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[†]

Martha E. Christiansen and Brent M. Znosko*

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

Received October 17, 2007; Revised Manuscript Received January 7, 2008

ABSTRACT: Because of the availability of an abundance of RNA sequence information, the ability to rapidly and accurately predict the secondary structure of RNA from sequence is becoming increasingly important. A common method for predicting RNA secondary structure from sequence is free energy minimization. Therefore, accurate free energy contributions for every RNA secondary structure motif are necessary for accurate secondary structure predictions. Tandem mismatches are prevalent in naturally occurring sequences and are biologically important. A common method for predicting the stability of a sequence asymmetric tandem mismatch relies on the stabilities of the two corresponding sequence symmetric tandem mismatches [Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) *J. Mol. Biol. 288*, 911–940]. To improve the prediction of sequence asymmetric tandem mismatches, the experimental thermodynamic parameters for the 22 previously unmeasured sequence symmetric tandem mismatches are reported. These new data, however, do not improve prediction of the free energy contributions of sequence asymmetric tandem mismatches. Therefore, a new model, independent of sequence symmetric tandem mismatch free energies, is proposed. This model consists of two penalties to account for destabilizing tandem mismatches, two bonuses to account for stabilizing tandem mismatches, and two penalties to account for A-U and G-U adjacent base pairs. This model improves the prediction of asymmetric tandem mismatch free energy contributions and is likely to improve the prediction of RNA secondary structure from sequence.

The three most common base pairs in RNA are the Watson–Crick pairs, G-C and A-U, and the wobble G-U pair. These canonical base pairs are the components of the helical portions of RNA, and they have a regular structure and hydrogen bonding pattern. However, canonical base pairs account for only approximately half of the nucleotides found in RNA (*1*). The other half are involved in other secondary structure motifs, such as hairpins, bulges, and internal loops. One common RNA secondary structure motif is a tandem mismatch, or 2 × 2 internal loop. Tandem mismatches occur when two adjacent, noncanonical pairs are situated within a helical portion of canonical base pairs. The presence of tandem mismatches has been confirmed in a variety of RNA secondary structures (*2–10*) ranging from bacteria to trinucleotide repeats in human neurological diseases.

Fortunately, due to the pioneering efforts of projects such as the Human Genome Project (*11, 12*), entire genomes can now be sequenced accurately and efficiently. In recent years, thousands of RNA nucleotide sequences have been made publicly available (*13*). After a RNA sequence has been determined, the next logical step in better understanding structure and function is to determine an accurate method for predicting the secondary structure of RNA from its primary sequence.

The ability to predict secondary structure of RNA from sequence is important for several reasons. The determination of secondary structure can aid in the determination of tertiary structure. Also, because of the direct relationship between structure and function, the ability to predict secondary structure of RNA gives insight into the different functions and roles that RNA may have. In addition, being able to predict the secondary structure of RNA can help with the design of pharmaceuticals by providing an accurate target site for recognition by drugs.

The overwhelming importance of being able to predict RNA secondary structure from sequence has led to the development of several computer algorithms (*1, 14–18*). These algorithms use the method of free energy minimization to predict secondary structure from sequence. In this method, for a given sequence, all possible (non-pseudoknotted) conformations are tested by recursion. For each possible conformation, the free energy parameters of all secondary structure motifs (experimental or predicted) in that conformation are added to give a total free energy for that conformation. The total free energies for all possible secondary structure conformations are compared. The conformation with the lowest free energy is predicted to be the predominant species in solution. Computer programs used to predict RNA secondary structure from sequence, such as *RNAstructure* (*1, 14, 15*), *mfold* (*16, 17*), and the Vienna RNA

[†] Partial funding for this project was provided by the Saint Louis University College of Arts and Sciences, the Saint Louis University Department of Chemistry, a Saint Louis University Summer Research Award (B.M.Z.), the Saint Louis University Faculty Development Fund (B.M.Z.), and two Sigma Xi Grants-in-Aide of Research (M.E.C.).

* To whom correspondence should be addressed. Phone: (314) 977-8567. Fax: (314) 977-2521. E-mail: znoskob@slu.edu.

10.1021/bi7020876 CCC: $40.75
Published on Web 03/11/2008

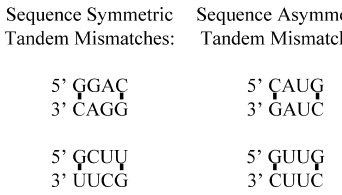

FIGURE 1: Examples of SSTM (left) and SATM (right).

Package (*18*), have become quite influential and popular within the scientific community. For example, the original article describing *mfold* (*17*) has been cited more than 1300 times (*19*), and the article describing the underlying algorithm has been cited 1244 times (*19*). In addition, *RNAstructure* has been downloaded more than 12000 times (D. H. Mathews, personal communication, February 27, 2006), and the original article describing this program (*1*) has been cited more than 1440 times (*19*).

Although influential and popular, these algorithms have room for improvement. If a tandem mismatch has been studied thermodynamically, the assigned value for the free energy contribution of that tandem mismatch is an experimental value. However, if a tandem mismatch has not been studied thermodynamically, a predicted value is assigned for the free energy contribution. When G-U pairs are classified as canonical pairs and when the identities of the mismatch nucleotides and the adjacent nearest neighbors are considered, there are 1830 possible tandem mismatches. Currently, only 112 (~6%) have been studied thermodynamically (*20–29*). Thus, the remaining 1718 combinations of tandem mismatches have free energy contributions that are predicted.

Tandem mismatches can be classified as sequence symmetric or sequence asymmetric. In sequence symmetric tandem mismatches (SSTM),[1] the 5′ to 3′ sequence of the top strand of the loop is identical to the 5′ to 3′ sequence of the bottom strand of the loop (Figure 1). All other sequence combinations of tandem mismatches are classified as sequence asymmetric tandem mismatches (SATM) (Figure 1). Of the 1830 possible tandem mismatches, 60 can be classified as SSTM and 1770 can be classified SATM.

