3′ Terminal Nucleotides Determine Thermodynamic Stabilities of Mismatches at the Ends of RNA Helices[†]

Koree Clanton-Arrowood,[‡] John McGurk,[‡] and Susan J. Schroeder*

Department of Chemistry and Biochemistry, University of Oklahoma, 620 Parrington Oval, Norman, Oklahoma 73019

Received August 25, 2008; Revised Manuscript Received October 25, 2008

ABSTRACT: The thermodynamic stabilities of consecutive mismatches at the ends of RNA helices are determined by the 3′ terminal nucleotides. More than 40 RNA duplexes containing terminal motifs of 3 or more nucleotides were studied by optical melting experiments. Up to three noncanonical pairs of nucleotides at the end of RNA helices provide additional thermodynamic stability. 3′ nucleotides contribute more stability than 5′ nucleotides, and purines contribute more stability than pyrimidines. The additional stability of a second or third 3′ nucleotide stacking on a purine is the same for both dangling ends and consecutive terminal mismatches. Current predictions underestimate RNA duplex stabilities with terminal motifs by 1.4 kcal/mol on average, which is an order of magnitude in a binding constant at 37 °C. Accurate thermodynamic parameters for these terminal motifs will contribute to improvements in RNA secondary structure predictions, identification of microRNA targets, and design of siRNA therapeutics with fewer off-target effects.

The binding of a small RNA to an mRNA target is intrinsic to RNA interference (RNAi) (*1*−*4*). Currently off-target effects limit the implementation of RNAi[1] therapeutics (*5*−*7*). Fully realizing the potential of RNAi to cure human disease will require an understanding of the stability and specificity of RNA interactions and RNA secondary structure (*8*−*11*). This paper presents measurements of the thermodynamic stabilities of the terminal mismatch motifs common in siRNA and miRNA interactions and an improved method to estimate the stabilities of these RNA motifs. These results can be incorporated into the database of thermodynamic parameters that form the core of RNA prediction programs such as mfold, RNAStructure, Vienna RNA Web Suite, RNAStar, and SFOLD (*12*−*16*).

The thermodynamic parameters presented here will contribute to improvements in the estimates of siRNA:mRNA interactions and target accessibility. The thermodynamic stability of the RNA duplexes is a critical factor in several steps of the RNA interference phenomenon, including target recognition, 5′ seed region formation, determining mRNA cleavage or translation inhibition pathways, and product release (*17*−*21*). Although protein−RNA interactions are an important factor in siRNA phenomena, the stabilities of different RNA duplexes are necessary to explain the changes in RNA regulation efficiency due to synonymous substitutions in the miRNA target sequence, for example (*22*). The new data and prediction methods presented in this paper improve the estimates for a set RNA duplex stabilities that correlate with siRNA efficacy in controlled *in vitro* assays (*18*).

Thermodynamic stability and secondary structure of the mRNA target are important factors for effective RNA therapeutics (*9, 10, 17, 23*−*25*). Possible conformational changes in the mRNA secondary structure, such as the possibilities considered by programs such as Oligowalk and siRNAfold (*8, 11, 25, 26*), can affect RNAi efficacy and specificity. For example, changes in RNA secondary structure are one strategy of HIV viral resistance to RNAi therapeutics (*10*). Thus, accurate thermodynamic parameters and RNA secondary structure prediction are essential for design of siRNA therapeutics and target selection.

Despite much research, the thermodynamic parameters for the formation of noncanonical motifs and the sequence dependence of irregular structures remain incomplete (*15, 27*). Previous research on the sequence dependence of thermodynamic parameters for a variety of loops in RNA shows a wide range of stabilities for noncanonical regions of RNA (*27*−*42*). Incorporating more accurate thermodynamic parameters for loop motifs has improved the overall prediction of RNA secondary structure (*15, 27*). Dangling ends are the nucleotides that occur at one end of an RNA helix. The thermodynamic stability of a single dangling end nucleotide on the 3′ end of the helix is more stable than on the 5′ end (*43*−*47*). This additional stability extends to 4 single nucleotides (*45*). The sequence dependence of these single dangling ends continues to two nucleotides (*48, 49*). Terminal mismatches are a pair of noncanonical nucleotides at the end of one RNA helix. Current thermodynamic models do not include any additional stability beyond a single terminal mismatch. This research shows that more than one consecutive terminal mismatch and dangling ends extending from terminal mismatches provide additional thermodynamic stability.

[†] This work was supported by grants from the Oklahoma Center for the Advancement of Science and Technology and the Pharmaceutical Research and Manufacturers of America Foundation.

* To whom correspondence should be addressed. E-mail: susan.schroeder@ou.edu. Phone: (405) 325-3092. Fax: (405) 325-6111.

[‡] These authors contributed equally to this work.

[1] Abbreviations: RNAi, RNA interference; siRNA, small interfering RNA; miRNA, microRNA.

10.1021/bi801594k CCC: $40.75
Published on Web 11/17/2008

EXPERIMENTAL PROCEDURES

Oligoribonucleotides were purchased from Dharmacon and prepared according to the manufacturer's instructions. Greater than 90% purity was confirmed by radioactive ^{32}P 5′ labeling and gel electrophoresis. If necessary, oligomers were purified by thin layer chromatography as previously described (*28*).

Oligomer concentration was determined from the UV absorbance at 280 nm at high temperature (*50*). The RNA was dissolved in 1 M NaCl, 10 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 7 melting buffer. Some oligomer duplexes formed aggregates in high salt and were melted in 10 mM NaCl. Motifs that contain the potential to form protonated A^+C or C^+C base pairs were also melted in pH 5 melting buffer. Melting curves were measured at 280 nm on a Beckman DU800 spectrometer with a modified microT_m cell holder and a heating rate of 1 °C/min. The melting curves were fit to a two-state model with sloping baselines (*51*) and analyzed with the Meltwin program (*52*). Thermodynamic parameters were obtained by fitting plots of inverse melting temperature (T_M^{-1}) versus the natural log of the total strand concentration (C_T) to the equation

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

where R is the gas constant (*53*).

Imino proton NMR were acquired with a 500 MHz Varian VNMR spectrometer and console in order to verify the expected duplex formation and test for possible hydrogen-bonded base pairs in terminal nucleotides. RNA samples for NMR were dissolved in 10 mM NaCl, 10 mM sodium phosphate, 0.5 mM Na_2EDTA, pH 6 to at least 0.15 mM concentration.

Linear regression analysis was done using the linest function in Excel software and includes data for all previous terminal motifs (*36, 43, 45, 48, 49, 54*−*57*) as well as the new data presented in this paper. Statistical significance was calculated using a Student's t test in Excel.

