BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 8669–8679 8669
DOI: 10.1021/bi100146z

Positional and Neighboring Base Pair Effects on the Thermodynamic Stability of RNA Single Mismatches†

Amber R. Davis and Brent M. Znosko*

Department of Chemistry, Saint Louis University, Saint Louis, Missouri 63103

Received January 29, 2010; Revised Manuscript Received August 2, 2010

ABSTRACT: Many naturally occurring RNA structures contain single mismatches, many of which occur near the ends of helices. However, previous thermodynamic studies have focused their efforts on thermodynamically characterizing centrally placed single mismatches. Additionally, algorithms currently used to predict secondary structure from sequence are based on two assumptions for predicting the stability of RNA duplexes containing this motif. It has been assumed that the thermodynamic contribution of small RNA motifs is independent of both its position in the duplex and the identity of the non-nearest neighbors. Thermodynamically characterizing single mismatches three nucleotides from both the 3′ and 5′ ends (i.e., off-center) of an RNA duplex and comparing these results to those of the same single mismatch−nearest neighbor combination centrally located have allowed for the investigation of these effects. The thermodynamic contributions of 13 single mismatch−nearest neighbor combinations are reported, but only nine combinations are studied at all three duplex positions and are used to determine trends and patterns. In general, the 5′- and 3′-shifted single mismatches are relatively similar, on average, and more favorable in free energy than centrally placed single mismatches. However, close examination and comparison shows there are several associated idiosyncrasies with these identified general trends. These peculiarities may be due, in part, to the identities of the single mismatch, the nearest neighbors, and the non-nearest neighbors, along with the effects of the single mismatch position in the duplex. The prediction algorithm recently proposed by Davis and Znosko [Davis, A. R., and Znosko, B. M. (2008) *Biochemistry 47*, 10178−10187] is used to predict the thermodynamic parameters of single mismatch contribution, and those values are compared to the measured values presented here. This comparison suggests the proposed model is a good approximation but could be improved by the addition of parameters that account for positional and/or non-nearest neighbor effects. However, more data are required to improve our understanding of these effects and to accurately account for them.

The known functions and roles of RNA in nature are vast. Similarly, the types of secondary structure motifs present in RNA are also diverse. These include canonical helices and noncanonical regions, such as internal loops, bulges, hairpins, and multibranch loops. Single mismatches, or 1 × 1 internal loops, are the most frequently occurring secondary structure motif in rRNA (*1*) and often times serve integral structural and/or functional roles (*2−12*). Consequently, single mismatches have been utilized in therapeutic techniques as a target (*13−16*), an aptamer drug (*17*, *18*), and a probe (*19−22*).

One example of a therapeutic technique utilizing this secondary structure motif is demonstrated in recent studies in which the positional effect of single mismatches on the efficacy of RNA interference (RNAi)[1] activity was examined via the placement of mismatches at the center and 5′ and 3′ ends of the sense-stranded small interfering RNA (ss-siRNA) component (*19−22*). siRNA duplexes with single mismatches placed at the 3′ terminus of the sense strand showed increased RNAi activity when compared to perfectly matched siRNA duplexes or those containing mismatches at the center or 5′ end. These enhanced siRNAs are known as "fork-siRNA duplexes" (*19*, *22*). Furthermore, the activity of short hairpin RNAs (shRNAs) has also been shown to be increased by the incorporation of 3′ terminal single mismatches and a decreased overall thermodynamic stability (ΔG). Westerhout and Berkhout further demonstrated shRNAs were most effective if they possessed a free energy value within a defined window, while also containing 3′ terminal mismatches (*21*). Synthetic fork-siRNAs and shRNAs are effective therapeutics for suppressing gene expression by interacting with the RNA-induced silencing complexes (RISCs) and thereby invoking sequence specific RNAi activity. The 3′ terminal mismatches allow for recognition and duplex unwinding by the RISC helicase activity (*23−27*). It has been proposed they also minimize off-target gene silencing by resulting in direction specific dissociation of the siRNA and act as sequence specific RNAi mediators in RISC (*19*).

The algorithms most commonly used to predict secondary structure from sequence are based on free energy minimization (*28−34*) using nearest neighbor parameters and have been incorporated into user-friendly, computer programs. In this method, a given sequence is folded into possible conformations. The total free energy values for each conformation are calculated by summing together the free energy parameters of all secondary structure motifs (experimental or predicted). This results in an optimal

†This work was supported by Grant R15GM085699 from the National Institute of General Medical Sciences. A.R.D. has been supported by a Monsanto Scholars Graduate Fellowship and the Saint Louis University Graduate School Dissertation Fellowship.
*To whom correspondence should be addressed. Phone: (314) 977-8567. Fax: (314) 977-2521. E-mail: znoskob@slu.edu.
[1]Abbreviations: R, purine nucleotide; RISC, RNA-induced silencing complexes; RNAi, RNA interference; shRNA, short hairpin RNA; SM, single mismatch; ss-siRNA, sense-stranded small interfering RNA; Y, pyrimidine nucleotide.

structure and a series of suboptimal structures. The optimal structure has the lowest free energy and is predicted to be the predominate structure in solution. These prediction algorithms utilize two methods when assigning free energy parameters to noncanonical regions. If thermodynamic parameters for a given motif are available, the experimentally determined free energy value is assigned. If such parameters have not been experimentally determined, a predicted free energy value is assigned.

Much work has been done to thermodynamically characterize single mismatches placed in the center of a duplex (*1*, *35*–*37*). These studies have shown the contribution of single mismatches to duplex thermodynamics to be dependent on the identity of the nearest neighbors and the identity of the mismatched nucleotides (*1*, *35*–*37*). For example, we (*36*) recently proposed a single mismatch specific algorithm that utilizes three parameters consisting of a total of nine variables. The free energy of an RNA duplex containing a single mismatch that has not been thermodynamically characterized can be calculated by

$$\Delta G^\circ_{37,\,\text{single mismatch}} = \Delta G^\circ_{37,\,\text{mismatch nt}} + \Delta G^\circ_{37,\,\text{mismatch}-\text{NN interaction}} + \Delta G^\circ_{37,\,\text{AU}} + \Delta G^\circ_{37,\,\text{GU}} \quad (1)$$

where $\Delta G^\circ_{37,\text{mismatch nt}}$ is −0.3, −2.1, or −0.6 kcal/mol for A·G, G·G, or U·U mismatches, respectively; $\Delta G^\circ_{37,\text{mismatch}-\text{NN interaction}}$ is 0.6, 0.0, 0.6, −0.5, or −0.9 kcal/mol for $[\substack{5'\mathrm{Y\underline{R}R}3'\\3'\mathrm{R\underline{R}Y}5'}]$, $[\substack{5'\mathrm{R\underline{Y}Y}3'\\3'\mathrm{Y\underline{Y}R}5'}]$, $[\substack{5'\mathrm{Y\underline{Y}R}3'\\3'\mathrm{R\underline{Y}Y}5'}]$, $[\substack{5'\mathrm{Y\underline{R}Y}3'\\3'\mathrm{R\underline{Y}R}5'}]$, or $[\substack{5'\mathrm{R\underline{R}Y}3'\\3'\mathrm{Y\underline{Y}R}5'}]$ mismatch and nearest neighbor combinations, respectively, when A and G are categorized as purines (R) and C and U are categorized as pyrimidines (Y); $\Delta G^\circ_{37,\text{AU}}$ is a penalty of 1.1 kcal/mol for replacing a G-C closing base pair with an A-U base pair; and $\Delta G^\circ_{37,\text{GU}}$ is a penalty of 1.4 kcal/mol for replacing a G-C closing base pair with a G-U base pair. All other combinations of single mismatch nucleotides and nearest neighbors are assumed to contribute no favorable or unfavorable contributions to duplex stability and are assigned a free energy value of zero (*36*).

In addition to the identity of the nearest neighbors and mismatched nucleotides, it is important to note studies have reported the dependence of the thermodynamic stability of small RNA motifs on the duplex position and the identity of non-nearest neighbors (*37*–*42*). An example of the thermodynamic dependence on the motif's duplex position was demonstrated by Kierzek and co-workers, who investigated the thermodynamics of single mismatches (*37*). A·A and U·U single mismatches had increased stability the closer they were placed toward the end of the duplex, while G·G single mismatches were unaffected by the position in the duplex (*37*). An investigation of bulges of one nucleotide (*38*) demonstrated a thermodynamic dependence on the identity of non-nearest neighbors and further showed a clear and direct relationship between the thermodynamic stability of the parental duplex and the thermodynamic contribution of the bulge. For example, the bulge $[\substack{5'\mathrm{G\underline{C}U}3'\\3'\mathrm{C\,\underline{\ \ }\,A}5'}]$ was placed in the center of two different duplex sequences, and a 3.0 kcal/mol difference in free energy contribution between the two duplexes was obtained (*38*). Similarly, a recent thermodynamic study on 1 × 2 loops (*43*) showed a strong dependence on the identity of non-nearest neighbors. Placing the 1 × 2 loop $[\substack{5'\mathrm{C\ \underline{C}\ C}3'\\3'\mathrm{G\underline{AU}G}5'}]$ in the center of two different duplex sequences resulted in a difference in free energy contribution of 2.8 kcal/mol (*43*, *44*). Additionally, for tetraloops, or hairpins of four, non-nearest neighbor effects were observed upon comparison of the thermodynamics of the tetraloop contribution to duplex stability when placed in the sequences 5′GCCNNNNGGC3′ and 5′GGCNNNNGCC3′. When tetraloops were placed in the latter stem sequence, they were, on average, 0.6 kcal/mol more stable than in the former sequence (*44*). Because current secondary structure prediction algorithms assume the thermodynamic contribution of small RNA motifs is independent of both its position in the duplex and the identity of the non-nearest neighbors, these results suggest an improved understanding of positional and non-nearest neighbor effects may lead to improved algorithms for predicting secondary structure from sequence.

(a) 5'G **UAC** ACCUG3'
3'C **AGG** UGGAC5'

(b) 5'GAC **UAC** CUG3'
3'CUG **AGG** GAC5'

(c) 5'GACCU **UAC** G3'
3'CUGGA **AGG** C5'

FIGURE 1: Three duplexes each containing the same single mismatch–nearest neighbor combination (boxed) at three different duplex positions: (a) 5′-shifted, (b) central, and (c) 3′-shifted. All three duplexes contain the same stem sequence, but when shifted among the three different positions, the single mismatch has different non-nearest neighbors.