An algorithm for predicting the free energy contribution of any unmeasured SATM is (*14*)

$$\Delta G^\circ_{37}\begin{pmatrix}5'\text{JMXW}\\3'\text{KNYZ}\end{pmatrix} = \left[\Delta G^\circ_{37}\begin{pmatrix}5'\text{JMNK}\\3'\text{KNMJ}\end{pmatrix} + \Delta G^\circ_{37}\begin{pmatrix}5'\text{ZYXW}\\3'\text{WXYZ}\end{pmatrix}\right]/2 + \Delta_P + \Delta_{GG} \quad (1)$$

where J-K and W-Z represent the canonical nearest neighbors, M·N and X·Y represent noncanonical pairs, Δ_P is a penalty of 0.6 kcal/mol assigned to tandem mismatches with an A·G or G·A pair adjacent to a U·C, C·U, or C·C pair and tandem mismatches with an A·A pair adjacent to a U·U pair, and Δ_{GG} is a bonus of −1.3 kcal/mol assigned to tandem mismatches with a G·G pair adjacent to either an A·A pair or any noncanonical pair with a pyrimidine. Therefore, the free energy contribution of a SATM is approximated by averaging the free energy contribution of the two corresponding SSTM and adding a penalty or a bonus for specific mismatch combinations. This model is based on the assumption that the energetics of a closing base pair and a mismatch in a SSTM is similar to that of the same closing base pair and mismatch in a SATM.

The periodic table of tandem mismatches (*1, 20, 29*) contains the free energy contributions of the SSTM. Of the 60 possible SSTM, only 38 have experimental free energy contributions. The 22 remaining SSTM have predicted free energy contributions. Therefore, predicted values for SSTM are often used to predict the free energy contributions of SATM. To improve prediction of SATM, the free energies for the 22 unmeasured SSTM are reported here, and a complete periodic table of tandem mismatches is constructed. However, the new experimental values for the SSTM do not improve the prediction of the free energy contributions of SATM. On the basis of these results, a new model for predicting the stability of SATM, independent of SSTM, is proposed.

MATERIALS AND METHODS

Design of Sequences for Optical Melting Studies. Sequences of tandem mismatches and closing base pairs were designed to represent those SSTM in the periodic table of tandem mismatches (*1, 20, 29*) that had predicted values for their free energy contribution. Most stems were chosen to be the same reference duplexes (*30–36*) as those used in previous tandem mismatch studies (*20–29*). The combinations of stems and tandem mismatches were chosen so that the duplexes had melting temperatures between 30 and 70 °C and so that minimal formation of hairpin structures or misaligned duplexes resulted. When placed into available stem sequences, two tandem mismatches studied here, $\left(\begin{smallmatrix}5'\text{UCCA}\\3'\text{ACCU}\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}5'\text{UACA}\\3'\text{ACAU}\end{smallmatrix}\right)$, resulted in RNA strands that could form competing structures. Therefore, these two tandem mismatches were studied within new stem sequences.

RNA Synthesis and Purification. Oligonucleotides were ordered from Integrated DNA Technologies, Inc. (Coralville, IA). The synthesis and purification of the oligonucleotides followed standard procedures that were described previously (*37–39*).

Concentration Calculations and Duplex Formation. Concentrations of the single-stranded oligoribonucleotides were calculated using Beer's law. The concentration of each individual strand was calculated from an absorbance measured at 280 nm and the single-strand extinction coefficient, which was calculated using *RNACalc* (*40*). To ensure that the absorbance was between 0.2 and 2.0, the samples were diluted. Furthermore, the absorbances of the oligoribonucleotides were measured at 80 °C to disrupt any single-strand folding. The tandem mismatch of interest was formed when a single-stranded oligoribonucleotide formed a bimolecular duplex with itself.

Optical Melting Experiments. 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 using a heating rate of 1 °C/min from 10 to 90 °C on a Beckman-Coulter DU800 spectrometer with a Beckman-Coulter high-performance temperature controller. The absorbance was measured at 280 nm. At least nine different concentrations were melted, representing a >50-fold concentration range.

[1] Abbreviations: R, purine nucleotides; SATM, sequence asymmetric tandem mismatches; SSTM, sequence symmetric tandem mismatches; Y, pyrimidine nucleotides.

Determination of Thermodynamic Parameters for Duplexes. Meltwin (*41*) was used to fit melting curves to a two-state model, assuming linear sloping baselines and temperature-independent ΔH° and ΔS° values (*35, 42*). Additionally, T_M values at different concentrations were used to calculate thermodynamic parameters of self-complementary duplexes according to the method of Borer et al. (*43*):

$$T_M^{-1} = (2.303R/\Delta H^\circ) \log C_T + (\Delta S^\circ/\Delta H^\circ) \quad (2)$$

where R is the gas constant, 1.987 cal mol^{-1} K^{-1}. For transitions that conform to the two-state model, ΔH° values from the two methods generally agree within 10%, which indicates that the two-state model is a good estimate of the transition (*22, 44*). To calculate the Gibbs free energy change at 37 °C, the following equation was used:

$$\Delta G^\circ_{37} = \Delta H^\circ - (310.15\ \text{K})\Delta S^\circ \quad (3)$$