RESULTS

Table 1 lists thermodynamic parameters derived from melting experiments for over 40 duplexes containing dangling ends and terminal mismatches. The duplexes are grouped by number of terminal nucleotides and then in order of decreasing stability of the terminal motif. The number of 5′ nucleotides is listed first followed by an x and then the number of 3′ nucleotides. For example, a 2x3 motif has two 5′ nucleotides opposite three 3′ nucleotides extending from the helix. The free energy of the terminal motif is calculated according to the following equation:

$$\Delta G^\circ_{\text{terminal motif}} = [\Delta G^\circ_{\text{duplex with terminal motif}} - \Delta G^\circ_{\text{duplex without terminal motif}}]/2 \quad (2)$$

For example, in this 2x3 motif,

$$\Delta G^\circ \begin{matrix} 5'\,\mathrm{C}\mathbf{AAA} \\ 3'\,\mathbf{GAA} \end{matrix} = [\Delta G^\circ \begin{matrix} 5'\,\mathbf{AA}\mathrm{GCGC}\mathbf{AAA}3' \\ 3'\,\mathbf{AAA}\mathrm{CGCG}\mathbf{AA}5' \end{matrix} - \Delta G^\circ \begin{matrix} 5'\,\mathrm{GCGC}3' \\ 3'\,\mathrm{CGCG}5' \end{matrix}]/2 \quad (3)$$

The $\Delta G^\circ_{\text{terminal motif}}$ value is the free energy increment attributed to the terminal mismatches and dangling ends. The values for the enthalpy and entropy changes are similarly derived and included in Supporting Information Table 1. The values for $\Delta G^\circ_{\text{duplex without terminal motif}}$ are experimentally measured values when possible or calculated using the internal nearest neighbor-hydrogen bonding (INN-HB) model (*58*). The INN-HB model implicitly includes any "end effects" for the helix without terminal mismatches in the helix initiation term. For helices in which the last base pair is an AU or GU base pair, the terminal AU penalty, which accounts for the different number of hydrogen bonds in a helix ending with GC pairs versus a helix with the same nearest-neighbors but ending with AU pairs, is included in the calculation of $\Delta G^\circ_{\text{duplex without terminal motif}}$.

Thermodynamic parameters for the duplexes with terminal motifs derived from T_M^{-1} vs $\ln(C_T)$ plots and from curve fitting agree within 15%, consistent with the two-state transition model unless otherwise noted (*59*). Sequences and data in italics have differences between enthalpies from van't Hoff plots and curve fitting slightly above but near 15%. Fourteen sequences based on naturally occurring siRNA or miRNA duplexes that have been studied *in vitro* (*18, 19, 21, 60*) are noted in Table 1. As previously observed with other RNA motifs, there is no apparent direct correlation between thermodynamic stability and natural occurrence (*29, 33, 41*). Duplexes melted in different buffers are noted in the table, but were not included in the linear regression analysis. No additional stability was observed for terminal motifs containing potential A^+C or C^+C base pairs at a lower pH that favors protonated base pair formation (*29, 35*).

Comparison of the stabilities of the same motif 5′U**CA**/3′A**AG** measured with different duplex stems in different sodium chloride concentrations, 1.9 and 1.0 kcal/mol at 1 M and 10 mM NaCl respectively, demonstrates that salt concentrations can have a significant effect on terminal motif stability. The salt-dependent difference in stability per nearest neighbor in the terminal mismatches, 0.45 kcal/mol, is similar in magnitude to the difference in stability per nearest neighbor for a Watson−Crick duplex melted at 1 M and 10 mM NaCl, 0.46 kcal/mol for the sequence 5′AGCGCU. These results are within error of recent estimates for salt dependence of an average AU base pair, although salt effects can vary for different types of secondary structure motifs (*61*).

Duplexes analyzed by imino proton NMR are noted by superscripts in Table 1. All the NMR spectra showed the correct number of peaks and chemical shifts expected for the designed duplex formation but showed no evidence for hydrogen bond formation in the terminal motifs. Imino protons in stable hydrogen-bonded base pairs may be protected from exchange with water and resonate between 9−15 ppm. The absence of an imino proton signal does not preclude hydrogen bond formation, but implies that the hydrogen bonding may be too transient or unstable to significantly affect exchange rates with water. The exceptions are a single additional low, broad resonance around 11.5 ppm for the sequences 5′**UUU**GCGC**UUU** and 5′**UGC**GAUC-**CUG**.

Table 2 shows the trends in free energy when adding either adenine or uracil nucleotides to the 3′ or 5′ end of a helix. Consistent with previous studies of single dangling ends, addition of nucleotides to the 3′ end contributes far more additional stability than 5′ dangling ends (*43*−*45*). Motifs with the same number of 3′ nucleotides have the same

Table 1: Thermodynamic Parameters of Duplex Formation[a]