This work investigates the positional and non-nearest neighbor effects on the thermodynamic contribution of single mismatches by thermodynamically characterizing the same single mismatch–nearest neighbor combinations at three duplex positions within the same stem. Results show positional and/or non-nearest neighbor effects play a role in defining the thermodynamic contribution of single mismatches to duplex stability.

MATERIALS AND METHODS

Sequence Design. Single mismatches chosen for this study were those that occur frequently in nature (*35*). Two single mismatches outside the 30 most frequently occurring, $[\substack{5'\mathrm{A\underline{A}G}3'\\3'\mathrm{U\underline{A}C}5'}]$ and $[\substack{5'\mathrm{A\underline{C}G}3'\\3'\mathrm{U\underline{C}C}5'}]$, were also chosen to allow at least one example of each of the seven combinations of single mismatches to be represented. Single mismatches and nearest neighbors were placed in three different positions within the same stem (Figure 1). The single mismatches were placed in the center or off-center (both 5′- and 3′-shifted). Although the identity of the single mismatch and nearest neighbors are kept constant, moving the single mismatch–nearest neighbor combination between the three duplex positions changes the mismatch's non-nearest neighbors. Further details for the design of sequences were described previously (*35*, *43*).

RNA Synthesis and Purification. Oligonucleotides were ordered from Integrated DNA Technologies (Coralville, IA). The synthesis and purification of the oligonucleotides followed standard procedures and have been previously described (*35*, *45*).

NMR Sample Preparation. Five representative duplexes, $[\substack{5'\mathrm{GAC\ U\underline{U}G\ CUG}3'\\3'\mathrm{CUG\ G\underline{U}C\ GAC}5'}]$, $[\substack{5'\mathrm{G\ A\underline{U}C\ ACCUG}3'\\3'\mathrm{C\ U\underline{U}G\ UGGAC}5'}]$, $[\substack{5'\mathrm{G\ G\underline{C}U\ ACCUG}3'\\3'\mathrm{C\ C\underline{A}A\ UGGAC}5'}]$, $[\substack{5'\mathrm{GACCU\ G\underline{C}U\ G}3'\\3'\mathrm{CUGGA\ C\underline{A}A\ C}5'}]$, and $[\substack{5'\mathrm{G\ A\underline{A}G\ ACCUG}3'\\3'\mathrm{C\ U\underline{A}C\ UGGAC}5'}]$, were studied by NMR spectroscopy. NMR was used to confirm the formation of the single mismatch-containing duplex conformation as the predominate structure in solution. The total concentration of each single strand was calculated from the extinction coefficient and the measured absorbance at 280 nm and 25 °C using Beer's law. An equal molar ratio of non-self-complementary strands was mixed to form a duplex containing a single mismatch, and the total duplex concentration was calculated using the same method previously described for calculating single strand concentrations (*35*, *43*). All duplex concentrations were 1−2 mM. The resulting duplexes were lyophilized and redissolved in 225 μL of 80 mM NaCl, 3 mM NaH_2PO_4, 7 mM Na_2HPO_4, 0.5 mM EDTA (pH 7.0), and 25 μL of 99.9% D_2O (Sigma-Aldrich, St. Louis, MO) for exchangeable proton NMR experiments.

NMR Spectroscopy. All spectra were recorded on a Bruker Avance III 400 MHz NMR spectrometer with a 5 mm broadband probe, two radiofrequency channels with pulse field gradient waveform generators, and a digital variable-temperature control unit. Exchangeable proton spectra were recorded using a jump-and-return pulse sequence (*46*) optimized for water suppression and for maximum peak intensity of the imino proton resonances. Experiments were conducted at 5 °C intervals, with temperatures ranging from 0 to 45 °C. The data were processed using TOPSPIN (Bruker BioSpin, Billerica, MA).

Optical Melting Experiments and Thermodynamics. The methods used to determine the concentration of the single strands and to form duplexes from the single strands are standard and were described previously (*35*, *43*). Optical melting experiments were performed in 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA (pH 7.0). Melting curves (absorbance vs temperature) were obtained, and duplex thermodynamics were determined as described previously (*35*). The thermodynamic contributions of single mismatches to duplex thermodynamics ($\Delta G^\circ_{\text{single mismatch}}$, $\Delta H^\circ_{\text{single mismatch}}$, and $\Delta S^\circ_{\text{single mismatch}}$) were determined by subtracting the canonical Watson−Crick contribution from the measured duplex thermodynamics. This type of calculation has been described previously (*35*). To explicitly demonstrate this type of calculation, the following explanation and examples are given. The total free energy change for duplex formation can be approximated by a nearest neighbor model (*47*) that is the sum of energy increments for helix initiation, nearest neighbor interactions between base pairs, and the single mismatch contribution. For example

$$\Delta G^\circ_{37}\begin{bmatrix} {}^{5'}\text{G U}\underline{\text{A}}\text{C ACCUG}^{3'} \\ {}^{3'}\text{C A}\underline{\text{G}}\text{G UGGAC}^{5'} \end{bmatrix} = \Delta G^\circ_{37,\,\text{i}} + \Delta G^\circ_{37,\,\text{single mismatch}} + \Delta G^\circ_{37}\begin{bmatrix}\text{GU}\\\text{CA}\end{bmatrix} + \Delta G^\circ_{37}\begin{bmatrix}\text{CA}\\\text{GU}\end{bmatrix} + \Delta G^\circ_{37}\begin{bmatrix}\text{AC}\\\text{UG}\end{bmatrix} + \Delta G^\circ_{37}\begin{bmatrix}\text{CC}\\\text{GG}\end{bmatrix} + \Delta G^\circ_{37}\begin{bmatrix}\text{CU}\\\text{GA}\end{bmatrix} + \Delta G^\circ_{37}\begin{bmatrix}\text{UG}\\\text{AC}\end{bmatrix} \quad (2)$$

where $\Delta G^\circ_{37,\text{i}}$ is the free energy change for duplex initiation [4.09 kcal/mol (*47*)], $\Delta G^\circ_{37,\text{single mismatch}}$ is the free energy contribution from the single mismatch, and the remainder of the terms are individual nearest neighbor values (*47*). Therefore, rearranging eq 2 can solve for the contribution of the single mismatch to duplex stability:

$$\Delta G^\circ_{37,\,\text{single mismatch}} = G^\circ_{37}\begin{bmatrix} {}^{5'}\text{G U}\underline{\text{A}}\text{C ACCUG}^{3'} \\ {}^{3'}\text{C A}\underline{\text{G}}\text{G UGGAC}^{5'} \end{bmatrix} - \Delta G^\circ_{37,\,\text{i}} - \Delta G^\circ_{37}\begin{bmatrix}\text{GU}\\\text{CA}\end{bmatrix} - \Delta G^\circ_{37}\begin{bmatrix}\text{CA}\\\text{GU}\end{bmatrix} - \Delta G^\circ_{37}\begin{bmatrix}\text{AC}\\\text{UG}\end{bmatrix} - \Delta G^\circ_{37}\begin{bmatrix}\text{CC}\\\text{GG}\end{bmatrix} - \Delta G^\circ_{37}\begin{bmatrix}\text{CU}\\\text{GA}\end{bmatrix} - \Delta G^\circ_{37}\begin{bmatrix}\text{UG}\\\text{AC}\end{bmatrix} \quad (3)$$

where $\Delta G^\circ_{37}\left[\begin{smallmatrix} {}^{5'}\text{G U}\underline{\text{A}}\text{C ACCUG}^{3'} \\ {}^{3'}\text{C A}\underline{\text{G}}\text{G UGGAC}^{5'} \end{smallmatrix}\right]$ is the value determined by optical melting experiments, $\Delta G^\circ_{37,\text{i}}$ is the free energy change for duplex initiation [4.09 kcal/mol (*47*)], and $\Delta G^\circ_{37,\text{single mismatch}}$ is the free energy contribution of the mismatch. More explicitly

$$\Delta G^\circ_{37,\,\text{single mismatch}} = -9.74 - 4.09 - (-2.24) - (-2.11) - (-2.24) - (-3.26) - (-2.08) - (-2.11) = 0.21 \text{ kcal/mol} \quad (4)$$

A second example of this type of calculation in which the same single mismatch−nearest neighbor sequence combination is placed in the center of the duplex is as follows:

$$\Delta G^\circ_{37}\begin{bmatrix} {}^{5'}\text{GAC U}\underline{\text{A}}\text{C CUG}^{3'} \\ {}^{3'}\text{CUG A}\underline{\text{G}}\text{G GAC}^{5'} \end{bmatrix} = \Delta G^\circ_{37,\,\text{i}} + \Delta G^\circ_{37}\begin{bmatrix}\text{GA}\\\text{CU}\end{bmatrix} + \Delta G^\circ_{37}\begin{bmatrix}\text{AC}\\\text{UG}\end{bmatrix} + \Delta G^\circ_{37}\begin{bmatrix}\text{CU}\\\text{GA}\end{bmatrix} + \Delta G^\circ_{37,\,\text{single mismatch}} + \Delta G^\circ_{37}\begin{bmatrix}\text{CC}\\\text{GG}\end{bmatrix} + \Delta G^\circ_{37}\begin{bmatrix}\text{CU}\\\text{GA}\end{bmatrix} + \Delta G^\circ_{37}\begin{bmatrix}\text{UG}\\\text{AC}\end{bmatrix} \quad (5)$$

$$\Delta G^\circ_{37,\,\text{single mismatch}} = \Delta G^\circ_{37}\begin{bmatrix} {}^{5'}\text{GAC U}\underline{\text{A}}\text{C CUG}^{3'} \\ {}^{3'}\text{CUG A}\underline{\text{G}}\text{G GAC}^{5'} \end{bmatrix} - \Delta G^\circ_{37,\,\text{i}} - \Delta G^\circ_{37}\begin{bmatrix}\text{GA}\\\text{CU}\end{bmatrix} - \Delta G^\circ_{37}\begin{bmatrix}\text{AC}\\\text{UG}\end{bmatrix} - \Delta G^\circ_{37}\begin{bmatrix}\text{CU}\\\text{GA}\end{bmatrix} - \Delta G^\circ_{37}\begin{bmatrix}\text{CC}\\\text{GG}\end{bmatrix} - \Delta G^\circ_{37}\begin{bmatrix}\text{CU}\\\text{GA}\end{bmatrix} - \Delta G^\circ_{37}\begin{bmatrix}\text{UG}\\\text{AC}\end{bmatrix} \quad (6)$$

$$\Delta G^\circ_{37,\,\text{single mismatch}} = -10.67 - 4.09 - (-2.35) - (-2.24) - (-2.08) - (-3.26) - (-2.08) - (-2.11) = -0.64 \text{ kcal/mol} \quad (7)$$

It is important to note that in these examples, the stem sequence remains constant. However, when the single mismatch and nearest neighbors are moved from the 5′-shifted position to the central position, some of the individual nearest neighbor combinations within the stem change. This change in nearest neighbor combinations in the stem is accounted for by subtraction of the free energy contribution of each nearest neighbor combination from the raw data for the entire duplex and calculation of the single mismatch free energy contribution. Errors in these single mismatch contributions (Table 2) were propagated from the errors for the measured duplex (obtained from the analysis of the T_M dependence of the melting curves) (Table 1) and the errors reported for the nearest neighbor parameters (*47*).