Determination of the Contribution of Tandem Mismatches to Duplex Thermodynamics. The free energy contribution of most tandem mismatches was calculated using an experimentally determined free energy value for a reference duplex, or the same duplex without the tandem mismatch. For example

$$\Delta G^\circ_{37,\text{tandem mismatch}} = \Delta G^\circ_{37}\begin{pmatrix}5'\text{GAUGGGCAUC}\\3'\text{CUACGGGUAG}\end{pmatrix} - \Delta G^\circ_{37}\begin{pmatrix}5'\text{GAUGCAUC}\\3'\text{CUACGUAG}\end{pmatrix} + \Delta G^\circ_{37}\begin{pmatrix}5'\text{GC}\\3'\text{CG}\end{pmatrix} \quad (4)$$

where $\Delta G^\circ_{37}\left(\begin{smallmatrix}5'\text{GAUGGGCAUC}\\3'\text{CUACGGGUAG}\end{smallmatrix}\right)$ is the free energy measured for the duplex containing the tandem mismatch of interest, $\Delta G^\circ_{37}\left(\begin{smallmatrix}5'\text{GAUGCAUC}\\3'\text{CUACGUAG}\end{smallmatrix}\right)$ is the free energy measured previously (*31*) for the reference duplex without the tandem mismatch, and $\Delta G^\circ_{37}\left(\begin{smallmatrix}5'\text{GC}\\3'\text{CG}\end{smallmatrix}\right)$ is the nearest-neighbor interaction (*31*) disrupted by the insertion of the tandem mismatch into the stem duplex. For duplexes without available stem thermodynamics, the free energy contribution of the tandem mismatch was approximated by a nearest-neighbor model (*31*):

$$\Delta G^\circ_{37,\text{tandem mismatch}} = \Delta G^\circ_{37}\begin{pmatrix}5'\text{GAUGGGCAUC}\\3'\text{CUACGGGUAG}\end{pmatrix} - \Delta G^\circ_{37,i} - \Delta G^\circ_{37,\text{symm}} - \Delta G^\circ_{37}\begin{pmatrix}5'\text{GA}\\3'\text{CU}\end{pmatrix} - \Delta G^\circ_{37}\begin{pmatrix}5'\text{AU}\\3'\text{UA}\end{pmatrix} - \Delta G^\circ_{37}\begin{pmatrix}5'\text{UG}\\3'\text{AC}\end{pmatrix} - \Delta G^\circ_{37}\begin{pmatrix}5'\text{CA}\\3'\text{GU}\end{pmatrix} - \Delta G^\circ_{37}\begin{pmatrix}5'\text{AU}\\3'\text{UA}\end{pmatrix} - \Delta G^\circ_{37}\begin{pmatrix}5'\text{UC}\\3'\text{AG}\end{pmatrix} \quad (5)$$

where $\Delta G^\circ_{37,i}$ is the free energy change for duplex initiation [4.09 kcal/mol (*31*)], $\Delta G^\circ_{37,\text{symm}}$ is a symmetry correction for self-complementary duplexes [0.43 kcal/mol (*31*)], and the remainder of the terms are individual nearest-neighbor values (*31*). The values of $\Delta H^\circ_{\text{tandem mismatch}}$ and $\Delta S^\circ_{\text{tandem mismatch}}$ were calculated in a similar manner, using reference duplexes if available and the nearest-neighbor model if reference duplexes were not available.

Predicting Stability of SATM by Using Experimental SSTM Values. The free energies of 18 SATM, all of which contain a closing pair and adjacent mismatch of the same sequence as a SSTM studied here, were found in the literature (*24, 26, 28, 29*). The $\Delta G^\circ_{37,\text{loop}}$ values for these SATM were predicted using two different methods. The first method utilized eq 1 with the predicted SSTM contributions from previously published periodic tables of tandem mismatches (*1, 20, 29*). The second method utilized eq 1 with the experimental SSTM contributions reported here. To determine if the experimental values for the SSTM reported here improve prediction of SATM free energy contributions, the results of these two methods were compared. Surprisingly, a better prediction of ΔG°_{37} values resulted from the predicted rather than the experimental SSTM free energy contributions.

Improved Model for the Prediction of SATM Free Energy Contributions. Since the experimental SSTM data did not improve prediction of SATM free energy contributions, a new model was derived using linear regression. A database of 74 SATM studied previously (*21, 24, 26, 28, 29*) was compiled. Seven variables, based on the predicted model for internal loops with sequence 5′GXYG3′/3′CWZC5′ (where W, X, Y, and Z are nucleotides in a tandem mismatch) (*45*), were used for linear regression: (1) tandem mismatches with adjacent U•U and R•R pairs, tandem mismatches with one G•A or A•G pair and one Y•Y pair, or tandem mismatches with adjacent A•C, U•C, C•U, C•C, C•A, or A•A pairs; (2) tandem mismatches with any combination of adjacent G•A and A•G pairs or adjacent U•U pairs; (3) tandem mismatches with a U•U pair adjacent to a Y•Y (not U•U), C•A, or A•C pair; (4) tandem mismatches with a single G•G pair not adjacent to a U•U pair; (5) tandem mismatches with an A-U nearest neighbor; and (6) tandem mismatches with a G-U nearest neighbor. The calculated experimental contribution of the tandem mismatch to duplex stability was used as a constant when doing linear regression. To simultaneously solve for each variable, the LINEST function of *Microsoft Excel* was used for linear regression. Many combinations of variables were tried, but this combination of variables produced a model that agreed closely with the experimental data and had error values that were comparable to those of the *RNAstructure* algorithm (*1, 14, 15*).

RESULTS

Thermodynamic Parameters of Duplexes Containing SSTM. Table 1 shows the thermodynamic parameters of duplex formation that 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. As evidenced by the agreement of ΔH° values determined by the two different methods, all duplexes melted in a two-state manner.

Contribution of Tandem Mismatches to Duplex Thermodynamics. The contributions of the 22 SSTM to duplex stability are listed in Table 2. These contributions are described in Materials and Methods and are further defined by eqs 4 and 5. The SSTM studied here contribute −0.64 to 6.00 kcal/mol, −23.5 to 36.1 kcal/mol, and −73.7 to 97.4 eu to duplex free energy, enthalpy, and entropy, respectively. Table 2 also shows the free energy contributions predicted in previous periodic tables of tandem mismatches (*1, 20, 29*). On average, the predicted values differ from the experimental values by 1.4 kcal/mol, and differences range from 0.1 to 3.7 kcal/mol.