				T_m^{-1} vs ln(c_t/a)				melt curve fits			
5′x3′	sequence	NaCl	$-\Delta\Delta G^\circ_{37}$[b] (kcal/mol)	$-\Delta G^\circ_{37}$[c] (kcal/mol)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m[d] (°C)	$-\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m[d] (°C)
4x4	5′***AAAA***GCGC***AAAA***	1 M	*2.3 ± 0.3*	*9.3 ± 0.2* (7.2)	*60.6 ± 3.5*	*165.6 ± 10.8*	*56.5*	*8.7 ± 0.1*	*50.5 ± 1.6*	*134.6 ± 5.2*	*57.2*
4x4	5′**GACG**AGCU**AUCA**[e]	1 M	1.6 ± 0.3	5.4 ± 0.2 (4.4)	48.2 ± 5.6	137.9 ± 18.2	35.5	5.6 ± 0.4	43.4 ± 8.3	121.8 ± 26.8	36.6
4x4	5′**GACG**AGGCCU**AUCA**[e,i]	10 mM		5.6 ± 0.1	69.3 ± 1.6	205.3 ± 5.3	36.6	5.7 ± 0.1	61.7 ± 3.3	180.7 ± 11.0	36.9
4x3	5′**AAAA**GCGC**AAA**	1 M	2.2 ± 0.2	9.1 ± 0.1 (7.2)	55.4 ± 1.7	149.3 ± 5.2	57.2	8.7 ± 0.1	48.8 ± 1.5	129.2 ± 4.8	57.6
3x4	5′**AAA**GCGC**AAAA**	1 M	2.1 ± 0.2	8.9 ± 0.1 (7.2)	57.0 ± 2.1	155.0 ± 6.4	55.6	8.7 ± 0.2	51.6 ± 2.3	138.4 ± 7.1	56.1
3x3	5′**GGA**GUAC**GAA**	1 M	2.7 ± 0.2	6.6 ± 0.1 (4.1)	58.2 ± 2.8	166.3 ± 9.0	41.9	6.6 ± 0.2	49.6 ± 4.9	138.7 ± 16.2	42.7
3x3	5′**AGA**CUAG**GAG**	1 M	2.4±0.2	5.8 ± 0.1 (4.2)	44.8 ± 3.0	125.7 ± 9.7	38.2	6.1 ± 0.3	51.8 ± 14.5	147.5 ± 46.4	39.4
3x3	5′**AAA**GCGC**AAA**	1 M	2.2±0.4	8.9 ± 0.3 (7.2)	58.3 ± 5.0	159.2 ± 15.4	55.4	8.9 ± 0.5	62.4 ± 27.5	172.7 ± 87.4	53.7
3x3	5′**UUU**GCGC**UUU**[h]	1 M	1.7 ± 0.2	7.9 ± 0.1 (6.6)	55.5 ± 2.7	153.3 ± 8.7	49.9	7.8 ± 0.1	52.1 ± 3.1	142.8 ± 10.1	50.3
3x3	5′**CUA**UCGA**GCA**	1 M	1.6 ± 0.2	4.7 ± 0.1 (3.2)	41.1 ± 1.4	117.4 ± 4.6	29.8	4.9 ± 0.3	38.7 ± 6.5	109.1 ± 21.8	30.7
3x3	5′**ACU**ACCGGU**CAG**	1 M	0.8 ± 0.3	9.8 ± 0.2 (9.6)	61.0 ± 4.2	165.1 ± 12.8	59.5	9.75 ± 0.5	57.2 ± 10.9	153.0 ± 33.8	60.8
3x2	5′**AAA**GCGC**AA**	1 M	2.4 ± 0.4	9.5 ± 0.3 (7.2)	62.3 ± 4.2	170.4 ± 13.0	57.2	9.0 ± 0.3	53.4 ± 2.7	143.2 ± 8.0	57.7
2x3	5′**AA**GCGC**AAA**	1 M	2.4 ± 0.2	9.1 ± 0.1 (7.2)	60.8 ± 2.0	166.6 ± 6.1	55.7	8.7 ± 0.2	51.7 ± 2.6	138.7 ± 8.4	56.2
2x5	5′**GA**AGCGCU**CACUU**(pH7)[e]	10 mM	1.1 ± 0.1	7.6 ± 0.1	74.8 ± 1.5	216.7 ± 4.6	46.2	7.6 ± 0.1	72.6 ± 3.9	209.6 ± 12.5	45.4
2x5	5′**GA**AGCGCU**CACUU**(pH5)[e,i]	10 mM	0.9 ± 0.1	7.3 ± 0.1	58.0 ± 2.7	163.4 ± 8.5	45.2	7.6 ± 0.3	60.1 ± 11.5	169.5 ± 36.4	47.0
2x2	5′**AA**GCGC**AA**	1 M	2.2 ± 0.3	8.9 ± 0.2 (7.2)	58.2 ± 3.8	158.9 ± 11.7	55.2	8.5 ± 0.2	50.6 ± 1.4	135.9 ± 4.6	55.3
2x2	5′**GA***UCGAGA*	*1 M*	*2.2 ± 0.2*	*5.9 ± 0.1* (3.7)	*56.9 ± 2.7*	*164.4 ± 8.6*	*38.3*	*6.0 ± 0.2*	*47.8 ± 3.8*	*135.1 ± 12.6*	*38.8*
2x2	5′**GA**CUAG**GA**[c]	1 M	2.2 ± 0.2	5.3 ± 0.1 (4.2)	48.2 ± 1.7	138.2 ± 5.5	35.1	5.5 ± 0.2	47.6 ± 7.9	135.5 ± 25.6	36.0
2x2	5′**AG**GUAC**AG**	1 M	2.2 ± 0.2	5.6 ± 0.1 (4.1)	46.8 ± 1.9	132.9 ± 6.1	36.2	5.7 ± 0.1	46.2 ± 8.7	130.7 ± 28.2	37.1
2x2	5′**CG**AGCU**AU**[e]	1 M	2.0 ± 1.4	6.1 ± 1.4 (4.1)	21.8 ± 7.0	50.3 ± 21.0	43.8	6.2 ± 1.2	25.7 ± 6.7	63.0 ± 23.1	43.3
2x2	5′**GA**AGCU**GA**[e]	1 M	1.9 ± 0.2	6.0 ± 0.1 (4.4)	48.5 ± 1.5	137.1 ± 4.7	39.0	6.1 ± 0.2	49.4 ± 8.7	139.6 ± 27.7	39.9
2x2	5′***GA****AGCUCA*[e,j]	1 M	1.9 ± 0.2	*6.0 ± 0.2* *(3.6)*	*43.9 ± 3.3*	*122.3 ± 11.0*	*38.7*	*6.0 ± 0.1*	*51.1 ± 2.0*	*145.4 ± 6.3*	*34.3*
2x2	5′**UG**UGCA**GU**[d]	1 M	2.0 ± 0.2	6.2 ± 0.1 (3.8)	63.1 ± 4.8	183.5 ± 14.1	39.7	6.2 ± 0.2	59.2 ± 2.4	170.7 ± 8.1	39.9
2x2	5′**GA**UGCA**GA**	1 M	1.7 ± 0.2	5.5 ± 0.1 (3.8)	53.2 ± 2.3	153.6 ± 7.6	*36.1*	5.6 ± 0.2	43.9 ± 2.9	123.6 ± 9.4	36.5
2x2	5′**UU**GCGC**UU**[i]	1 M	1.6 ± 0.2	7.8 ± 0.1 (6.6)	58.0 ± 2.3	161.8 ± 7.3	*48.9*	7.6 ± 0.1	50.5 ± 2.5	138.3 ± 8.1	49.0
2x2	5′**UA**UCGA**GC**	1 M	1.6 ± 0.2	4.8 ± 0.1 (3.7)	45.0 ± 2.0	129.4 ± 6.1	31.3	5.0 ± 0.2	44.3 ± 7.0	126.8 ± 23.3	32.0
2x2	5′**AA**AGCU**AA**	1 M	1.6 ± 0.2	5.1 ± 0.1 (4.2)	37.3 ± 2.0	103.8 ± 6.5	38.3	5.5 ± 0.3	38.3 ± 6.1	105.9 ± 20.3	35.3
2x2	5′**GG**AGCU**GG**	1 M	1.5 ± 0.2	5.1 ± 0.1 (4.6)	37.3 ± 2.0	103.8 ± 6.5	*32.1*	5.2 ± 0.2	42.4 ± 8.8	119.9 ± 28.4	33.4
2x2	5′**AG**UGCAA**G**	1 M	1.3 ± 0.2	4.7 ± 0.1 (3.8)	42.3 ± 2.2	121.2 ± 7.2	30.3	4.9 ± 0.2	40.0 ± 4.0	113.3 ± 13.3	30.7
2x2	5′**UU**CGCG**UC**	1M	1.2 ± 0.2	5.9 ± 0.1 (5.1)	52.4 ± 2.2	149.9 ± 7.1	38.3	5.9 ± 0.2	45.0 ± 3.0	125.9 ± 9.6	38.7
2x2	5′**UU**AGCUUU[h]	1 M	1.0 ± 0.3	4.2 ± 0.2 (3.2)	48.2 ± 3.5	142.0 ± 11.6	27.6	4.4 ± 0.4	43.3 ± 7.6	125.3 ± 25.5	28.5
2x2	5′**UU**AGCU**UC**	1 M	1.0 ± 0.2	4.2 ± 0.1 (3.2)	48.2 ± 1.6	141.9 ± 5.2	27.5	4.4 ± 0.2	42.7 ± 3.7	123.6 ± 12.2	27.9
2x2	5′**GA**AGCGCU**CA**	10 mM	1.0 ± 0.1	6.3 ± 0.1	57.6 ± 1.6	165.2 ± 5.2	35.9	7.5 ± 0.1	67.6 ± 1.8	193.7 ± 5.6	45.6
2x2	5′**GA**AGCCU**CA** 3′**AC**UCCGGAA**G**	10 mM		6.3 ± 0.1	57.6 ± 1.6	165.2 ± 5.2	35.9	6.4 ± 0.1	62.2 ± 4.2	179.9 ± 13.5	36.2
2x1	5′**AA**GCGC**A**	1 M	1.7 ± 0.2	8.0 ± 0.1 (7.2)	52.9 ± 2.1	144.6 ± 6.4	51.7	7.7 ± 0.1	45.5 ± 1.1	121.6 ± 3.8	51.7
2x1	5′**UU**GCGC**U**	1 M	1.2 ± 0.2	7.0 ± 0.1 (6.6)	49.6 ± 1.8	137.3 ± 5.8	45.3	6.9 ± 0.1	44.7 ± 2.0	121.9 ± 7.3	45.6
1x2	5′**G**GUAC**AG**	1 M	2.1 ± 0.2	5.4 ± 0.1 (4.1)	41.3 ± 2.7	115.8 ± 8.7	34.8	5.5 ± 0.2	45.3 ± 9.2	128.3 ± 30.0	35.5
1x2	5′**A**GCGC**AA**	1 M	2.0 ± 0.2	8.6 ± 0.1 (7.2)	57.8 ± 2.7	158.6 ± 8.4	53.8	8.3 ± 0.1	50.2 ± 2.0	135.0 ± 6.6	53.9