RESULTS

Confirmation of Single Mismatch Formation by NMR. Five representative duplexes were studied by NMR. The thermodynamics of the first duplex, $\left[\begin{smallmatrix} {}^{5'}\text{GAC U}\underline{\text{U}}\text{G CUG}^{3'} \\ {}^{3'}\text{CUG G}\underline{\text{U}}\text{C GAC}^{5'} \end{smallmatrix}\right]$, were studied previously (*35*), but the data were not used in the previous study to determine averages, trends, etc., because of the possible formation of a competing structure. The NMR data collected here confirm the presence of a competing structure. The imino proton region of the NMR spectrum contains more resonances (at least 14) than expected (11, one from each Watson−Crick pair, two from the G-U pair, and two from the two uracils in the mismatch) if the duplex containing the single mismatch was the sole conformation

Table 1: Thermodynamic Parameters for Duplex Formation[a]

	analysis of melt curve fit/errors				analysis of T_M dependence/ errors (ln plot)			
sequence[b]	ΔH° (kcal/mol)	ΔS° (cal/K·mol)	ΔG°$_{37}$ (kcal/mol)	T_M[c] (°C)	ΔH° (kcal/mol)	ΔS° (cal/K·mol)	ΔG°$_{37}$ (kcal/mol)	T_M[c] (°C)
G UAC ACCUG C AGG UGGAC	-63.7 ± 6.5	-173.8 ± 20.0	-9.77 ± 0.38	53.6	-64.6 ± 2.2	-177.0 ± 6.6	-9.74 ± 0.10	53.2
GAC UAC CUG[d] CUG AGG GAC	-82.7 ± 4.7	-232.9 ± 14.5	-10.46 ± 0.26	52.5	-88.8 ± 10.1	-251.8 ± 31.3	-10.67 ± 0.46	52.2
GACCU UAC G CUGGA AGG C	-79.0 ± 12.1	-224.6 ± 38.2	-9.29 ± 0.29	48.3	-77.3 ± 2.3	-219.2 ± 7.2	-9.29 ± 0.07	48.5
G CAC ACCUG C GGG UGGAC	-81.7 ± 6.4	-226.8 ± 19.4	-11.33 ± 0.36	56.4	-82.1 ± 3.5	-228.2 ± 10.8	-11.32 ± 0.20	56.2
GAC CAC CUG[d,e] CUG GGG GAC	(-80.9)	(-225.3)	(-11.00)	(55.1)	(-215.3)	(-638.3)	(-17.30)	(53.3)
GACCU CAC G[e] CUGGA GGG C	(-72.8)	(-198.1)	(-11.40)	(59.2)	(-82.0)	(-226.2)	(-11.88)	(58.6)
G UAG ACCUG C AGC UGGAC	-70.5 ± 4.6	-191.7 ± 13.6	-10.99 ± 0.35	57.9	-70.2 ± 2.3	-190.9 ± 6.9	-10.98 ± 0.13	58.0
GAC UAG CUG[d] CUG AGC GAC	-76.4 ± 4.5	-217.4 ± 13.3	-8.95 = 0.47	47.1	-76.8 ± 11.1	-218.9 ± 34.6	-8.93 ± 0.56	47.0
GACCU UAG G CUGGA AGC C	-69.7 ± 5.2	-195.2 ± 16.0	-9.13 ± 0.23	49.0	-69.3 ± 4.2	-194.1 ± 13.2	-9.13 ± 0.14	49.1
G UAU ACCUG C AGA UGGAC	-63.8 ± 5.0	-180.5 ± 16.3	-7.80 ± 0.11	43.3	-63.4 ± 2.8	-179.3 ± 8.7	-7.78 ± 0.05	43.2
GAC UAU CUG[d] CUG AGA GAC	-79.5 ± 7.8	-232.7 ± 24.6	-7.33 = 0.22	40.1	-67.6 ± 3.9	-194.9 ± 12.7	-7.20 ± 0.06	40.1
GACCU UAU G CUGGA AGA C	-61.0 ± 10.0	-169.8 ± 31.1	-8.39 ± 0.37	46.8	-63.5 ± 5.4	-177.7 ± 16.9	-8.39 ± 0.19	46.4
G UUG ACCUG C GUC UGGAC	-68.1 ± 4.7	-187.1 ± 14.4	-10.02 ± 0.24	53.7	-68.8 ± 1.4	-189.4 ± 4.4	-10.02 ± 0.07	53.6
GAC UUG CUG[d,f,g] CUG GUC GAC	-94.9 ± 7.3	-263.7 ± 22.0	-13.12 ± 0.51	60.1	-94.5 ± 7.2	-262.5 ± 21.7	-13.04 ± 0.48	60.0
GACCU UUG G CUGGA GUC C	-74.6 ± 2.6	-211.9 ± 8.3	-8.86 ± 0.06	47.0	-75.1 ± 1.0	-213.6 ± 3.0	-8.86 ± 0.03	46.9
G AUC ACCUG[f] C UUG UGGAC	-82.0 ± 6.1	-229.6 ± 18.7	-10.72 ± 0.30	53.7	-82.4 ± 0.7	-231.1 ± 2.2	-10.71 ± 0.03	53.6
GAC AUC CUG[d] CUG UUG GAC	-87.4 ± 2.3	-250.2 ± 7.1	-9.84 = 0.10	49.2	-84.1 ± 1.6	-239.9 ± 5.0	-9.73 ± 0.06	49.2
GACCU AUC G CUGGA UUG C	-76.1 ± 7.2	-215.2 ± 22.1	-9.38 ± 0.33	49.1	-74.5 ± 2.3	-210.0 ± 7.0	-9.33 ± 0.08	49.1
G GCU ACCUG[f] C CAA UGGAC	-76.6 ± 5.6	-216.4 ± 17.7	-9.50 ± 0.17	49.5	-76.8 ± 2.1	-217.0 ± 6.6	-9.49 ± 0.08	49.4
GAC GCU CUG[d] CUG CAA GAC	-84.6 ± 12.8	-242.7 ± 39.7	-9.29 ± 0.49	47.4	-83.9 ± 1.6	-240.8 ± 5.0	-9.23 ± 0.05	47.3
GACCU GCU G[f] CUGGA CAA C	-70.8 ± 6.4	-197.4 ± 20.2	-9.54 ± 0.19	50.8	-72.3 ± 3.3	-202.4 ± 10.2	-9.56 ± 0.12	50.6
G CAG ACCUG C GCC UGGAC	-74.3 ± 5.6	-204.3 ± 17.2	-10.91 ± 0.30	56.4	-76.5 ± 2.3	-211.1 ± 7.2	-11.00 ± 0.12	56.3
GAC CAG CUG[d] CUG GCC GAC	-67.3 = 13.8	-182.0 ± 42.8	-10.82 ± 0.56	58.1	-66.8 ± 8.2	-180.3 ± 24.9	-10.84 ± 0.54	58.4
GAC CAG CUG[d] CUG GCC GAC	-74.5 ± 13.5	-203.9 ± 40.9	-11.24 ± 0.85	57.9	-76.4 ± 3.3	-210.0 ± 9.9	-11.26 ± 0.19	57.5
GACCU CAG G CUGGA GCC C	-74.6 ± 10.4	-203.4 ± 31.6	-11.50 ± 0.65	59.1	-75.7 ± 4.2	-206.9 ± 12.7	-11.50 ± 0.27	58.8
G AAG ACCUG C UCC UGGAC	-66.3 ± 7.3	-182.7 ± 22.2	-9.64 ± 0.42	52.2	-66.4 ± 2.4	-183.1 ± 7.5	-9.65 ± 0.11	52.3
GAC AAG CUG[d] CUG UCC GAC	-71.2 ± 4.2	-201.2 ± 13.1	-8.77 ± 0.20	47.1	-69.5 ± 2.8	-196.0 ± 8.9	-8.71 ± 0.08	47.0
G UCU ACCUG C AUA UGGAC	-64.5 ± 2.2	-183.4 ± 7.0	-7.57 ± 0.06	42.1	-64.4 ± 1.8	-183.3 ± 5.6	-7.59 ± 0.02	42.2
GAC UCU CUG[d] CUG AUA GAC	-78.4 ± 8.4	-230.6 ± 26.6	-6.90 ± 0.23	38.5	-66.6 ± 5.2	-192.5 ± 16.9	-6.86 ± 0.11	38.5
GACCU UCU G CUGGA AUA C	-57.3 ± 7.8	-158.9 ± 24.6	-7.97 ± 0.21	45.0	-57.2 ± 2.5	-158.6 ± 7.9	-7.97 ± 0.06	45.0
G AAG ACCUG[f] C UAC UGGAC	-60.0 ± 3.2	-162.7 ± 9.8	-9.57 ± 0.15	53.5	-60.6 ± 1.8	-164.5 ± 5.6	-9.57 ± 0.08	53.4
GAC AAG CUG CUG UAC GAC	-76.1 ± 6.4	-218.0 ± 20.2	-8.45 ± 0.12	45.0	-79.2 ± 1.2	-228.0 ± 3.9	-8.45 ± 0.02	44.7
GACCU AAG G CUGGA UAC C	-68.7 ± 2.3	-193.1 ± 7.2	-8.82 ± 0.06	47.7	-68.6 ± 1.5	-192.6 ± 4.6	-8.82 ± 0.04	47.7
G AGG ACCUG C UGC UGGAC	-66.7 ± 7.4	-181.4 ± 22.7	-10.44 ± 0.42	56.3	-67.8 ± 4.8	-184.8 ± 14.6	-10.44 ± 0.28	56.0
GAC AGG CUG[d] CUG UGC GAC	-84.0 ± 5.0	-237.9 ± 15.3	-10.22 ± 0.32	51.2	-85.1 ± 5.4	-241.4 ± 16.6	-10.23 ± 0.23	51.1
GACCU AGG G CUGGA UGC C	-57.6 ± 12.8	-152.8 ± 39.3	-10.16 ± 0.63	57.9	-57.6 ± 2.4	-153.0 ± 7.4	-10.16 ± 0.14	57.8
G ACG ACCUG C UCC UGGAC	-66.0 ± 3.0	-181.2 ± 9.2	-9.82 ± 0.13	53.2	-66.7 ± 1.6	-183.4 ± 4.8	-9.82 ± 0.07	53.1
GAC ACG CUG CUG UCC GAC	-71.7 ± 4.6	-205.4 ± 14.5	-7.98 ± 0.09	43.4	-74.7 ± 1.5	-215.0 ± 4.7	-7.98 ± 0.02	43.2

[a]Measurements were taken in 1.0 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na_2EDTA (pH 7.0). Significant figures beyond error estimates are given to allow accurate calculation of T_M and other parameters. [b]The single mismatch is identified by bold letters. The nearest neighbors and the mismatch are set apart for easy identification. The top strand of each duplex is written 5′ to 3′ and each bottom strand 3′ to 5′. [c]Calculated at an oligomer concentration of 10^{-4} M. [d]From ref *35*. [e]Data derived from non-two-state melts and not included in trends and averages. [f]Duplexes investigated by one-dimensional NMR spectroscopy. [g]Duplex not included in averages and trends because a bimolecular association of one of the strands with itself may be a competing structure.