These measured free energy contributions were combined with data from previously measured SSTM, resulting in an updated, complete periodic table of tandem mismatches

Table 1: Thermodynamic Parameters for Duplex Formation[a]

	curve fit parameters				$1/T_M$ vs log C_T parameters			
sequence[b]	$-\Delta G°_{37}$ (kcal/mol)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	T_m[c] (°C)	$-\Delta G°_{37}$ (kcal/mol)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	T_m[c] (°C)
GGCG^{AA}UGCC CCGU$_{AA}$GCGG	9.66 ± 0.28	45.2 ± 2.7	114.5 ± 8.0	67.0	9.67 ± 0.16	45.1 ± 1.8	114.2 ± 5.5	67.2
GGCG^{AC}UGCC CCGU$_{CA}$GCGG	9.41 ± 0.56	43.8 ± 7.8	111.0 ± 23.4	65.9	9.44 ± 0.31	43.1 ± 3.7	108.5 ± 10.9	66.7
GCAU^{AC}AUGC CGUA$_{CA}$UACG	5.37 ± 0.29	45.5 ± 13.7	129.5 ± 44.8	34.9	5.20 ± 0.18	47.3 ± 5.1	135.7 ± 16.5	33.9
GGAU^{AC}GUCC CCUG$_{CA}$UAGG	6.44 ± 0.18	58.4 ± 8.9	167.4 ± 28.6	41.1	6.49 ± 0.04	55.7 ± 2.8	158.5 ± 9.0	41.6
CGGG^{CA}UCCG GCCU$_{AC}$GGGC	6.46 ± 0.14	35.5 ± 4.8	93.7 ± 15.7	44.0	6.49 ± 0.04	39.5 ± 2.0	106.3 ± 6.5	43.5
GAGU^{CA}GCUC CUCG$_{AC}$UGAG	6.16 ± 0.10	64.6 ± 5.6	188.4 ± 18.1	39.3	6.16 ± 0.01	74.1 ± 2.1	219.1 ± 6.9	39.0
GCA^{CC}UGC CGU$_{CC}$ACG	2.65 ± 0.31	34.7 ± 4.1	103.2 ± 14.2	12.1	2.66 ± 0.14	34.5 ± 1.4	102.5 ± 5.1	12.1
CCG^{CC}CGG GGC$_{CC}$GCC	5.34 ± 0.40	41.2 ± 9.7	115.6 ± 31.6	34.5	5.31 ± 0.05	33.7 ± 1.5	91.6 ± 5.0	33.7
GUCG^{CC}UGAC CAGU$_{CC}$GCUG	5.34 ± 0.45	48.2 ± 15.8	138.3 ± 52.2	34.8	5.34 ± 0.04	49.6 ± 1.9	142.8 ± 6.2	34.9
CUGU^{CC}ACAG GACA$_{CC}$UGUC	3.40 ± 0.49	52.4 ± 6.7	158.0 ± 23.0	24.1	3.42 ± 0.20	51.0 ± 2.9	153.5 ± 9.9	23.9
GAGU^{CC}GCUC CUCG$_{CC}$UGAG	7.06 ± 0.72	53.7 ± 25.2	150.5 ± 79.1	45.2	7.05 ± 0.14	53.6 ± 4.2	150.2 ± 13.3	45.2
GAGA^{CU}UCUC CUCU$_{UC}$AGAG	6.09 ± 0.09	70.0 ± 8.7	205.9 ± 28.0	38.8	6.03 ± 0.04	65.2 ± 3.3	190.8 ± 10.5	38.7
CGGG^{CU}UCCG GCCU$_{UC}$GGGC	6.58 ± 0.21	31.0 ± 5.0	78.7 ± 15.7	46.3	6.55 ± 0.03	30.7 ± 1.2	77.8 ± 3.8	46.1
CCAU^{CU}GUGG GGUG$_{UC}$UACC	4.53 ± 0.28	33.0 ± 9.1	91.9 ± 30.2	26.6	4.53 ± 0.25	29.4 ± 3.2	80.3 ± 10.8	25.4
GAUG^{GG}CAUC CUAC$_{GG}$GUAG	7.29 ± 0.37	75.2 ± 9.7	219.0 ± 30.2	43.8	7.34 ± 0.12	81.4 ± 5.9	238.7 ± 18.6	43.5
GGCG^{GG}UGCC CCGU$_{GG}$GCGG	9.62 ± 0.37	59.2 ± 7.7	159.8 ± 23.8	59.1	9.61 ± 0.16	59.2 ± 2.7	159.9 ± 8.1	59.1
GGAU^{GG}GUCC CCUG$_{GG}$UAGG	7.39 ± 0.17	43.9 ± 8.6	117.8 ± 27.5	49.6	7.41 ± 0.23	47.5 ± 5.3	129.1 ± 16.7	48.8
GCA^{UC}UGC CGU$_{CU}$ACG	3.37 ± 0.40	36.0 ± 5.6	105.3 ± 19.0	18.4	3.38 ± 0.20	35.2 ± 2.6	102.5 ± 8.9	18.0
GCG^{UC}CGC CGC$_{CU}$GCG	5.81 ± 0.15	57.3 ± 6.1	166.0 ± 20.1	37.7	5.72 ± 0.08	58.3 ± 3.1	169.6 ± 10.1	37.2
CCUG^{UC}UAGG GGAU$_{CU}$GUCC	7.72 ± 0.15	37.9 ± 4.5	97.4 ± 14.0	54.7	7.72 ± 0.04	38.1 ± 1.0	97.9 ± 3.1	54.6
CCAU^{UC}GUGG GGUG$_{CU}$UACC	7.80 ± 0.49	60.6 ± 15.4	170.2 ± 48.3	48.3	7.87 ± 0.14	66.5 ± 4.3	189.1 ± 13.4	47.6
GGCG^{UU}UGCC CCGU$_{UU}$GCGG	9.68 ± 0.33	54.0 ± 4.3	142.8 ± 13.0	61.9	9.67 ± 0.09	53.6 ± 1.3	141.5 ± 3.7	62.0

[a] Measurements were taken in 1.0 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na_2EDTA (pH 7.0). [b] The top strand of each duplex is written from 5′ to 3′, and each bottom strand is written from 3′ to 5′. [c] Calculated at an oligomer concentration of 10^{-4} M.