Table 1: (Continued)

				T_m^{-1} vs $\ln(c_t/a)$				melt curve fits			
5′x3′	sequence	NaCl	$-\Delta\Delta G°_{37}$[b] (kcal/mol)	$-\Delta G°_{37}$[c] (kcal/mol)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	T_m[d] (°C)	$-\Delta G°_{37}$ (kcal/mol)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	T_m[d] (°C)
1x2	5′**A**CUAG**GA**	1 M	1.9 ± 0.2	4.8 ± 0.1 (3.8)	43.9 ± 2.2	126.3 ± 7.4	30.6	4.8 ± 0.2	46.6 ± 7.8	134.7 ± 25.8	31.5
1x2	5′**U**GCGC**UU**	1 M	1.7 ± 0.2	7.8 ± 0.1 (6.6)	58.4 ± 2.0	163.1 ± 6.2	48.8	7.6 ± 0.1	52.1 ± 2.3	143.5 ± 7.5	48.9
1x2	5′**A**UGCAG**G**[f]	1 M	1.3 ± 0.2	4.8 ± 0.1 (3.8)	39.2 ± 1.8	110.7 ± 6.1	30.2	5.0 ± 0.2	37.5 ± 6.1	104.9 ± 20.2	31.3
1x1	5′**A**GCUAGU**G**	1 M	1.2 ± 0.3	7.4 ± 0.1 (7.3)	73.9 ± 2.0	214.2 ± 6.3	44.5	7.2 ± 0.1	62.3 ± 1.7	177.6 ± 5.8	44.7
1x1	5′**A**GGCU**G**[d]	1 M	1.2 ± 0.4	4.5 ± 0.2 (4.4)	49.2 ± 4.9	144.0 ± 16.1	30.2	4.8 ± 0.3	42.0 ± 5.4	119.9 ± 17.8	30.6
0x1	5′**U**UGCA**GU**[f]	1 M	0.8 ± 0.3	4.9 ± 0.1 (4.5)	42.2 ± 2.0	120.5 ± 6.4	31.2	5.0 ± 0.2	40.0 ± 3.0	112.8 ± 10.0	31.6
0x0	5′**GU**UGCA**GU**[d,i]	1 M	1.6 ± 0.2	5.4 ± 0.1 (3.3)	45.2 ± 3.7	128.3 ± 12.1	35.2	5.5 ± 0.2	46.2 ± 5.2	131.3 ± 16.7	35.6
0x0	5′**UGC**GAUC**CUG**[g,i]	1 M	2.9 ± 0.3	7.0 ± 0.2 (3.5)	69.5 ± 6.9	201.6 ± 21.9	57.2	6.9 ± 0.6	60.3 ± 17.7	172.3 ± 55.2	43.4
0x0	5′AGCGCU	1 M		7.8 ± 0.1 (7.9)	49.5 ± 2.5	134.5 ± 7.9	50.6	7.9 ± 0.2	50.8 ± 4.2	138.5 ± 13.1	51.0
0x0	5′AGCGCU	10 mM		5.5 ± 0.1	42.3 ± 0.9	118.6 ± 2.9	35.7	5.6 ± 0.2	46.3 ± 3.9	131.3 ± 12.2	36.3

[a] Melting experiments were done in 1 M NaCl, 10 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 7.0 buffer solutions unless otherwise noted. Sequences are listed by size of terminal motif and then decreasing order of terminal motif stability. Sequences and data in italics have differences in enthalpies from van't Hoff plots and curve fitting near, but slightly above, 15%. Listed errors are standard deviations from reported measurements assuming no correlation of errors in the slope and intercept and are therefore overestimates of this source of error. Estimated errors from all sources are ± 10%, ± 10%, and ± 2% for $\Delta H°$, ΔS °, and $\Delta G°$, respectively. [b] Values are calculated from T_m^{-1} vs $\ln(c_t/a)$ plots using the following equation: $\Delta G°$terminal motif = [$\Delta G°$duplex with terminal motif − $\Delta G°$duplex without terminal motif]/2. [c] Values in parentheses are predicted stabilities of the entire duplex using the rules currently implemented in prediction programs such as mfold (*85*), RNAstructure (*15*) and Vienna websuite (*16*). The duplex stabilities are predicted using nearest-neighbor parameters (*58*) and rules for single terminal mismatches and dangling ends (*15, 43, 54, 56, 57*). [d] Melting temperatures are given for total strand concentration of 1 × 10−4 M. [e] Terminal motif sequence based on naturally occurring miRNA or siRNA sequences from ref *19*, [f] Terminal motif sequence based on naturally occurring miRNA or siRNA sequences from ref *21*, [g] Terminal motif sequence based on naturally occurring miRNA or siRNA sequences from ref *60*, [h] Terminal motif sequence based on naturally occurring miRNA or siRNA sequences from ref *18*. [i] Sequences were analyzed by one-dimensional imino proton NMR. [j] Duplex was melted in pH 5 buffer with no significant change in thermodynamic parameters.