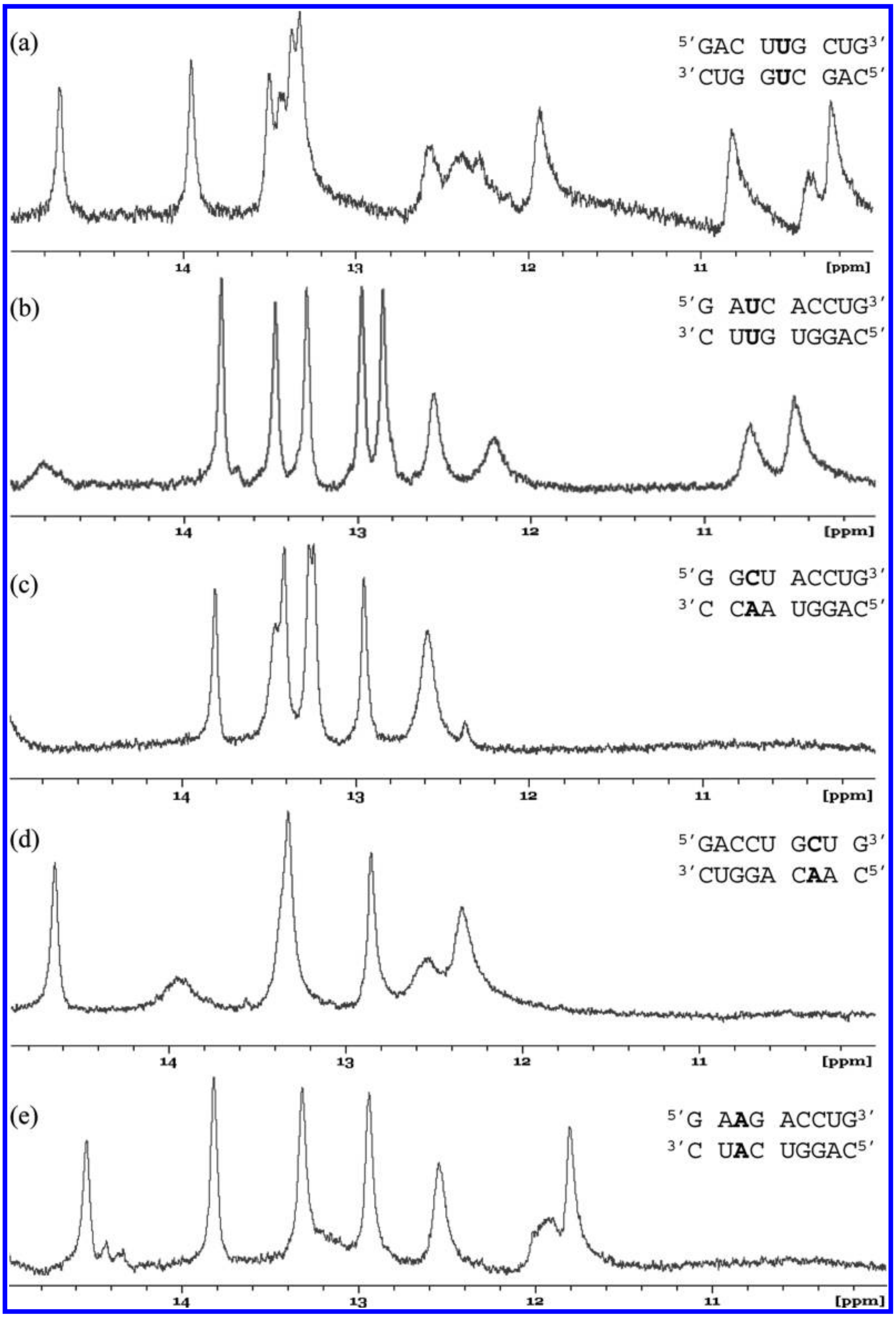

FIGURE 2: Imino proton region of the one-dimensional NMR spectra of duplexes studied here. (a) This spectrum suggests the duplex containing the single mismatch was not the sole conformation in solution. (b−e) These spectra are suggestive of single mismatch formation. Spectra a−d were recorded at 10 °C, and spectrum e was recorded at 0 °C.

in solution (Figure 2a). The spectra for the other four duplexes studied, however, are suggestive of single mismatch formation (Figure 2b−e).

For $\left[\begin{smallmatrix}5'\text{G A}\underline{\text{U}}\text{C ACCUG}3'\\ 3'\text{C U}\underline{\text{U}}\text{G UGGAC}5'\end{smallmatrix}\right]$, eight hydrogen-bonded imino resonances are expected and all eight are observed. In addition, two upfield imino resonances from the two uracils of the mismatch are also expected, and both are observed (Figure 2b). For $\left[\begin{smallmatrix}5'\text{G G}\underline{\text{C}}\text{U ACCUG}3'\\ 3'\text{C C}\underline{\text{A}}\text{A UGGAC}5'\end{smallmatrix}\right]$, eight hydrogen bonded imino resonances are expected and all eight are observed, with two resonances overlapping at 12.6 ppm. No imino resonances are expected from the A·C mismatch, and none are observed (Figure 2c). For $\left[\begin{smallmatrix}5'\text{GACCU G}\underline{\text{C}}\text{U G}3'\\ 3'\text{CUGGA C}\underline{\text{A}}\text{A C}5'\end{smallmatrix}\right]$, eight hydrogen-bonded imino resonances are expected. Only seven are observed (with two overlapping at 13.3 ppm). It is likely one of the terminal imino protons is exchanging rapidly with the solvent, and this resonance has broadened into the baseline (Figure 2d). Similarly, eight hydrogen-bonded imino resonances are expected for $\left[\begin{smallmatrix}5'\text{G A}\underline{\text{A}}\text{G ACCUG}3'\\ 3'\text{C U}\underline{\text{A}}\text{C UGGAC}5'\end{smallmatrix}\right]$; however, only seven are observed. Again, it is likely one of the terminal imino

Table 2: Contributions of 13 Single Mismatch–Nearest Neighbor Sequence Combinations to Duplex Thermodynamics[a]

		$\Delta H°_{single\ mismatch}$ (kcal/mol)			$\Delta S°_{single\ mismatch}$ (cal/K·mol)			$\Delta G°_{37,\ single\ mismatch}$ (kcal/mol)		
sequence[b]	mismatch[c]	predicted[d]	measured[e]	abs. diff.[f]	predicted[d]	measured[e]	abs. diff.[f]	predicted[d]	measured[e]	abs. diff.[f]
G UAC ACCUG C AGG UGGAC			-0.7	4.1		-2.9	22.5		0.21	0.6
GAC UAC CUG[g] CUG AGG GAC	UAC AGG	-4.8	-23.8	19.0	-25.4	-74.5	49.1	0.8	-0.64	1.4
GACCU UAC G CUGGA AGG C			-15.7	10.9		-50.2	24.8		-0.16	1.0
G CAC ACCUG C GGG UGGAC			-14.7	13.9		-46.7	36.3		-0.19	0.1
GAC CAC CUG[g,h] CUG GGG GAC	CAC GGG	-0.8	(-13.0)	(12.2)	-10.4	(-42.4)	(32.0)	-0.3	(0.21)	(0.5)
GACCU CAC G[h] CUGGA GGG C			(-5.6)	(4.8)		(-15.6)	(5.2)		(-0.85)	(0.6)
G UAG ACCUG C AGC UGGAC			-4.3	0.5		-11.2	8.0		-0.79	2.2
GAC UAG CUG[g] CUG AGC GAC	UAG AGC	-4.8	-10.3	5.5	-19.2	-37.4	18.2	1.4	1.26	0.1
GACCU UAG G CUGGA AGC C			-5.0	0.2		-19.1	0.1		0.90	0.5
G UAU ACCUG C AGA UGGAC			-2.2	6.6		-11.6	28.8		1.39	0.5
GAC UAU CUG[g] CUG AGA GAC	UAU AGA	-8.8	-3.5	5.3	-40.4	-17.8	22.6	1.9	1.92	0.0
GACCU UAU G CUGGA AGA C			-2.1	6.7		-8.5	31.9		0.49	1.4
G UUG ACCUG C GUC UGGAC			-1.7	14.6		-6.7	51.8		0.44	1.0
GAC UUG CUG[g,i] CUG GUC GAC	UUG GUC	-16.3	(-26.4)	(10.1)	-58.5	(-75.9)	(17.4)	1.4	(-2.82)	(4.2)
GACCU UUG G CUGGA GUC C			-4.8	11.5		-20.3	38.2		1.51	0.1
G AUC ACCUG C UUG UGGAC			-17.4	1.8		-54.0	7.9		-0.65	1.2
GAC AUC CUG[g] CUG UUG GAC	AUC UUG	-19.2	-19.1	0.1	-61.9	-62.8	0.9	0.5	0.33	0.2
GACCU AUC G CUGGA UUG C			-12.1	7.1		-39.5	22.4		0.20	0.3
G GCU ACCUG C CAA UGGAC			-13.6	7.7		-46.1	6.6		0.70	0.5
GAC GCU CUG[g] CUG CAA GAC	GCU CAA	-21.3	-19.7	1.6	-39.5	-64.1	24.6	0.2	0.17	0.0
GACCU GCU G CUGGA CAA C			-7.3	14.0		-25.3	14.2		0.50	0.3
G CAG ACCUG C GCC UGGAC			-7.1	7.1		-24.0	24.0		0.37	0.4
GAC CAG CUG[g,j] CUG GCC GAC	CAG GCC	0.0	-2.2	2.2	0.0	-8.1	8.1	0.0	0.32	0.3
GACCU CAG G CUGGA GCC C			-5.8	5.8		-18.4	18.4		-0.05	0.1
G AAG ACCUG C UCC UGGAC	AAG UCC	-4.0	0.6	4.6	-15.0	-0.4	14.6	1.1	0.65	0.5
GAC AAG CUG[g] CUG UCC GAC			-3.0	1.0		-14.7	0.3		1.51	0.4
G UCU ACCUG C AUA UGGAC			-3.2	4.8		-15.6	14.4		1.58	0.6
GAC UCU CUG[g] CUG AUA GAC	UCU AUA	-8.0	-2.5	5.5	-30.0	-15.4	14.6	2.2	2.26	0.1
GACCU UCU G CUGGA AUA C			4.2	12.2		10.6	40.6		0.91	1.3
G AAG ACCUG C UAC UGGAC			6.4	10.4		18.2	33.2		0.73	0.4
GAC AAG CUG CUG UAC GAC	AAG UAC	-4.0	-12.7	8.7	-15.0	-46.7	31.7	1.1	1.77	0.7
GACCU AAG G CUGGA UAC C			-3.4	0.6		-16.1	1.1		1.61	0.5
G AGG ACCUG C UGC UGGAC			-0.8	4.1		-2.1	65.1		-0.14	0.9
GAC AGG CUG[g] CUG UGC GAC	AGG UGC	-4.9	-18.6	13.7	-67.2	-60.1	7.1	-1.0	-0.01	1.0
GACCU AGG G CUGGA UGC C			7.6	12.5		23.5	90.7		0.27	1.3
G ACG ACCUG C UCC UGGAC	ACG UCC	-4.0	0.3	4.3	-15.0	-0.7	14.3	1.1	0.48	0.6
GAC ACG CUG CUG UCC GAC			-8.2	4.2		-33.7	18.7		2.24	1.1

