sugimoto, N. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.
5. Zuker, M. (1989) *Science 244*, 48−52.
6. Peattie, D. A., Douthwaite, S., Garrett, R. A., and Noller, H. F. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 7331−7335.
7. Christiansen, J., Douthwaite, S. R., Christensen, A., and Garrett, R. A. (1985) *EMBO J. 4*, 1019−1024.
8. Romaniuk, P. J., Lowary, P., Wu, H. N., Stormo, G., and Uhlenbeck, O. C. (1987) *Biochemistry 26*, 1563−1568.
9. Wu, H. N., and Uhlenbeck, O. C. (1987) *Biochemistry 26*, 8221−8227.
10. Baudin, F., and Romaniuk, P. J. (1989) *Nucleic Acids Res. 17*, 2043−2056.
11. Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., Lowe, A. D., Singh, M., and Skinner, M. A. (1990) *EMBO J. 9*, 4145−4153.
12. Roy, S., Delling, U., Chen, C. H., Rosen, C. A., and Sonenberg, N. (1990) *Genes Dev. 4*, 1365−1373.
13. Harper, J. W., and Logsdon, N. J. (1991) *Biochemistry 30*, 8060−8066.
14. Schmelzer, C., and Schweyen, R. J. (1986) *Cell 46*, 557−565.
15. Parker, R., Siliciano, P. G., and Guthrie, C. (1987) *Cell 49*, 229−239.
16. Climie, S. C., and Friesen, J. D. (1987) *J. Mol. Biol. 198*, 371−381.
17. Woese, C. R., and Gutell, R. R. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3119−3122.
18. Michel, F., and Dujon, B. (1983) *EMBO J. 2*, 33−38.
19. Gutell, R. R., Weiser, B., Woese, C. R., and Noller, H. F. (1985) *Prog. Nucleic Acid Res. Mol. Biol. 32*, 155−216.
20. Gutell, R. R., and Fox, G. E. (1988) *Nucleic Acids Res. 16*, R175−R269.
21. Erdmann, V. A., and Wolters, J. (1986) *Nucleic Acids Res. 14*, R1−R59.
22. Michel, F., and Jacquier, A. (1987) *Cold Spring Harbor Symp. Quantum Biol. 52*, 201−212.
23. Woodson, S. A., and Crothers, D. M. (1987) *Biochemistry 26*, 904−912.
24. Leblanc, D. A., and Morden, K. M. (1991) *Biochemistry 30*, 4042−4047.
25. Zieba, K., Chu, T. M., Kupke, D. W., and Marky, L. A. (1991) *Biochemistry 30*, 8018−8026.
26. Fink, T. R., and Crothers, D. M. (1972) *J. Mol. Biol. 66*, 1−12.
27. Fink, T. R., and Krakauer, H. (1975) *Biopolymers 14*, 433−436.
28. Groebe, D. R., and Uhlenbeck, O. C. (1989) *Biochemistry 28*, 742−747.
29. Longfellow, C. E., Kierzek, R., and Turner, D. H. (1990) *Biochemistry 29*, 278−285.
30. Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., and Gralla, J. (1973) *Nature 246*, 40−41.
31. Papanicolaou, C., Gouy, M., and Ninio, J. (1984) *Nucleic Acids Res. 12*, 31−44.
32. Zuker, M., and Stiegler, P. (1981) *Nucleic Acids Res. 9*, 133−148.
33. Ke, S. H., and Wartell, R. M. (1995) *Biochemistry 34*, 4593−4600.
34. Zhu, J., and Wartell, R. M. (1999) *Biochemistry 38*, 15986−15993.
35. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. J. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.
36. Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., and Usman, N. (1995) *Nucleic Acids Res. 23*, 2677−2684.
37. McDowell, J. A., and Turner, D. H. (1996) *Biochemistry 35*, 14077−14089.
38. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843−853.
39. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) *Science 289*, 905−920.
40. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) *Nature 407*, 327−339.
41. Gutell, R. R., Gray, M. W., and Schnare, M. N. (1993) *Nucleic Acids Res. 21*, 3055−3074.
42. Gutell, R. R. (1994) *Nucleic Acids Res. 22*, 3502−3507.
43. Freier, S. M., Petersheim, M., Hickey, D. R., and Turner, D. H. (1984) *J. Biomol. Struct. Dyn. 1*, 1229−1242.
44. White, S. A., and Draper, D. E. (1989) *Biochemistry 28*, 1892−1897.
45. Kierzek, R., Burkard, M. E., and Turner, D. H. (1999) *Biochemistry 38*, 14214−14223.
46. Turner, D. H. (1992) *Curr. Opin. Struct. Biol. 2*, 334−337.
47. Miller, M., Harrison, R. W., Wlodawer, A., Appella, E., and Sussman, J. L. (1988) *Nature 334*, 85−86.
48. Joshuator, L., Frolow, F., Appella, E., Hope, H., Rabinovich, D., and Sussman, J. L. (1992) *J. Mol. Biol. 225*, 397−431.
49. Portmann, S., Grimm, S., Workman, C., Usman, N., and Egli, M. (1996) *Chem. Biol. 3*, 173−184.
50. Clemons, W. M., May, J. L. C., Wimberly, B. T., McCutcheon, J. P., Capel, M. S., and Ramakrishnan, V. (1999) *Nature 400*, 833−840.
51. Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B., and Steitz, T. A. (1999) *Nature 400*, 841−847.

52. Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N., and Noller, H. F. (1999) *Science 285*, 2095−2104.
53. Xiong, Y., and Sundaralingam, M. (2000) *RNA 6*, 1316−1324.
54. Berglund, J. A., Rosbash, M., and Schultz, S. C. (2001) *RNA 7*, 682−691.
55. Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Itakura, K., and Breslauer, K. J. (1982) *Biochemistry 21*, 445−451.
56. Morden, K. M., Chu, Y. G., Martin, F. H., and Tinoco, I. (1983) *Biochemistry 22*, 5557−5563.
57. Woodson, S. A., and Crothers, D. M. (1988) *Biochemistry 27*, 3130−3141.
58. Woodson, S. A., and Crothers, D. M. (1989) *Biopolymers 28*, 1149−1177.
59. Kalnik, M. W., Norman, D. G., Swann, P. F., and Patel, D. J. (1989) *J. Biol. Chem. 264*, 3702−3712.
60. Kalnik, M. W., Norman, D. G., Zagorski, M. G., Swann, P. F., and Patel, D. J. (1989) *Biochemistry 28*, 294−303.
61. Kalnik, M. W., Norman, D. G., Li, B. F., Swann, P. F., and Patel, D. J. (1990) *J. Biol. Chem. 265*, 636−647.
62. Thiviyanathan, V., Guliaev, A. B., Leontis, N. B., and Gorenstein, D. G. (2000) *J. Mol. Biol. 300*, 1143−1154.
63. Gutell, R. R., Cannone, J. J., Shang, Z., Du, Y., and Serra, M. J. (2000) *J. Mol. Biol. 304*, 335−354.
64. Roy, S., Sklenar, V., Appella, E., and Cohen, J. S. (1987) *Biopolymers 26*, 2041−2052.
65. Joshuator, L., Rabinovich, D., Hope, H., Frolow, F., Appella, E., and Sussman, J. L. (1988) *Nature 334*, 82−84.

BI025781Q