Table 2: Incremental Stabilities for Additional Purine or Pyrimidine Sequences[a]

5′x3′		$-\Delta\Delta G°$ (kcal/mol)	5′x3′		$-\Delta\Delta G°$ (kcal/mol)	5′x3′		$-\Delta\Delta G°$ (kcal/mol)			$-\Delta\Delta G°$ (kcal/mol)	5′x3′		$-\Delta\Delta G°$ (kcal/mol)
0x1	5′CA 3′G	–1.5[b]				1x1	5′CA ↔ 5′AG 3′GA 3′AC	–1.5[b]				1x0	5′AG 3′C	–0.3[b]
0x2	5′CAA 3′G	–1.9[b]	1x2	5′CAA 3′GA	–2.0	2x2	5′CAA ↔ 5′AAG 3′GAA 3′AAC	–2.0	2x1	5′AAG 3′AC	–1.7	2x0	5′AAG 3′C	–0.5[b]
0x3	5′CAAA 3′G	–2.4[b]	2x3	5′CAAA 3′GAA	*–2.4*	3x3	5′CAAA ↔ 5′AAAG 3′GAAA 3′AAAC	–2.2	3x2	5′AAAG 3′AAC	–2.4	3x0	5′AAAG 3′C	–0.8[b]
0x4	5′CAAAA 3′G	–2.5[b]	3x4	5′CAAAA 3′GAAA	–2.1	4x4	5′CAAAA ↔ 5′AAAAG 3′GAAAA 3′AAAAC	*–2.3*	4x3	5′AAAAG 3′AAAC	–2.2	4x0	5′AAAAG 3′C	–0.9[b]
0x1	5′CU 3′G	–1.2[d]				1x1	5′CU ↔ 5′UG 3′GU 3′UC	–1.2[d]				1x0	5′UG 3′C	–0.1[d]
0x2	5′CUU 3′G	–1.3[c]	1x2	5′CUU 3′GU	–1.6	2x2	5′CUU ↔ 5′UUG 3′GUU 3′UUC	–1.6	2x1	5′UUG 3′UC	–1.2			
						3x3	5′CUUU ↔ 5′UUUG 3′GUUU 3′UUUC	–1.7						

[a] Melting experiments were done in 1 M NaCl, 10 mM sodium cacodylate, 0.5 mM Na2EDTA, pH 7.0 buffer solutions. [b] Values are from ref *45*. [c] Values are from ref *48*. [d] Values are from ref *62*.

stability regardless of the number of 5′ nucleotides. For example 0x1, 1x1, and 2x1 motifs with adenine nucleotides have stabilities of −1.5, −1.5, and −1.7 kcal/mol, respectively. 0x2, 1x2, and 2x2 motifs with adenine nucleotides have stabilities of −1.9, −2.0, and −2.1 kcal/mol, respectively. Extra 5′ nucleotides provide additional stability only as dangling ends without any opposing 3′ nucleotides. For example, the free energies of 1x0, 2x0, 3x0, and 4x0 motifs with adenines are −0.3, −0.5, −0.8, and −0.9 kcal/mol respectively (*45*). The additional stability of adenines tapers off after three nucleotides; for example, the stabilities of 3x3 and 4x4 motifs with adenines are −2.2 and −2.3 kcal/mol, respectively. The additional stability of uridines tapers off sooner after the addition of two nucleotides; for example, the stabilities of 2x2 and 3x3 motifs with uridines are −1.6 and −1.7 kcal/mol, respectively. The amount of additional stability of a 3′ nucleotide stacking on a purine is the same for both dangling ends and consecutive mismatches.

Parameters for predicting terminal motifs (Table 3, Figure 1) were derived from a multiple linear regression analysis

Table 3: Parameters for Terminal Motif Prediction[a]

Table 3: Parameters for Terminal Motif Prediction[a]			
0x1	kcal/mol	**1x1**	kcal/mol
5'C**R** 3'G	-1.6 ± 0.1	5'C**R** = 5'G**R** 3'G**N** 3'C**N**	-1.5 ± 0.1
5'G**R** 3'C	-1.0 ± 0.1		
5'C**Y** 3'G	-1.1 ± 0.1	5'C**Y** = 5'G**Y** 3'G**N** 3'C**N**	-1.2 ± 0.1
5'G**Y** 3'C	-0.5 ± 0.1		
5'A**N** =5'U**N** 3'U 3'A	-0.6 ± 0.1	5'A**R** = 5'U**R** 3'U**N** 3'A**N**	-1.0 ± 0.1
		5'A**Y** = 5'U**Y** 3'U**N** 3'A**N**	-0.7 ± 0.1
Parameters for Nucleotides Beyond the First 3' Dangling End or Mismatch			
	kcal/mol		
2nd 3'nucleotide	-0.6 ± 0.1		
3rd 3'nucleotide	-0.2 ± 0.1		
5'CUU bonus 3'GU	-0.4 ± 0.1		
5'CGA bonus 3'GAG	-0.4 ± 0.2		
5' nucleotide	-0.2 ± 0.1		

[a] Values are derived from a multiple linear regression analysis of 187 experimentally measured free energies for terminal motifs measured in 1 M NaCl, 10 mM cacodylate, 0.5 mM Na_2EDTA, pH 7.0. R is an abbreviation for purine nucleotides A or G. Y is an abbreviation for pyrimidine nucleotides C or U. N represents any nucleotide that does not form a Watson−Crick pair. The terms for the 2nd and 3rd 3′ nucleotide, which may be either a purine or a pyrimidine, are added only if the additional nucleotide *stacks* on a purine. The term for 5′ nucleotides is added for each 5′ dangling nucleotide beyond the number of 3′ dangling nucleotides.