[a]Calculations were based on the data obtained from T_M^{-1} vs $\ln(C_T/4)$ plots. Values in parentheses may not be accurate due to non-two-state melting, or a bimolecular association of one of the strands with itself may be a competing structure. [b]The single mismatch is identified by bold letters. The mismatch and nearest neighbors are set apart for easy identification. The top strand of each duplexis written 5′ to 3′ and each bottom strand 3′ to 5′. [c]The mismatch and nearest neighbors common for each set of duplexes are indicated and written as described in footnote b . [d]Values predicted by the single mismatch specific model published by Davis and Znosko (*36*). This value is dependent of the identity of the mismatch and nearest neighbors but is independent of mismatch duplex position. [e]Measured values were calculated by subtracting the nearest neighbor contribution for the canonical base pairs (*67*) from the optical melting data resulting from duplex formation. Errors associated with these values are approximately ±6.3 kcal/mol, ±19.4 cal/K·mol, and ±0.31 kcal/mol for $\Delta H°_{single\ mismatch}$, $\Delta S°_{single\ mismatch}$, and $\Delta G°_{37,single\ mismatch}$, respectively. [f]Absolute differences between the measured and predicted values. [g]From ref *35*. [h]Data derived from non-two-state melts and not included in trends or averages. [i]Duplex not included in trends and averages because a bimolecular association of one of the strands with itself may be a competing structure. [j]The duplex sequence was measured twice, and the resulting thermodynamic parameters were averaged.

Table 3: Averages and Ranges of Thermodynamic Parameters for Single Mismatch−Nearest Neighbor Sequence Combinations[a]

sm position[b]		$\Delta H°_{SM}$ (kcal/mol)	$\Delta S°_{SM}$ (cal/K·mol)	$\Delta G°_{37,SM}$ (kcal/mol)
center	average	−12.5	−43.0	0.82
	range	−23.8 to −2.2	−74.5 to 8.1	−0.64 to 2.26
off-center[c]	average	−4.6	−16.3	0.45
	range	−17.4 to 7.6	−54.0 to 23.5	−0.79 to 1.61
5′-shifted	average	−4.1	−14.3	0.36
	range	−17.4 to 6.4	−54.0 to 18.2	−0.79 to 1.58
3′-shifted	average	−5.1	−18.2	0.54
	range	−15.7 to 7.6	−50.2 to 23.5	−0.16 to 1.61

[a]Averages and ranges are based on the data obtained from T_M^{-1} vs $\ln(C_T/4)$ plots of the nine complete sets of data as described in Materials and Methods. Errors associated with the individual $\Delta H°_{SM}$, $\Delta S°_{SM}$, and $\Delta G°_{37,SM}$ values used to calculate the average values listed here are approximately ±6.3 kcal/mol, ±19.4 cal/K·mol, and ±0.31 kcal/mol, respectively. [b]The duplex position of the single mismatch as described in Materials and Methods. [c]The off-center values are an average of the 5′- and 3′-shifted single mismatch data.

Table 4: Averages of the Absolute Difference between Predicted and Measured Single Mismatch Thermodynamic Parameters[a]

sm position[b]		$\Delta\Delta H°_{SM}$ (kcal/mol)	$\Delta\Delta S°_{SM}$ (cal/K·mol)	$\Delta\Delta G°_{37,SM}$ (kcal/mol)
center	average[c]	6.8	19.7	0.43
	standard deviation	6.1	14.6	0.50
off-center[d]	average[c]	6.5	25.3	0.76
	standard deviation	4.3	22.6	0.53
5′-shifted	average[c]	5.2	23.4	0.80
	standard deviation	3.0	18.4	0.58
3′-shifted	average[c]	7.8	27.1	0.73
	standard deviation	5.1	27.2	0.51

[a]Averages and standard deviations are based on the data obtained from T_M^{-1} vs $\ln(C_T/4)$ plots of the nine complete sets of data as described in Materials and Methods. [b]The duplex position of the single mismatch as described in Materials and Methods. [c]The average of the absolute difference between the predicted and measured thermodynmic values. Predicted values are calculated using the single mismatch specific algorithm (*36*). [d]The off-center values are an average of the 5′- and 3′-shifted single mismatch data.

protons is exchanging rapidly with the solvent, and this resonance has broadened into the baseline (Figure 2e). The number of imino proton resonances in these spectra suggests the duplex with the single mismatch is the predominate structure in solution.

Thermodynamic Parameters. The thermodynamic parameters for duplex formation, which were obtained from fitting each melting curve to the two-state model and from the van't Hoff plot of T_M^{-1} versus $\log(C_T/4)$, are listed in Table 1. Data for 38 duplexes containing 13 single mismatch−nearest neighbor sequence combinations are shown because most combinations were melted at three duplex positions. One central single mismatch nearest−neighbor combination, $\left[\begin{smallmatrix}\text{5'C}\underline{\text{A}}\text{G3'}\\ \text{3'G}\underline{\text{C}}\text{C5'}\end{smallmatrix}\right]$, with the same stem was studied twice by melting the same duplex sequence from two separate samples. Two duplexes, $\left[\begin{smallmatrix}\text{5'GAC C}\underline{\text{A}}\text{C CUG3'}\\ \text{3'CUG G}\underline{\text{G}}\text{G GAC5'}\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}\text{5'GACCU C}\underline{\text{A}}\text{C G3'}\\ \text{3'CUGGA G}\underline{\text{G}}\text{GC5'}\end{smallmatrix}\right]$, melted in a non-two-state manner. Non-two-state melting was assessed when the enthalpy values resulting from the two methods used to analyze the melting curves did not agree within 10% (*48*, *49*). It is interesting to note both of these duplexes contain the same single mismatch−nearest neighbor combination but at different duplex positions. The non-two-state melting observed here may be due to the formation of a guanine tetraplex or aggregation, which is a result of having three or more consecutive guanine residues (*50*). The melt transitions of all other duplexes are most likely two-state. Two combinations, $\left[\begin{smallmatrix}\text{5'A}\underline{\text{A}}\text{G3'}\\ \text{3'U}\underline{\text{C}}\text{C5'}\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}\text{5'A}\underline{\text{C}}\text{G3'}\\ \text{3'U}\underline{\text{C}}\text{C5'}\end{smallmatrix}\right]$, were only studied at the 5′-shifted and central positions because if the mismatch was placed at the 3′-shifted position, multiple structures were likely to compete with the formation of the single mismatch structure (*28*−*30*). Lastly, as suggested by NMR data, $\left[\begin{smallmatrix}\text{5'CAG C}\underline{\text{U}}\text{G GUC3'}\\ \text{3'GUC G}\underline{\text{U}}\text{U CAG5'}\end{smallmatrix}\right]$ may not be the only conformation in solution. Perhaps the bimolecular association of the top strand with itself is a competing structure. The resulting data from those sequences that melted in a non-two-state manner, $\left[\begin{smallmatrix}\text{5'GAC C}\underline{\text{A}}\text{C CUG3'}\\ \text{3'CUG G}\underline{\text{G}}\text{G GAC5'}\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}\text{5'GACCU C}\underline{\text{A}}\text{C G3'}\\ \text{3'CUGGA G}\underline{\text{G}}\text{G C5'}\end{smallmatrix}\right]$, and the duplex sequence possibly forming multiple conformations, $\left[\begin{smallmatrix}\text{5'CAG C}\underline{\text{U}}\text{G GUC3'}\\ \text{3'GUC G}\underline{\text{U}}\text{U CAG5'}\end{smallmatrix}\right]$, were not included in trends or averages and are denoted in Table 1. Taken together, there are four single mismatch−nearest neighbor combinations that do not provide viable thermodynamic data at each duplex position and, therefore, are not included trends or averages. Consequently, nine of the 13 single mismatch−nearest neighbor sequence combinations investigated here have viable thermodynamic data at each of the three duplex positions and are used to determine trends and averages.