(Table 3). On average, the ranking of the tandem mismatches in SSTM from most stable to least stable is as follows: $(5'GA)_2$ at −0.1 kcal/mol > $(5'UU)_2$ at 0.8 kcal/mol > $(5'AG)_2$ at 1.1 kcal/mol > $(5'GG)_2$ at 1.2 kcal/mol > $(5'UC)_2$ at 1.5 kcal/mol > $(5'AA)_2$ at 2.0 kcal/mol > $(5'AC)_2$ at 2.2 kcal/mol > $(5'CA)_2$ at 2.6 kcal/mol > $(5'CC)_2$ at 2.8 kcal/mol > $(5'CU)_2$ at 2.9 kcal/mol. On average, the ranking of closing base pairs adjacent to SSTM from most to least stable is as follows: G-C at 0.2 kcal/mol > C-G at 0.8 kcal/mol > A-U at 2.2 kcal/mol ≈ U-G at 2.2 kcal/mol > U-A at 2.4 kcal/mol ≈ G-U at 2.4 kcal/mol.

Updated Model for Predicting Free Energy Contributions of SATM. The free energies of 18 SATM, all of which contain a closing pair and adjacent mismatch of the same sequence as a SSTM studied here, were found in the literature (*24, 26, 28, 29*). The $\Delta G°_{37,loop}$ values for these SATM were predicted using two different methods. The first method utilized eq 1 with the predicted SSTM contributions from previously published periodic tables of tandem mismatches (*1, 20, 29*). The average absolute deviation between the measured and predicted free energy values using this method was 0.5 kcal/mol (Table S1 of the Supporting Information). The second method utilized eq 1 with the experimental SSTM contributions reported here. The average absolute deviation between the measured and predicted free energy values using this method was 0.8 kcal/mol (Table S1). Surprisingly, the measured SSTM data did not improve the prediction of SATM free energy contributions. Shankar et al. (*45*) recently suggested that the assumption that energetics of a closing base pair and mismatch in a SSTM would be similar to the energetics of the same closing base pair and mismatch in a SATM is not a reasonable assumption. Therefore, a new model for predicting the stability of SATM without utilizing the comprising SSTM was derived.

Although several models were generated, a model based on a table of predicted 2 × 2 free energy increments proposed by Shankar et al. (*45*) resulted in free energy contributions that agreed closely with the experimental data and error values that were comparable to those of the *RNAstructure* algorithm (*1, 14, 15*). This model is shown in Table 4. For the 74 SATM studied previously (*21, 24, 26, 28, 29*), the average absolute difference between the measured values and the values predicted with this new model is 0.48 ± 0.32 kcal/mol (data not shown). This is an improvement over the average absolute difference between the measured values and the values predicted with eq 1 and the experimentally determined SSTM, 0.65 ± 0.55 kcal/mol (data not shown). This newly proposed model should be used only when experimental data of SATM are unavailable. Since all SSTM have now been measured, experimental values, rather than this model, should always be used for SSTM. A similar model for calculating enthalpy contributions of SATM is available in Table S2 of the Supporting Information.

Table 2: Contributions of Tandem Mismatches to Duplex Thermodynamics

sequence[a]	$\Delta G^\circ_{37,\text{tandem mismatch}}$[b] (kcal/mol)	$\Delta H^\circ_{\text{tandem mismatch}}$ (kcal/mol)	$\Delta S^\circ_{\text{tandem mismatch}}$ (eu)
GGCG^{AA}UGCC CCGU$_{AA}$GCGG	1.34 (4.1)	13.7	41.5
GGCG^{AC}UGCC CCGU$_{CA}$GCGG	1.57 (2.5)	15.7	47.2
GCAU^{AC}AUGC CGUA$_{CA}$UACG	3.97 (2.8)	18.5	46.8
GGAU^{AC}GUCC CCUG$_{CA}$UAGG	2.21 (2.8)	8.1	19.1
CGGG^{CA}UCCG GCCU$_{AC}$GGGC	6.00 (2.3)	27.4	68.9
GAGU^{CA}GCUC CUCG$_{AC}$UGAG	3.54 (1.9)	−0.4	−12.5
GCA^{CC}UGC CGU$_{CC}$ACG	3.62 (2.2)	18.5	48.0
CCG^{CC}CGG GGC$_{CC}$GCC	1.11 (1.0)	12.2	35.8
GUCG^{CC}UGAC CAGU$_{CC}$GCUG	2.00 (2.2)	4.9	9.3
CUGU^{CC}ACAG GACA$_{CC}$UGUC	5.35 (2.8)	10.0	15.0
GAGU^{CC}GCUC CUCG$_{CC}$UGAG	2.65 (2.8)	20.1	56.4
GAGA^{CU}UCUC CUCU$_{UC}$AGAG	2.98 (2.2)	0.4	−8.3
CGGG^{CU}UCCG GCCU$_{UC}$GGGC	5.94 (2.2)	36.1	97.4
CCAU^{CU}GUGG GGUG$_{UC}$UACC	3.57 (2.2)	31.8	91.0
GAUG^{GG}CAUC CUAC$_{GG}$GUAG	−0.64 (0.8)	−23.5	−73.7
GGCG^{GG}UGCC CCGU$_{GG}$GCGG	1.40 (2.1)	−0.4	−4.2
GGAU^{GG}GUCC CCUG$_{GG}$UAGG	1.28 (2.1)	16.3	48.5
GCA^{UC}UGC CGU$_{CU}$ACG	2.90 (2.2)	17.8	48.0
GCG^{UC}CGC CGC$_{CU}$GCG	1.48 (1.0)	−7.2	−28.0
CCUG^{UC}UAGG GGAU$_{CU}$GUCC	0.38 (2.2)	18.4	58.2
CCAU^{UC}GUGG GGUG$_{CU}$UACC	0.23 (2.8)	−5.3	−17.8
GGCG^{UU}UGCC CCGU$_{UU}$GCGG	1.34 (0.6)	5.3	14.2

[a] The top strand of every duplex is written from 5′ to 3′, and each bottom strand is written from 3′ to 5′. [b] Numbers in parentheses are the predicted values found in the periodic table of tandem mismatches (*1, 20, 29*).