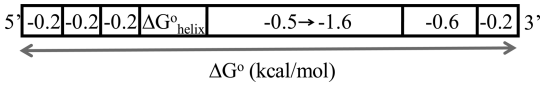

FIGURE 1: Additional stabilities of terminal nucleotides. Each box represents one nucleotide added onto the helix. The first 3′ nucleotide shows a wide range of stabilities determined by the closing pair, nucleotide identity, and whether the nucleotide is a dangling end or part of a noncanonical pair. The stabilities of the second and third 3′ nucleotides are independent of these factors and are added only when stacking on a purine nucleotide. The stability of 5′ nucleotides are added only as a dangling end but not in the context of terminal mismatches.

of 187 terminal motifs from this work and previous literature (*36, 43, 45, 48, 49, 54−57, 62*). The first 3′ nucleotide stacking on the last Watson−Crick base pair is the most significant determinant of sequence dependence for all terminal motifs. Dangling ends and terminal mismatches show a different sequence dependence pattern for the first 3′ nucleotide. Additional stability for subsequent 3′ nucleotides is the same for both dangling ends and terminal mismatches. Terminal motifs containing only 3′ dangling ends, i.e. 0x1 or 0x2 terminal motifs, show a dependence on the orientation of a CG closing base pair and a wide range of stabilities from −0.4 to −2.5 kcal/mol (Supporting Information Table 2). Thus four terms describe the stacking effects of a purine or a pyrimidine on each possible CG closing base pair orientation. A much narrower range of free energies is observed when a single 3′ nucleotide stacks on an AU closing pair. Thus only one term is necessary to describe this interaction. Terminal mismatches, or 1x1 motifs, show less dependence on the orientation of CG or AU closing base pairs and a narrow range of stabilities, −0.8 to −1.5 and −0.5 to −1.2 kcal.mol, respectively (Supporting Information Table 2). Terminal mismatch (1x1) stability mainly depends on whether the 3′ stacking nucleotide is a pyrimidine or purine. Thus, 4 terms describe the stacking effects of a 3′ purine or pyrimidine nucleotide on either CG or AU closing base pairs. As the 3′ nucleotides extend farther, additional sequence dependence is seen only when the nucleotides stack on a purine. This is consistent with previous studies of 0x2 motifs (*48*). A single sequence-independent term adds a bonus for each 5′ dangling end beyond the number of 3′ nucleotides. Two bonus terms are included to describe the two of the exceptions to the overall trends in terminal motif stability. The bonus for 5′**CUU**/3′**GU** and 5′ C**GA**/3′G**AG** motifs are statistically significant in the linear regression analysis, although there are only 3 or 2 occurrences, respectively. While some parameters have similar values, combining these values into a single parameter predicts the experimental data more poorly and is less consistent with a physical model of stacking interactions. These 14 parameters accurately predict the values of all 187 measured terminal motifs within experimental error. The R^2 value for the linear regression is 0.96.

Table 4 compares the free energies of formation of terminal motifs, internal loops, and hairpin loops of adenine nucleotides. Terminal motifs have favorable energies of formation unlike hairpin loops or most internal loops. The change in free energy of loop formation when adding a nucleotide to a loop is approximately 0.5 kcal/mol or less for all types of loops with 3 or more nucleotides except when adding a fourth nucleotide to a hairpin loop, which reduces significant backbone strain. The unfavorable stability of internal and hairpin loops results mainly from the entropic cost of closing a loop.

Table 5 shows a comparison of predicted thermodynamic stabilities for the siRNA-mRNA pairing in the chemokine receptor 4 (CXCR4) gene (*18*). Free energies are calculated for the siRNA 5′ seed region using thermodynamic parameters in mfold as reported by Doench and Sharp and new parameters presented here and by Mathews et al. (*15*). When more than one base pairing pattern could occur as a result of the mutation, the lowest free energy conformation was chosen. Accurate thermodynamic parameters for noncanonical motifs have a significant impact on estimates of siRNA-mRNA duplex stabilities. The average change in free energies for the mutants is 1.5 kcal/mol. This subset of mutants analyzed by Doench and Sharp shows the overall correlation between thermodynamic stability of the siRNA 5′ seed region and translational repression in a firefly luciferase reporter assay. This is one example of how accurate thermodynamic parameters can contribute to understanding the physical basis of RNAi specificity.

DISCUSSION

3′ Terminal Purine Nucleotides Provide the Most Additional Thermodynamic Stability at the Ends of Helices. 3′ nucleotides provide more additional stability than 5′ nucleotides as dangling ends (*44, 45, 57*), terminal mismatches (*54*), and consecutive terminal mismatches. This effect is most likely a consequence of the geometry of stacking interaction at the 3′ and 5′ ends of the helix (*62, 63*). A 3′ dangling purine explores more stacked conformations more often than a 5′ dangling purine (*44*). Purine nucleotides bury more surface area when stacked on a helix (*64*). Thus, the

Table 4: Comparison of the Free Energies of Formation for Terminal, Internal, and Hairpin Loops[a]

N	terminal loop	ΔG° (kcal/mol)	internal loop	ΔG° (kcal/mol)	$\Delta\Delta G^\circ$ (kcal/mol)	hairpin loop	ΔG° (kcal/mol)	$\Delta\Delta G^\circ$ (kcal/mol)
3	5′C**AA** 3′G**A**	–2.0	5′C**AA**C 3′G **A** G	2.5[b]	4.5	5′C$^{\mathbf{A}}$**A** 3′G**A**	5.4[d]	7.4
4	5′C**AA** 3′G**AA**	–2.2	5′C**AA**C 3′G**AA**G	2.0[b]	4.2	5′C$^{\mathbf{A}}$**A** 3′G**AA**	4.1[d]	6.3
5	5′C**AAA** 3′G**AA**	–2.4	5′C**AAA**C 3′G **AA** G	2.5[b]	4.9	5′C$^{\mathbf{A}}$**A**$_{\mathbf{A}}$ 3′G**AA**	4.0[d]	6.4
6	5′C**AAA** 3′G**AAA**	–2.2	5′C**AAA**C 3′G**AAA**G	2.4[b]	4.6	5′C$^{\mathbf{A}}$**A**$_{\mathbf{A}}$ 3′G**AA**$^{\mathbf{A}}$	(3.9)	(6.1)
7	5′C**AAAA** 3′G**AAA**	–2.2	5′C**AAAA**C 3′G **AAA** G	(2.7)[c]	(4.9)	5′C**AAA**$_{\mathbf{A}}$ 3′G**AAA**	4.5[d]	6.7
8	5′C**AAAA** 3′G**AAAA**	*–2.3*	5′C**AAAA**C 3′G**AAAA**G	(2.7)[c]	(5.0)	5′C**AAA**$_{\mathbf{A}}$ 3′G**AAA**$^{\mathbf{A}}$	(4.0)	(6.3)
				average	4.7 ± 0.3		average	6.5 ± 0.5

[a] Melting experiments were done in 1 M NaCl, 10 mM sodium cacodylate, 0.5 mM Na2EDTA, pH 7.0 buffer solutions. *N* is the number of nucleotides in the loop. Parameters in parentheses are predicted numbers using rules and parameters described in ref *15*. ΔG° is the free energy for formation of the loop motif. $\Delta\Delta G^\circ$ is the free energy difference between a closed loop, either an internal loop or a hairpin loop, and the terminal loop motif. [b] Values are from ref *73* and adjusted for new nearest neighbor parameters (*58*). [c] Values are from ref *32*. [d] Values are from ref *38* and adjusted for new nearest neighbor parameters (*58*).

free energies of stacking are larger for purines than pyrimidines and may show more sequence dependence. Enhanced stability from additional terminal nucleotides requires stacking on a purine nucleotide. The lack of sequence dependence in larger terminal motifs suggests that stacking interactions dominate and hydrogen bonds contribute very little additional stability. Although little direct evidence for hydrogen bonding was observed in terminal motifs, terminal nucleotides may be sampling many conformations and forming transient hydrogen bond interactions.