Contribution of Single Mismatches to Duplex Thermodynamics. The contributions of the 13 single mismatch−nearest neighbor sequence combinations to duplex stability at the three duplex positions are listed in Table 2. For the nine complete sets, single mismatches placed at the 5′-shifted, central, and 3′-shifted positions contribute an average of 0.4 (range of −0.8 to 1.6 kcal/mol), 0.8 (range of −0.6 to 2.3 kcal/mol), and 0.5 (range of −0.2 to 1.6 kcal/mol) kcal/mol to duplex stability, respectively (Table 3). The corresponding entropy and enthalpy averages and ranges are listed in Table 3. These experimental free energy values are compared to those obtained by a predictive model (*36*) (Tables 2 and 4), resulting in average absolute free energy differences of 0.8, 0.4, and 0.7 kcal/mol for the 5′-shifted, central, and 3′-shifted positions, respectively.

DISCUSSION

Kierzek and co-workers did examine the positional effects on the stability of three single mismatch types (*37*); however, their investigation and other previous thermodynamic studies have mainly focused their efforts on characterizing single mismatches placed at the center of an RNA duplex (*1*, *35*−*37*, *50*−*52*). However, the analysis of the secondary structures of rRNA and group I introns (*37*, *53*−*61*) reveals many single mismatches do not occur toward the center of the duplex but are preferentially found near the ends of duplex regions. Additionally, characterization of this small motif at various duplex positions may be beneficial in the rational design of several types of therapeutic agents, such as fork-siRNAs and shRNA, which have both been found to have enhanced RNAi activity when single mismatches are placed at the 3′ end of ss-RNA (*19*−*22*). However, algorithms used to predict RNA secondary structure from sequence assume the thermodynamic contribution of a single mismatch is independent of its position within the duplex and independent of its non-nearest neighbors (*28*−*31*). Thirteen single mismatch−nearest neighbor combinations have been thermodynamically characterized at three duplex positions, the center and two

Table 5: Comparison of the Free Energy Contributions of the 13 Single Mismatch−Nearest Neighbor Sequence Combinations at Three Helical Positions[a]

		$\Delta G°_{37}$ (kcal/mol)		
sequence[b]	mismatch[c]	measured[d]	$\lvert \Delta G_{end\,SM} - \Delta G_{center\,SM} \rvert$[e]	$\lvert \Delta G_{5'\,SM} - \Delta G_{3'\,SM} \rvert$[f]
G UAC ACCUG C AGG UGGAC		0.21	0.85	
GAC UAC CUG[g] CUG AGG GAC	UAC AGG	-0.64		0.37
GACCU UAC G CUGGA AGG C		-0.16	0.48	
G CAC ACCUG C GGG UGGAC		-0.19	(-0.40)	
GAC CAC CUG[g,h] CUG GGG GAC	CAC GGG	(0.21)		(0.66)
GACCU CAC G[h] CUGGA GGG C		(-0.85)	(-1.06)	
G UAG ACCUG C AGC UGGAC		-0.79	-2.05	
GAC UAG CUG[g] CUG AGC GAC	UAG AGC	1.26		-1.69
GACCU UAG G CUGGA AGC C		0.90	-0.36	
G UAU ACCUG C AGA UGGAC		1.39	-0.53	
GAC UAU CUG[g] CUG AGA GAC	UAU AGA	1.92		0.90
GACCU UAU G CUGGA AGA C		0.49	-1.43	
G UUG ACCUG C GUC UGGAC		0.44	(3.26)	
GAC UUG CUG[g,i] CUG GUC GAC	UUG GUC	(-2.82)		(-1.07)
GACCU UUG G CUGGA GUC C		1.51	(4.33)	
G AUC ACCUG C UUG UGGAC		-0.65	-0.98	
GAC AUC CUG[g] CUG UUG GAC	AUC UUG	0.33		-0.85
GACCU AUC G CUGGA UUG C		0.20	-0.13	
G GCU ACCUG C CAA UGGAC		0.70	0.53	
GAC GCU CUG[g] CUG CAA GAC	GCU CAA	0.17		0.20
GACCU GCU G CUGGA CAA C		0.50	0.33	
G CAG ACCUG C GCC UGGAC		0.37	0.05	
GAC CAG CUG[g,j] CUG GCC GAC	CAG GCC	0.32		0.42
GACCU CAG G CUGGA GCC C		-0.05	-0.37	
G AAG ACCUG C UCC UGGAC	AAG UCC	0.65	-0.86	
GAC AAG CUG[g] CUG UCC GAC		1.51		—
G UCU ACCUG C AUA UGGAC		1.58	-0.68	
GAC UCU CUG[g] CUG AUA GAC	UCU AUA	2.26		0.67
GACCU UCU G CUGGA AUA C		0.91	-1.35	
G AAG ACCUG C UAC UGGAC		0.73	-1.04	
GAC AAG CUG CUG UAC GAC	AAG UAC	1.77		-0.88
GACCU AAG G CUGGA UAC C		1.61	-0.16	
G AGG ACCUG C UGC UGGAC		-0.14	-0.13	
GAC AGG CUG[g] CUG UGC GAC	AGG UGC	-0.01		-0.41
GACCU AGG G CUGGA UGC C		0.27	0.28	
G ACG ACCUG C UCC UGGAC	ACG UCC	0.48	-1.76	
GAC ACG CUG CUG UCC GAC		2.24		—

[a]Calculations were based on the data obtained from T_M^{-1} vs $\ln(C_T/4)$ plots. Values in parentheses may not be accurate because of non-two-state melting, or a bimolecular association of one of the strands with itself may be a competing structure. [b]The single mismatch is identified by bold letters. The mismatch and nearest neighbors are set apart for easy identification. The top strand of each duplex is written 5′ to 3′ and each bottom strand 3′ to 5′. [c]The mismatch and nearest neighbors common for each set of duplexes are indicated and written as described in footnote b. [d]Measured values were calculated by subtracting the nearest neighbor contribution for the canonical base pairs (*67*) from the optical melting data resulting from duplex formation. [e]Difference in free energy contribution of the single mismatches at either the 5′- or 3′-shifted position and the center of the duplex. [f]Difference in single mismatch free energy contribution at the 5′- and 3′-shifted positions. [g]From ref *35*. [h]Data derived from non-two-state melts and not included in trends and averages. [i]Duplex not included in trends and averages because a bimolecular association of one of the strands with itself may be a competing structure. [j]The duplex sequence was measured twice, and the resulting thermodynamic parameters were averaged.

off-center positions (5′- and 3′-shifted). The resulting data are analyzed and compared to investigate the effects of duplex position and non-nearest neighbor identity on the thermodynamic contribution of the single mismatch to duplex stability.

Thermodynamic Contributions of Single Mismatches to Duplex Thermodynamics. Free energy minimization algorithms used to predict RNA secondary structure from sequence (*28−34*) utilize a measured value or an average of measured values if the thermodynamic parameters of a single mismatch have been experimentally determined. This study has thermodynamically characterized two previously unstudied single mismatch−nearest neighbor sequence combinations, $\left[\begin{smallmatrix}5'\mathrm{AAG}3'\\3'\mathrm{U\underline{A}C}5'\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}5'\mathrm{ACG}3'\\3'\mathrm{U\underline{C}C}5'\end{smallmatrix}\right]$, enabling the use of measured thermodynamic parameters instead of predictive values, which may help improve the accuracy of such predictive algorithms.

Assessment of the data in Tables 1 and 2 reveals a large variance in the obtained thermodynamic parameters. Table 3 compiles this data and shows, on average, the thermodynamic contributions of 5′-shifted single mismatches are relatively similar to the thermodynamic contributions of 3′-shifted single mismatches. Table 3 also shows, on average, the thermodynamic contributions of the off-center single mismatches are different from those of the central single mismatches and are less favorable enthalpically and more favorable in terms of both entropy and free energy.

Although Table 3 identifies these general trends, Table 5 shows there are idiosyncrasies associated with these general trends.

For example, Table 3 identifies the general trend in the similarity of the thermodynamic contributions of 5′- and 3′-shifted single mismatches. However, Table 5 shows the contribution of 5′-shifted single mismatches is not always comparable to that of the 3′-shifted single mismatches. For example, 5′-shifted $\left[\begin{smallmatrix}5'\text{U}\underline{\text{A}}\text{G}3'\\3'\text{A}\underline{\text{G}}\text{C}5'\end{smallmatrix}\right]$ is 1.7 kcal/mol more stable than when the same mismatch is 3′-shifted. On the other hand, 3′-shifted $\left[\begin{smallmatrix}5'\text{U}\underline{\text{A}}\text{U}3'\\3'\text{A}\underline{\text{G}}\text{A}5'\end{smallmatrix}\right]$ is 0.9 kcal/mol more stable than when the same mismatch is 5′-shifted. Table 3 also shows, on average, a central single mismatch is 0.4 kcal/mol less stable than an off-center single mismatch. However, individual examples in Table 5 reveal a central $\left[\begin{smallmatrix}5'\text{U}\underline{\text{A}}\text{G}3'\\3'\text{A}\underline{\text{G}}\text{C}5'\end{smallmatrix}\right]$ is *2.1* kcal/mol less stable than the same mismatch 5′-shifted, and a central $\left[\begin{smallmatrix}5'\text{G}\underline{\text{C}}\text{U}3'\\3'\text{C}\underline{\text{A}}\text{A}5'\end{smallmatrix}\right]$ is 0.5 kcal/mol *more* stable than the same mismatch 3′-shifted. In summary, Table 3 identifies some general trends associated with the effect of duplex position on the thermodynamic contribution of a single mismatch, but Table 5 reveals some idiosyncrasies that are unexpected on the basis of the general trends.

Effect of Single Mismatch Identity and Duplex Position on the Free Energy of Single Mismatches. Previous studies have found the stability of a single mismatch to be dependent upon the identity of the nucleotides involved in the mismatch and the duplex position (*1*, *35−37*). For example, U·U and A·A mismatches are found to be more stable when placed closer to the duplex terminus than when in the center of the duplex; however, G·G mismatches are found to be insensitive to positional effects (*37*), which is in accordance with the results found here (Tables 2 and 5).

To further compare these findings to the data presented here for the nine complete sets, the single mismatches were grouped by type of mismatch (data not shown), and the average free energies at each of the duplex positions were derived. The type of mismatch is defined by purine·purine (R·R, including A·G, G·G, and A·A), pyrimidine·pyrimidine (Y·Y, including C·C, C·U, and U·U), and purine·pyrimidine (R·Y, including A·C) single mismatches. For 5′-shifted single mismatches, average free energy values of 0.3, 0.5, and 0.4 kcal/mol were obtained for the R·R, Y·Y, and R·Y mismatches, respectively. For the centrally placed single mismatches, average free energy values of 0.9, 1.3, and 0.2 kcal/mol were obtained for R·R, Y·Y, and R·Y mismatches, respectively. For the 3′-shifted single mismatches, average free energy values of 0.6, 0.6, and 0.3 kcal/mol were obtained for the R·R, Y·Y, and R·Y mismatches, respectively.