Table 3: Complete Periodic Table of Tandem Mismatches[a]

closing base pair[c]	tandem mismatch[b] 5′GA 3′AG	5′UU 3′UU	5′AG 3′GA	5′GG 3′GG	5′UC 3′CU	5′AA 3′AA	5′AC 3′CA	5′CA 3′AC	5′CC 3′CC	5′CU 3′UC
5′G 3′C	**−2.6**[d]	−0.5[e]	−1.3[d]	−0.6	1.5	1.5[e]	0.9[e]	1.0[e]	1.1	**1.1**[e]
5′C 3′G	−0.7[f]	−0.4[g]	**−0.7**[f,h]	0.8[g]	1.4[g]	1.3[f]	2.0[g]	1.1[g]	1.7[g]	1.4[g]
5′A 3′U	**0.3**[d]	0.6[e]	1.7[i]	1.9[j]	2.9	2.8[e]	2.5[e]	2.3[e]	3.6	3.0
5′U 3′G	0.1[d]	2.5[e]	3.4[i]	1.3	0.2	2.1[e]	2.2	3.5	2.7	3.6
5′U 3′A	0.7[d]	1.1[e]	0.9[i]	2.3[j]	2.8[e]	2.8[e]	4.0	1.9[e]	5.4	2.2[e]
5′G 3′U	1.8[d]	1.3	2.6[i]	1.4	0.4	1.3	1.6	6.0	2.0	5.9

[a] The values are free energy values in kilocalories per mole. Values in bold are averages. [b] Tandem mismatches are arranged from most stable (left) to least stable (right). [c] Closing base pairs are arranged from most stable (top) to least stable (bottom). [d] From ref (*27*). [e] From ref (*20*). [f] From ref (*22*). [g] From ref (*25*). [h] From ref (*24*). [i] From ref (*29*). [j] From ref (*26*).

DISCUSSION

The foundation for the algorithm currently used to predict the free energy contribution of SATM is the periodic table of tandem mismatches (*1, 20, 29*). From this table, the free energy contribution of any SATM can be calculated by averaging the free energy contribution of the two SSTM that comprise the SATM. Despite the importance of this table, 22 SSTM had not been determined experimentally. Thus, since these predicted free energy contributions are used to predict free energy contributions for SATM, there was a need to thermodynamically characterize the complete set of SSTM.

Table 4: Model for Predicting the Free Energy Contribution of Sequence Asymmetric Tandem Mismatches

tandem mismatches with the following[a]	free energy increments (kcal/mol)
a U·U pair adjacent to an R·R pair, a G·A or A·G pair adjacent to a Y·Y pair, or any combination of A·C, U·C, C·U, C·C, C·A, or A·A pairs	1.1 ± 0.1
any combination of adjacent G·A and A·G pairs or two U·U pairs	−1.2 ± 0.3
a U·U pair adjacent to a Y·Y (not U·U), C·A, or A·C pair	0.8 ± 0.2
a G·G pair not adjacent to a U·U pair	−0.3 ± 0.2
per A-U nearest neighbor	0.5 ± 0.2
per G-U nearest neighbor	1.2 ± 0.1

[a] Any other base pair combinations in a tandem mismatch do not contribute to duplex stability.

Contribution of Tandem Mismatches to Duplex Thermodynamics. The contributions of 22 SSTM to duplex stability are listed in Table 2. Also in Table 2, a comparison between the measured contributions and the contributions predicted in previous periodic tables of tandem mismatches is made (*1, 20, 29*). This comparison shows that half of the predicted values are ≥1 kcal/mol different from the measured values, suggesting that previous assumptions were not accurate.

Updated, Complete Periodic Table of Tandem Mismatches. Now that all of the SSTM have experimental free energy contributions, the periodic table of tandem mismatches has been updated (Table 3). Similar to previous representations of the periodic table of tandem mismatches (*1, 20, 29*), this table was arranged so that the tandem mismatches were ordered from most stable to least stable. It is important to note, however, that the ranking of tandem mismatches is different from the previous ranking (*1*). In agreement with earlier studies (*1, 20, 29*), tandem mismatches with A·G, U·U, and G·A tandem mismatches are the most stable mismatches. Similarly, the table was arranged so that the closing base pairs are ordered from most stable to least stable. As expected, G-C and C-G closing base pairs are the most stable closing pairs adjacent to symmetric tandem mismatches. However, these trends do not hold true for the entire table. For example, when considering all possible tandem mismatches with a particular nearest neighbor (any row of Table 3), the free energy contributions do not always increase for every column when moving from left to right. Similarly, when considering all possible closing base pairs for a particular tandem mismatch (any column of Table 3), the free energy contributions increase for each row when moving from top to bottom for only one tandem mismatch, $(5'GA)_2$. Therefore, it is obvious that idiosyncrasies are prevalent throughout the table. For example, $(5'UC)_2$, although it does not contain an A·G, G·A, or U·U mismatch, is the most stable tandem mismatch adjacent to a G-U closing pair. From the thermodynamic trends alone, however, it is difficult to explain the oddities observed throughout the table. These irregularities may be due to hydrogen bonding, stacking,

solvation, backbone distortion, and/or the inter-relationship between these properties. More data, such as structural data from NMR or X-ray crystallography and data from electrostatic calculations, may be helpful in further understanding the results seen here.