The new prediction rules in Table 3 provide a method for estimating the thermodynamic stabilities of terminal motifs with fewer numbers than the many tables previously compiled. Several terms previously listed separately are now described by one term, for example a single 3′ dangling nucleotide on an AU or UA closing base pair. Helices with CG terminal base pairs show more sequence dependence because the CG base pairs have a larger dipole moment than AU pairs and stronger stacking interactions. The dipole moments for C and G bases and A and U bases are 6 and 3 D, respectively (*65*). The prediction rules are consistent with the results from previous measurements of terminal motifs (*45, 48, 49, 54, 57, 63, 66*). Previous studies on a subset of loop sequences have generated accurate estimates and prediction rules for later thermodynamic studies (*28, 29, 32, 34*). Thus, these new measurements and prediction rules should similarly improve future predictions of RNA thermodynamic stability.

The three exceptions to the trends in terminal motif stability are 5′C**UU**/3′G**U**, and 5′C**GA**/G**AG** and 5′U**CA**/3′A**AG**. These three motifs show sequence dependence that varies with the nucleotides in the 5′ position of the mismatch unlike other terminal motifs studied. 5′C**UU**/3′G**U,** 5′C**UU**/3′G**UU** and 5′C**UUU**/3′G**UUU**, −1.6, −1.6, and −1.7 kcal/mol respectively, are more stable than 5′C**UU**/3′G or 5′C**U**/3′G**U**, −1.2 and −1.3 kcal/mol, respectively. Other 2x2 terminal motifs containing U nucleotides, 5′G**UC**/3′C**UU**, 5′U**UC**/3′A**UU**, or 5′U**UU**/3′A**UU**, have stabilities of −1.0 kcal/mol (Table 1). The imino proton NMR of 5′C**UUU**/3′G**UUU** shows one weak upfield peak that could possibly result from hydrogen bonding in a UU pair. Perhaps the particular orientation of the CG terminal Watson−Crick pair favors stacking and hydrogen bonding in UU pairs such that this motif has a greater population of conformations with hydrogen bonds. Thus, the particular closing base pair, UU pairs, and potential hydrogen bonding in motifs with 5′C**UU**/3′G**U** rationalize the bonus term. UU pairs in other contexts such as internal loops and hairpin loops also show additional thermodynamic stability (*28, 29, 35, 67*). Similarly, the motif 5′C**GA**/3′G**AG** and 5′C**GAA**/3′G**AGG,** −2.4 and −2.7 kcal/mol, are more stable than 5′C**GA**/3′G or 5′C**G**/3′G**A**, −2.1 and −1.4 kcal/mol respectively (Supporting Information Table 2). These sequences are also found in the most stable 2x2 and 3x3 internal loops (*31, 36*), and are thus given a bonus also in terminal motifs. 5′U**CA**/3′A**AG** is much more stable, −1.9 kcal/mol, than 5′U**CA**/3′A or 5′U**C**/3′A**A**, −0.7 kcal/mol each. This motif, 5′U**CA**/3′A**AG,** also has the possibility of forming a Watson Crick base pair with a single nucleotide bulge and a single 3′ dangling adenine. This alternate conformation may explain its anomalous stability; thus this motif was not included in the linear regression analysis. This motif shows a large salt dependence but no change in stability between pH 7 and pH 5, which favors the formation of protonated A^+C mismatches. The 5′U**CA**/3′A**AG** motif did not show exceptional stability in other contexts such as 5′U**AUCA**/3′A**GCAG** or 5′U**CACUU**/3′A**AG**, and thus was not given a separate bonus term. These exceptions occur in a total of 6 measured terminal motifs; the other 182 terminal motif measurements are predicted within experimental error by the parameters in Table 3.

The Extent of Additional Stability in Terminal Motifs Correlates with Measurements of RNA Persistence Length. The additional stability of terminal nucleotides tapers off after three 3′ purines or two 3′ pyrimidines, which is consistent with experimental measurements and theoretical estimates of persistence length in RNA (*68*−*70*). The persistence length is a measurement of the stiffness and flexibility of a polymer. In single molecule pulling experiments on polyA for example, the persistence length was measured to be 1.8 nm with a 0.6 nm phosphate-phosphate distance, or approximately 3 nucleotides (*69*). In optical melting experiments, the thermodynamic stabil-

Table 5: Correlation of 5′siRNA Thermodynamic Stability and Translational Repression[a]

duplex	fold repress	previous ΔG^{o} (kcal/mol)	new ΔG^{o} (kcal/mol)
5′AAGUUUUCAC^{AA} **AGCUAACAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UCGAUUGU**	11.9	–9.1	–10.4
5′AAGUUUUCAC^{AA} **AGCUAAC*CA*AUC…** 3′ TTCAAAAGUG$_{AGG}$**UCGAUUG*U***	12.8	–8.2	–8.8
5′AAGUUUUCAC^{AA} **AGCUAA*GU*AUC…** 3′ TTCAAAAGUG$_{AGG}$**UCGAUU*GU***	11.2	–5.7	–6.8
5′AAGUUUUCAC^{AA} **AGC^{C}UAACAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UCG AUUGU**	10.6	–5.3	–7.9
5′AAGUUUUCAC^{AA} **AGCU^{C}AACAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UCGA UUGU**	10.4	–5.3	–6.6
5′AAGUUUUCAC^{AA} **AGCUA^{C}ACAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UCGAU UGU**	8.9	–5.3	–6.6
5′AAGUUUUCAC^{AA} **AGCUA*U*CAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UCGAU*U*GU**	8.9	–5.0	–6.4
5′AAGUUUUCAC^{AA} **AG*G*UAACAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UC*G*AUUGU**	7.8	–4.6	–6.3
5′AAGUUUUCAC^{AA} **AGCUA*CA*AUC…** 3′ TTCAAAAGUG$_{AGG}$**UCGAU*UGU***	7.5	–4.5	–6.1
5′AAGUUUUCAC^{AA} **AGC*A*AACAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UCG*A*UUGU**	7.4	–4.6	–5.8
5′AAGUUUUCAC^{AA} **AGCUA^{G}ACAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UCGAU UGU**	7.3	–5.4	–6.6
5′AAGUUUUCAC^{AA} **AGCU*U*ACAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UCGA*U*UGU**	7.1	–5.1	–6.2
5′AAGUUUUCAC^{AA} **AGCU*UU*CAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UCGA*UU*GU**	6	–2.8	–5.6
5′AAGUUUUCAC^{AA} **AG*GA*AACAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UC*GA*UUGU**	5.3	–1.2	–3.1
5′AAGUUUUCAC^{AA} **AGC AACAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UCG$_{A}$UUGU**	3.9	–4.0	–5.3
5′AAGUUUUCAC^{AA} **AG UAACAAUC…** 3′ TTCAAAAGUG$_{AGG}$**UC$_{G}$AUUGU**	3.5	–2.5	–3.3