Regardless of duplex position, Y·Y mismatches are on average the most destabilizing, while R·Y mismatches are on average the least destabilizing. Additionally, centrally placed R·R and Y·Y single mismatches are the most destabilizing to duplex thermodynamics, while centrally placed R·Y single mismatches are the least destabilizing to duplex thermodynamics. These results are in concordance with our initial hypotheses; R·Y mismatches would be the least destabilizing to duplex thermodynamics overall, and of the three positions studied, R·R and Y·Y single mismatches would be the most destabilizing in the center of the duplex. This can be explained by realizing R·Y mismatches are similar in size to a canonical base pair since they are comprised of one purine and one pyrimidine; therefore, R·Y single mismatches are not likely disrupting the duplex backbone. R·R and Y·Y single mismatches are likely to disrupt the duplex backbone by causing the backbone to bulge out and in, respectively, to accommodate the mismatched nucleotides; however, it is unclear why Y·Y single mismatches are more destabilizing than R·R single mismatches. It is likely the duplex can better accommodate single mismatches near the end of the duplex than in the center. These results suggest the thermodynamic stability of a single mismatch is dependent upon the identity of the mismatched nucleotides and duplex position.

Effect of Nearest Neighbor Identity on the Free Energy of Single Mismatches. It is interesting to note previous studies on various small RNA motifs, such as 1 × 2 (*39*, *43*, *62*), 1 × 3 (*39*), 2 × 3 (*39*), and 2 × 2 (*50*, *51*, *63−67*) centrally placed internal loops, have shown a thermodynamic dependence on the identity of the nearest neighbors. Specifically for single mismatches, previous thermodynamic investigations have demonstrated a correlation between the number of G-C base pairs adjacent to the single mismatch and the thermodynamic contribution of the single mismatch to duplex stability was identified (decreasing in thermodynamic stability: two G-C nearest neighbors > one G-C nearest neighbor > no G-C nearest neighbor) (*35*, *36*). A similar correlation is found for the single mismatches placed at each of the three duplex positions characterized in this work. These relationships are further defined in Table S1 of the Supporting Information. It is interesting to note the central single mismatches have the most unfavorable average free energy contribution, when compared to the average free energy values for the off-center positions (Table S1).

Kierzek and co-workers (*37*) investigated the thermodynamics of single mismatches and demonstrated the orientation of nearest neighbors can affect the thermodynamic contribution of the mismatch to duplex stability. Specifically, we compared the two nearest neighbors, $\left[\begin{smallmatrix}5'\text{G}\underline{\text{X}}\text{C}3'\\3'\text{C}\underline{\text{X}}\text{G}5'\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}5'\text{C}\underline{\text{X}}\text{G}3'\\3'\text{G}\underline{\text{X}}\text{C}5'\end{smallmatrix}\right]$, where the two X's are either both uracil (U) or both adenine (A) involved in a U·U or A·A single mismatch, respectively. For each case, the former set of nearest neighbors was found to have the most favorable free energy value. Via comparison of the two single mismatch−nearest neighbor combinations, $\left[\begin{smallmatrix}5'\text{U}\underline{\text{A}}\text{C}3'\\3'\text{A}\underline{\text{G}}\text{G}5'\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}5'\text{U}\underline{\text{A}}\text{G}3'\\3'\text{A}\underline{\text{G}}\text{C}5'\end{smallmatrix}\right]$, at each of the three duplex positions measured, on average the former is 0.7 kcal/mol more favorable.

Effect of Non-Nearest Neighbor Identity on the Thermodynamics of Single Mismatches. To further investigate the wide range of differences in single mismatch thermodynamics, we examined the free energies of the mismatch placed at the 5′- and 3′-shifted duplex positions. The average difference between the 5′- and 3′-shifted contribution is -0.14 ± 0.86 kcal/mol (Table 5); however, there are idiosyncrasies. For example, there is a -1.69 kcal/mol difference between the 5′- and 3′-shifted $\left[\begin{smallmatrix}5'\text{U}\underline{\text{A}}\text{G}3'\\3'\text{A}\underline{\text{G}}\text{C}5'\end{smallmatrix}\right]$ single mismatch. The only difference between $\left[\begin{smallmatrix}5'\text{U}\underline{\text{A}}\text{G}3'\\3'\text{A}\underline{\text{G}}\text{C}5'\end{smallmatrix}\right]$ at the 5′ position and $\left[\begin{smallmatrix}5'\text{U}\underline{\text{A}}\text{G}3'\\3'\text{A}\underline{\text{G}}\text{C}5'\end{smallmatrix}\right]$ at the 3′ position is the identity of the non-nearest neighbors, which suggests they are the origin of the observed idiosyncrasies between the same mismatch at these two duplex positions. However, the effect of non-nearest neighbors is not well understood and cannot be accounted for with the current size of the data set. Studies are currently underway to investigate this imperative research question.

Single Mismatch Specific Prediction Algorithm. The work recently published by Davis and Znosko (*35*, *36*) proposed a single mismatch specific algorithm for predicting the thermodynamic contribution to duplex stability. To allow for the comparison of the recently proposed predictive model (*35*, *36*) and the data obtained here (Table 2), the average absolute differences of the predicted and measured thermodynamic contributions of the nine complete sets of single mismatch-nearest neighbor sequence combinations are listed in Table 4. It is apparent centrally placed single mismatches are predicted most accurately, with a $\Delta\Delta G°_{37}$ of 0.4 kcal/mol, yet when considering the $\Delta\Delta G°_{37}$ values along with their standard deviations, 0.4 ± 0.5 kcal/mol for central single mismatches and 0.8 ± 0.5 kcal/mol for off-center single mismatches, it appears as if the previously proposed predictive model (*35*) works just as well for off-center as it does for central single mismatches. However, the data presented here suggest the addition of parameters that account for positional and/or non-nearest neighbor effects may improve prediction. A better understanding, along with more data, is required to accurately account for these observed effects in predictive models.

CONCLUSIONS

The effects of duplex position and identity of non-nearest neighbors were investigated for 13 single mismatch−nearest neighbor sequence combinations. Nine of these 13 single mismatches produced viable thermodynamic data at the three duplex positions studied, 5′-shifted, central, and 3′-shifted. It was found, on average, the thermodynamic contributions of 5′-shifted single mismatches are relatively equivalent to the thermodynamic contributions of 3′-shifted single mismatches. Additionally, on average, the thermodynamic contributions of the off-center single mismatches are quite different from those of the centrally placed single mismatches and are less favorable enthalpically and more favorable in both entropy and free energy. However, it is important to note there are several idiosyncrasies associated with these general trends upon comparison of the thermodynamic contributions of single mismatches on an individual basis. Overall, the stability of a single mismatch is dependent upon the identity of the mismatched nucleotides, the identity and orientation of the nearest neighbors, the identity of non-nearest neighbors, and the duplex position. The effects of non-nearest neighbors and duplex position are not fully understood, and work is currently underway to further investigate them.