Using SSTM To Predict the Free Energy Contribution of SATM. To determine if the experimental data for SSTM would improve the prediction of the free energy contribution of SATM, a database of 74 previously studied SATM (*20–29*) was compiled. Of these 74 duplexes, 18 (*24, 26, 28, 29*) contained tandem mismatches with a closing pair and an adjacent mismatch of the same sequence as a SSTM studied here. The free energy contributions of the 18 SATM were calculated using eq 1. First, the free energy contributions were calculated using the predicted values available in previous periodic tables of tandem mismatches (*1, 20, 29*). Then, the free energy contributions were calculated using the measured values in Table 3. Finally, the free energy contributions calculated by both methods were compared to the experimental values determined previously. Surprisingly, when the predicted symmetric tandem mismatch values (*1, 20, 29*) were used, the predicted free energy contributions were 0.3 kcal/mol closer to the experimental values than when the experimental symmetric free energy contributions (Table 3) were used. These results, the idiosyncrasies associated with Table 3, and the hypothesis (*45*) that assumptions associated with using SSTM to predict the stability of SATM were not reasonable suggested that a new model, independent of SSTM, may better predict the stability of SATM.

Updated Model for Predicting Free Energy Contributions of SATM. Using NMR evidence, Shankar et al. (*45*) proposed a model for predicting the free energy contribution of 5′GXYG3′/3′CWZC5′ (where W, X, Y, and Z are nucleotides in a tandem mismatch) that is independent of SSTM. Although several models were tested here (data not shown), a new model based on the model of Shankar et al. (*45*) resulted in free energy contributions that agreed closely with the experimental data and error values that were comparable to those of the *RNAstructure* algorithm (*1, 14, 15*). This model is shown in Table 4 and consists of six parameters.

Most of these parameters correspond to the categories used by Shankar et al. (*45*), and the reason for categorizing tandem mismatches into those categories has been described previously (*45*). The first parameter is for tandem mismatches with a U•U pair adjacent to an R•R pair, a G•A or A•G pair adjacent to a Y•Y pair, or any combination of A•C, U•C, C•U, C•C, C•A, or A•A pairs. The free energy increment derived here, 1.1 ± 0.1 kcal/mol, corresponds with the increment proposed previously, 1.2 ± 0.4 kcal/mol (*45*). The second parameter is for tandem mismatches with any combination of adjacent G•A and A•G pairs or for tandem mismatches with two U•U pairs. Although Shankar et al. (*45*) suggested that these loops could be calculated from the corresponding SSTM, a free energy parameter independent of the SSTM, −1.2 kcal/mol, was derived here. The third parameter contains tandem mismatches with a U•U pair adjacent to a Y•Y (not U•U), C•A, or A•C pair. Shankar et al. (*45*) included these tandem mismatches and those that have a G•A or A•G pair adjacent to an A•C, C•A, or A•A pair in the same category. However, in the linear regression done here, the latter did not contribute to the free energy and is not included in Table 4. The former contributes 0.8 ± 0.2 kcal/mol to duplex stability. The fourth parameter is for tandem mismatches that contain a G•G pair not adjacent to a U•U pair. The free energy increment derived here, -0.3 ± 0.2 kcal/mol, corresponds with the increment proposed previously, -0.2 ± 0.3 kcal/mol (*45*). Fifth and sixth parameters were added in addition to the categories proposed by Shankar et al. (*45*) to account for tandem mismatches adjacent to A-U and G-U pairs. Since the model proposed by Shankar et al. (*45*) was for tandem mismatches closed by G-C pairs, these parameters were not included in their model.

One interesting feature of the model proposed in Table 4 is the difference in the penalty per A-U nearest neighbor (0.5 kcal/mol) and the penalty per G-U nearest neighbor (1.2 kcal/mol). Previous models used to predict the free energy contribution of RNA secondary structure motifs have treated G-U pairs adjacent to secondary structure motifs as A-U pairs (*1, 14, 15*). Recent studies on the thermodynamics of single mismatches (*38*) and 1×2 internal loops (*39*) have found no significant difference between A-U and G-U pairs adjacent to these motifs. The results shown here, however, suggest that tandem mismatches may behave in a manner different from that of other RNA secondary structure motifs. A tandem mismatch may be flexible enough to allow adjacent G-U pairs to adopt a nonwobble conformation or a combination of conformations. More studies with A-U and G-U pairs adjacent to secondary structure motifs will determine if only tandem mismatches exhibit this property or if other secondary structure motifs also display this thermodynamic difference between motifs adjacent to A-U pairs and those adjacent to G-U pairs.

The new parameters in Table 4 were tested on the database of 74 previously measured SATM (*21, 24, 26, 28, 29*). The average absolute difference between predicted and measured values is 0.48 kcal/mol. This is slightly better than the value with the model using eq 1 and the previous periodic tables of tandem mismatches (*1, 20, 29*), which resulted in an average absolute difference of 0.65 kcal/mol between predicted and measured values. Although this new model results in reasonable predictions for most loops, there are four tandem mismatches that have absolute differences between predicted and measured values of >1 kcal/mol, 5′UAGU3′/3′GAAG5′, 5′UGAU3′/3′GAGG5′, 5′GACG3′/3′CACC5′, and 5′GCGG3′/3′CCAC5′. Evidently, there are still idiosyncrasies, such as non-nearest-neighbor effects, that have yet to be discovered.

ACKNOWLEDGMENT

We thank Neel Shankar and Doug Turner for providing a copy of their manuscript prior to publication (*45*).

SUPPORTING INFORMATION AVAILABLE

Comparison of two methods for calculating the free energy contributions of SATM (Table S1) and model for predicting the enthalpy contribution of sequence asymmetric tandem mismatches (Table S2). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol. 288*, 911–940.