SUPPORTING INFORMATION AVAILABLE

A table demonstrating the correlation between the number of G-C nearest neighbors and the free energy contribution of the single mismatch to duplex stability. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Peritz, A. E., Kierzek, R., Sugimoto, N., and Turner, D. H. (1991) Thermodynamic study of internal loops in oligoribonucleotides: Symmetric loops are more stable than asymmetric loops. *Biochemistry 30*, 6428–6436.
2. Calin-Jageman, I., and Nicholson, A. W. (2003) Mutational analysis of an RNA internal loop as a reactivity epitope for *Escherichia coli* ribonuclease III substrates. *Biochemistry 42*, 5025–5034.
3. Saito, H., and Richardson, C. C. (1981) Processing of mRNA by ribonuclease III regulates expression of gene 1.2 of bacteriophage T7. *Cell 27*, 533–542.
4. Du, T., and Zamore, P. D. (2005) MicroPrimer: The biogenesis and function of microRNA. *Development 132*, 4645–4652.
5. Bae, S. H., Cheong, H. K., Lee, J. H., Cheong, C., Kainosho, M., and Choi, B. S. (2001) Structural features of an influenza virus promoter and their implications for viral RNA synthesis. *Proc. Natl. Acad. Sci. U.S.A. 98*, 10602–10607.
6. Huthoff, H., and Berkhout, B. (2002) Multiple secondary structure rearrangements during HIV-1 RNA dimerization. *Biochemistry 41*, 10439–10445.
7. Schüler, M., Connell, S. R., Lescoute, A., Giesebrecht, J., Dabrowski, M., Schroeer, B., Mielke, T., Penczek, P. A., Westhof, E., and Spahn, C. M. T. (2006) Structure of the ribosome-bound cricket paralysis virus IRES RNA. *Nat. Struct. Mol. Biol. 13*, 1092–1096.
8. Wientges, J., Putz, J., Giege, R., Florentz, C., and Schwienhorst, A. (2000) Selection of viral RNA-derived tRNA-like structures with improved valylation activities. *Biochemistry 39*, 6207–6218.
9. Thunder, C., Witwer, C., Hofacker, I. L., and Stadler, P. F. (2004) Conserved RNA secondary structures in Flaviviridae genomes. *J. Gen. Virol. 85*, 1113–1124.
10. Shi, P.-Y., Brinton, M. A., Veal, J. M., Zhong, Y. Y., and Wilson, W. D. (1996) Evidence for the existence of a pseudoknot structure at the 3′ terminus of the Flavivirus genomic RNA. *Biochemistry 35*, 4222–4230.
11. Everett, C. M., and Wood, N. W. (2004) Trinucleotide repeats and neurodegenerative disease. *Brain 127*, 2385–2405.
12. Ranum, L. P. W., and Day, J. W. (2004) Myotonic dystrophy: RNA pathogenesis comes into focus. *Am. J. Hum. Genet. 74*, 793–804.
13. Tok, J. B.-H., Bi, L., and Saenz, M. (2005) Specific recognition of napthyridine-based ligands toward guanine-containing bulges in RNA duplexes and RNA−DNA heteroduplexes. *Bioorg. Med. Chem. Lett. 15*, 827–831.
14. Disney, M. D., Labuda, L. P., Paul, D. J., Poplawski, S. G., Pushechnikov, A., Tran, T., Velagapudi, S. P., Wu, M., and Childs-Disney, J. L. (2008) Two-dimensional combinational screening identifies specific aminoglycoside-RNA internal loop partners. *J. Am. Chem. Soc. 130*, 11185–11194.
15. Hobbie, S. N., Pfister, P., Brüll, C., Westhof, E., and Bröttger, E. C. (2005) Analysis of the contribution of individual subsitituents in 4,6-aminoglycoside-ribosome interactions. *Antimicrob. Agents Chemother. 49*, 5112–5118.
16. Vicens, Q., and Westhof, E. (2003) Crystal structure of Geneticin bound to a bacterial 16S ribosomal RNA A site oligonucleotide. *J. Mol. Biol. 326*, 1175–1188.
17. Hirao, I., Harada, Y., Nojima, T., Osawa, Y., Masaki, H., and Yokoyama, S. (2004) *In vitro* selection of RNA aptamers that bind to colicin E3 and structurally resemble the decoding site of 16S ribosomal RNA. *Biochemistry 43*, 3214–3221.
18. Rentmeister, A., Bill, A., Wahle, T., Walter, J., and Famulok, M. (2006) RNA aptamers selectively modulate protein recruitment to the cytoplasmic domain of β-secretase BACE1 in vitro. *RNA 12*, 1650–1660.
19. Hohjoh, H. (2004) Enhancement of RNAi activity by improved siRNA duplexes. *FEBS Lett. 557*, 193–198.
20. Schubert, S., Grünweller, A., Erdmann, V. A., and Kurreck, J. (2005) Local RNA target structure influences siRNA efficacy: Systematic analysis of intentionally designed binding regions. *J. Mol. Biol. 348*, 883–893.
21. Westerhout, E. M., and Berkhout, B. (2007) A systematic analysis of the effect of target RNA structure on RNA interference. *Nucleic Acids Res. 35*, 4322–4330.
22. Ohnishi, Y., Tokunaga, K., and Hohjoh, H. (2005) Influence of assembly of siRNA elements into RNA-induced silencing complex by fork-siRNA duplex carrying nucleotide mismatches at the 3′- or 5′-end of the sense-stranded siRNA element. *Biochem. Biophys. Res. Commun. 329*, 516–521.
23. Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature 404*, 293–296.
24. Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature 409*, 363–366.
25. Brummelkamp, T. R., Bernards, R., and Agami, R. (2002) A system for stable expression of short interfering RNas in mammalian cells. *Science 296*, 550–553.
26. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature 411*, 494–498.
27. Paul, C. P., Good, P. D., Winer, I., and Engelke, D. R. (2002) Effective expression of small interfering RNA in human cells. *Nat. Biotechnol. 20*, 505–508.
28. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol. 288*, 911–940.
29. Mathews, D. H., Disney, M. D., Childs, J. C., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U.S.A. 101*, 7287–7292.
30. Lu, Z. J., Turner, D. H., and Mathews, D. H. (2006) A set of nearest neighbor parameters for predicting the enthalpy change of RNA secondary structure formation. *Nucleic Acids Res. 34*, 4912–4924.
31. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res. 31*, 3406–3415.
32. Zuker, M. (1989) On finding all suboptimal foldings of an RNA molecule. *Science 244*, 48–52.

33. Hofacker, I. L. (2003) Vienna RNA secondary structure server. *Nucleic Acids Res. 31*, 3429–3431.
34. Lu, Z. J., Gloor, J. W., and Mathews, D. H. (2009) Improved RNA secondary structure prediction by maximizing expected pair accuracy. *RNA 15*, 1805–1813.
35. Davis, A. R., and Znosko, B. M. (2007) Thermodynamic characterization of single mismatches found in naturally occurring RNA. *Biochemistry 46*, 13425–13436.
36. Davis, A. R., and Znosko, B. M. (2008) Thermodynamic characterization of naturally occurring RNA single mismatches with G-U nearest neighbors. *Biochemistry 47*, 10178–10187.
37. Kierzek, R., Burkard, M. E., and Turner, D. H. (1999) Thermodynamics of single mismatches in RNA duplexes. *Biochemistry 38*, 14214–14223.
38. Blose, J. M., Manni, M. L., Klapec, K. A., Stranger-Jones, Y., Zyra, A. C., Sim, V., Griffith, C. A., Long, J. D., and Serra, M. J. (2007) Non-nearest-neighbor dependence of the stability for RNA bulge loops based on the complete set of group I single-nucleotide bulge loops. *Biochemistry 46*, 15123–15135.
39. Schroeder, S. J., and Turner, D. H. (2000) Factors affecting the thermodynamic stability of small asymmetric internal loops in RNA. *Biochemistry 39*, 9257–9274.
40. Siegfried, N. A., Metzger, S. L., and Bevilacqua, P. C. (2007) Folding cooperativity in RNA and DNA is dependent on position in the helix. *Biochemistry 46*, 172–181.
41. Longfellow, C. E., Kierzek, R., and Turner, D. H. (1990) Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides. *Biochemistry 29*, 278–285.
42. Ziomek, K., Kierzek, E., Biala, E., and Kierzek, R. (2002) The thermal stability of RNA duplexes containing modified base pairs at internal and terminal positions of the oligoribonucleotides. *Biophys. Chem. 97*, 233–241.
43. Badhwar, J., Karri, S., Cass, C. K., Wunderlich, E. L., and Znosko, B. M. (2007) Thermodynamic characterization of RNA duplexes containing naturally occurring 1 $\times$ 2 nucleotide internal loops. *Biochemistry 46*, 14715–14724.
44. Sheehy, J. P., Davis, A. R., and Znosko, B. M. (2010) Thermodynamic characterization of naturally occurring RNA tetraloops. *RNA 16*, 417–429.
45. Wright, D. J., Rice, J. L., Yanker, D. M., and Znosko, B. M. (2007) Nearest neighbor parameters for inosine-uridine pairs in RNA duplexes. *Biochemistry 46*, 4625–4634.
46. Plateau, P., and Gueron, M. (1982) Exchangeable protons without base line distortion using a new strong pulse sequence. *J. Am. Chem. Soc. 104*, 7310–7311.
47. Xia, T., SantaLucia, J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry 37*, 14719–14735.
48. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1990) Effects of GA mismatches on the structure and thermodynamics of RNA internal loops. *Biochemistry 29*, 8813–8819.
49. Marky, L. A., and Breslauer, K. J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers 26*, 1601–1620.
50. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1991) Stabilities of consecutive A.C, C.C, G.G, U.C, and U.U mismatches in RNA internal loops: Evidence for stable hydrogen-bonded U.U and $C.C^+$ pairs. *Biochemistry 30*, 8242–8251.
51. Xia, T. B., McDowell, J. A., and Turner, D. H. (1997) Thermodynamics of nonsymmetric tandem mismatches adjacent to G-C base pairs in RNA. *Biochemistry 36*, 12486–12497.
52. SantaLucia, J., Jr., and Turner, D. H. (1997) Measuring the thermodynamics of RNA secondary structure formation. *Biopolymers 44*, 309–319.
53. Mears, J. A., Sharma, M. R., Gutell, R. R., McCook, A. S., Richarson, P. E., Caulfield, T. R., Agrawal, R. K., and Harvey, S. C. (2006) A structural model for the large subunit of the mammalian mitochondrial ribosome. *J. Mol. Biol. 358*, 193–212.
54. Gillespie, J. J., McKenna, C. H., Yoder, M. J., Gutell, R. R., Johnston, J. S., Kathirithamby, J., and Cognato, A. I. (2005) Assessing the odd secondary structural properties of nuclear small subunit ribosomal RNA sequences (18S) of the twisted-wing parasites (insecta: strepsiptera). *Insect Mol. Biol. 14*, 625–643.
55. Gillespie, J. J., Johnston, J. S., Cannone, J. J., and Gutell, R. R. (2006) Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: hymenoptera): Structure, organization, and retrotransposable elements. *Insect Mol. Biol. 15*, 657–686.
56. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 angstrom resolution. *Science 289*, 905–920.
57. Elgavish, T., Cannone, J. J., Lee, J. C., Harvey, S. C., and Gutell, R. R. (2001) AA.AG@helix.ends: A:A and A:G base-pairs at the ends of 16 and 23 S rRNA helices. *J. Mol. Biol. 310*, 735–753.
58. Gutell, R. R. (1994) Collection of small-subunit (16s- and 16s-like) ribosomal-RNA structures: 1994. *Nucleic Acids Res. 22*, 3502–3507.
59. Gutell, R. R., and Fox, G. E. (1988) A Compilation of Large Subunit RNA Sequences Presented in a Structural Format. *Nucleic Acids Res. 16*, R175–R269.
60. Gutell, R. R., Gray, M. W., and Schnare, M. N. (1993) A compilation of large subunit (23s-like and 23s-like) ribosomal-RNA structures: 1993. *Nucleic Acids Res. 21*, 3055–3074.
61. Gutell, R. R., Weiser, B., Woese, C. R., and Noller, H. F. (1985) Comparative Anatomy of 16-S-Like Ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol. 32*, 155–216.
62. Schroeder, S., Kim, J., and Turner, D. H. (1996) G·A and U·U mismatches can stabilize RNA internal loops of three nucleotides. *Biochemistry 35*, 16105–16109.
63. SantaLucia, J., Kierzek, R., and Turner, D. H. (1991) Functional-group substitutions as probes of hydrogen-bonding between GA mismatches in RNA internal loops. *J. Am. Chem. Soc. 113*, 4313–4322.
64. Wu, M., McDowell, J. A., and Turner, D. H. (1995) A periodic table of symmetric tandem mismatches in RNA. *Biochemistry 34*, 3204–3211.
65. Walter, A. E., Wu, M., and Turner, D. H. (1994) The stability and structure of tandem GA mismatches in RNA depend on closing base pairs. *Biochemistry 33*, 11349–11354.
66. Christiansen, M. E., and Znosko, B. M. (2008) Thermodynamic characterization of the complete set of sequence symmetric tandem mismatches in RNA and an improved model for predicting the free energy contribution of sequence asymmetric tandem mismatches. *Biochemistry 47*, 4329–4336.
67. Christiansen, M. E., and Znosko, B. M. (2009) Thermodynamic characterization of tandem mismatches found in naturally occurring RNA. *Nucleic Acids Res. 37*, 4696–4706.