2. Gutell, R. R. (1994) Collection of small-subunit (16s- and 16s-like) ribosomal-RNA structures-1994. *Nucleic Acids Res. 22*, 3502–3507.
3. 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.
4. 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.
5. Gutell, R. R., and Woese, C. R. (1990) Higher order structural elements in ribosomal RNAs: Pseudo-knots and the use of noncanonical pairs. *Proc. Natl. Acad. Sci. U.S.A. 87*, 663–667.
6. 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.
7. Disney, M. D., Haidaris, C. G., and Turner, D. H. (2001) Recognition elements for 5′ exon substrate binding to the *Candida albicans* group I intron. *Biochemistry 40*, 6507–6519.
8. DeNap, J. C. B., Thomas, J. R., Musk, D. J., and Hergenrother, P. J. (2004) Combating drug-resistant bacteria: Small molecule mimics of plasmid incompatibility as antiplasmid compounds. *J. Am. Chem. Soc. 126*, 15402–15404.
9. Zhang, J. C., Zhang, G. H., Guo, R., Shapiro, B. A., and Simon, A. E. (2006) A pseudoknot in a preactive form of a viral RNA is part of a structural switch activating minus-strand synthesis. *J. Virol. 80*, 9181–9191.
10. Sobczak, K., de Mezer, M., Michlewski, G., Krol, J., and Krzyzosiak, W. J. (2003) RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res. 31*, 5469–5482.
11. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. Q. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J. H., Miklos, G. L. G., Nelson, C., Broder, S., Clark, A. G., Nadeau, C., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z. M., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W. M., Gong, F. C., Gu, Z. P., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z. X., Ketchum, K. A., Lai, Z. W., Lei, Y. D., Li, Z. Y., Li, J. Y., Liang, Y., Lin, X. Y., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B. X., Sun, J. T., Wang, Z. Y., Wang, A. H., Wang, X., Wang, J., Wei, M. H., Wides, R., Xiao, C. L., Yan, C. H., et al. (2001) The sequence of the human genome. *Science 291*, 1304.
12. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., et al. (2001) Initial sequencing and analysis of the human genome. *Nature 409*, 860–921.
13. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Wheeler, D. L. (2007) GenBank. *Nucleic Acids Res. 35*, D21–D25.
14. 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.
15. 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.
16. Zuker, M. (1989) On finding all suboptimal foldings of an RNA molecule. *Science 244*, 48–52.
17. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res. 31*, 3406–3415.
18. Hofacker, I. L. (2003) Vienna RNA secondary structure server. *Nucleic Acids Res. 31*, 3429–3431.
19. *ISI Web of Knowledge*, version 3.0, Thomson Corp., Stamford, CT. http://portal.isiknoweldge.com/portal.cgi (accessed October 2007).
20. Wu, M., McDowell, J. A., and Turner, D. H. (1995) A periodic table of symmetric tandem mismatches in RNA. *Biochemistry 34*, 3204–3211.
21. 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.
22. 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.
23. Schroeder, S. J., Burkard, M. E., and Turner, D. H. (2001) The energetics of small internal loops in RNA. *Biopolymers 52*, 157–167.
24. 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.
25. 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.
26. Burkard, M. E., Xia, T., and Turner, D. H. (2001) Thermodynamics of RNA internal loops with a guanosine-guanosine pair adjacent to another noncanonical pair. *Biochemistry 40*, 2478–2483.
27. 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.
28. Bourdelat-Parks, B. N., and Wartell, R. M. (2005) Thermodynamics of RNA duplexes with tandem mismatches containing a uracil-uracil pair flanked by C-G/G-C or G-C/A-U closing base pairs. *Biochemistry 44*, 16710–16717.
29. Schroeder, S. J., and Turner, D. H. (2001) Thermodynamic stabilities of internal loops with GU closing pairs in RNA. *Biochemistry 40*, 11509–11517.
30. Sugimoto, N., Kierzek, R., and Turner, D. H. (1987) Sequence dependence for the energetics of dangling ends and terminal base pairs in ribonucleic acid. *Biochemistry 26*, 4554–4558.
31. 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.
32. Freier, S. M., Sinclair, A., Nielson, T., and Turner, D. H. (1985) Improved free energies for G-C base pairs. *J. Mol. Biol. 185*, 645–647.
33. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) Improved free energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373–9377.
34. He, L., Kierzek, R., SantaLucia, J., Jr., Walter, A. E., and Turner, D. H. (1991) Nearest neighbor parameters for G-U mismatches: 5′GU3′/3′UG5′ is destabilizing in the contexts CGUG/GUGC, UGUA/AUGU, and AGUU/UUGA but stabilizing in GGUC/CUGG. *Biochemistry 30*, 11124–11132.
35. McDowell, J. A., and Turner, D. H. (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: Solution structure of $(rGAGGUCUC)_2$ by two-dimensional NMR and simulated annealing. *Biochemistry 35*, 14077–14089.
36. McDowell, J. A., He, L., Chen, X., and Turner, D. H. (1997) Investigation of the structural basis for thermodynamic stabilities of tandem GU wobble pairs: NMR structures of $(rGGAGUUCC)_2$ and $(rGGAUGUCC)_2$. *Biochemistry 36*, 8030–8038.

37. 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.
38. Davis, A. R., and Znosko, B. M. (2007) Thermodynamic characterization of single mismatches found in naturally occurring RNA. *Biochemistry 46*, 13425–13436.
39. Badhwar, J., Karri, S., Cass, C. K., Wunderlich, E. L., and Znosko, B. M. (2007) Thermodynamic characterization of RNA duplexes containing naturally occurring 1 × 2 nucleotide internal loops. *Biochemistry 46*, 14715–14724.
40. McDowell, J. A. (1995) *RNA Calculations*, version 1.1.
41. McDowell, J. A. (1996) *MeltWin*, version 3.5 (Melt Curve Processing Program). This program is available for download at www.meltwin.com.
42. Petersheim, M., and Turner, D. H. (1983) Base-stacking and base-pairing contributions to helix stability: Thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry 22*, 256–263.
43. Borer, P. N., Dengler, B., Tinoco, I., and Uhlenbeck, O. (1974) Stability of ribonucleic-acid double-stranded helices. *J. Mol. Biol. 86*, 843–853.
44. Marky, L. A., and Breslauer, K. J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers 26*, 1601–1620.
45. Shankar, N., Xia, T., Kennedy, S. D., Krugh, T. R., Mathews, D. H., and Turner, D. H. (2007) NMR reveals the absence of hydrogen bonding in adjacent UU and AG mismatches in an isolated internal loop from ribosomal RNA. *Biochemistry 46*, 12665–12678.

BI7020876