[a] The duplexes represent the interaction between the CXCR4 siRNA and the 3′UTR in the reporter firefly luciferase gene. Mutations are shown in italics. The fold repression of the gene and the first column of free energies calculated using mfold (*13*) were reported by Doench and Sharp (*18*). Free energy is calculated for the 5′siRNA seed region represented by nucleotides in bold. New thermodynamic parameters are calculated using the new data presented in this work and parameters in Mathews et al. (*15*). Errors in free energies are estimated to be ±0.2 kcal/mol. The r^2 value for the correlation between thermodynamic stability of 5′siRNA seed region and fold repression is 0.699 and 0.739 with the previous and new thermodynamic parameters, respectively.

ity is calculated from UV absorbance measurements that monitor the transition from an initial duplex conformation at low temperature to the final single-strand conformations at high temperature. At high temperatures, entropy dominates and any stacking interactions are likely to be only transiently sampled. At low temperatures, single strand adenines adopt a stacked conformation (*71*). More than three nucleotides extended from a helix may stack but provide no measureable stability difference between stacked and single strand conformations. In polyA, stacking is cooperative. The Zimm−Bragg cooperativity factor for polyA is 0.58 ± 0.15 measured by single molecule pulling experiments (*69*), which is consistent with previous measurements of cooperativity in polyA by calorimetric and spectroscopic methods (*72*). Propagating a helix stack is more favorable than starting a helix stack (*72*). In the case of terminal motifs containing adenines, the core Watson−Crick duplex has started the helix stack, which makes adenine stacking more favorable. The next 3′ adenine is also more likely to stack on an extended helix, and this effect apparently continues for approximately the persistence length of polyA.

The Stabilities of Terminal Motifs Provide a Benchmark for Estimates of the Entropic Cost of Closing Loops. Unlike hairpin loops or most internal loops, the free energies of terminal motifs are energetically favorable. At the ends of helices, nucleotides are flexible and can adopt many possible conformations with stacking interactions and possibly hydrogen bonding interactions. When the nucleotides are entropically restrained in internal loops or hairpin loops, large unfavorable penalties result from straining the phosphodiester backbone and limiting the possible conformations of the

bases and backbone angles. When comparing loops of adenine nucleotides in terminal motifs (Table 4, Supporting Information Table 3) versus hairpin loops (*38*) or internal loops (*32, 73*), the average free energy difference is 6.5 ± 0.5 kcal/mol and 4.7 ± 0.3 kcal/mol, respectively. Comparing sequences with all adenines minimizes sequence-dependent variations in stacking interactions. Although adenines can form hydrogen-bonded purine-purine pairs, the additional stability of adenine pairs is typically less than GA or UU pairs, which contribute additional, context-dependent thermodynamic stability to loops (for example refs *28, 35, 37*). The additional energy for increasing the number of nucleotides in a loop is small in all types of loops, but each type of loop has a different energy for initiating the loop. The free energy of forming a hairpin loop includes the helix initiation entropic penalty for a unimolecular unfolding reaction. The penalty for helix initiation for a bimolecular duplex is 4.09 kcal/mol in the INN-HB model (*58*), a smaller penalty than initiation of a unimolecular hairpin loop. The additional penalty for forming a hairpin loop beyond helix initiation partially reflects the entropic cost of restraining the nucleotides and backbone angles in a loop and the break in stacking interactions necessary to make a 180° turn. The free energy for initiating an internal loop includes contributions from the stacking and hydrogen bonding interactions in the closing base pairs. This may partly explain the larger sequence dependence in internal loops (3.6 to −2.6 kcal/mol, Supporting Information Table 3) than terminal motifs (0.0 to −2.7 kcal/mol, Supporting Information Table 3) and the wide range of differences in stabilities when comparing terminal and internal loops of different sequences (6.1 to −0.6 kcal/mol, Supporting Information Table 3). The dependence of the loop stability on the orientation of the closing base pairs suggests some enthalpic stacking contributions to loop formation. For example 5′CAAC/3′GAAG and 5′CAAG/3′GAAC have stabilities of 2.4 and 1.8 kcal/mol respectively, which cannot be explained by a model that considers loop formation to be entirely an entropic penalty. In the absence of energy parameters for terminal motifs, many theoretical models consider loop formation to be entirely entropic (*74*−*76*). The data presented here provide parameters to further improve theoretical models.

Surprisingly, two exceptional terminal motif sequences are less stable when compared to internal loops with similar sequences (Supporting Information Table 3). A terminal GG mismatch, 5′C**G**/3′G**G** (−1.6 kcal/mol) is less stable than an internal GG mismatch 5′C**GG**/3′G**GC** (−2.1 kcal/mol) (*77*). GG mismatches are idiosyncratically stable and dynamic even within an internal mismatch (*78*). A 2x2 terminal GA pair 5′G**GA**/3′C**AG** (−2.2 kcal/mol) is also less stable than an internal 2x2 loop 5′ G**GA**C/3′C**AG**G (−2.6 kcal/mol) (*39*). The NMR structure of this loop shows imino GA pairs and stacking that optimizes electrostatic overlap (*79*). In these two cases, perhaps the preferred conformation of the terminal motif nucleotides fits well with A-form helical structure and little entropic penalty is paid to close the loop with a base pair, resulting in only additional favorable stacking enthalpy. Alternatively, perhaps the stacking interactions with these sequences are so favorable that the entropic penalty of closing the loop with a Watson−Crick base pair is overcome by a large favorable enthalpy.

GU Pairs Are Idiosyncratically Stable. Thermodynamic stabilities of GU pairs are notoriously idiosyncratic and context dependent (*39, 40, 80*−*82*). Although imino proton NMR did not provide positive evidence for the formation of terminal GU pairs, the thermodynamic stabilities of terminal GU pairs with dangling ends and terminal mismatches have thermodynamic stabilities similar to the stabilities of AU terminal base pairs, which also often do not show imino proton resonances (*29*). This is consistent with rules for predicting the free energies of internal loops closed by GU pairs (*40*). The sequence 5′**UGC**GAUC**CUG** has an anomalously stable terminal motif free energy, −2.9 kcal/mol, and thus was not included in the linear regression analysis. Perhaps this sequence would be better described as a single mismatch and two terminal GU base pairs. The internal nearest neighbor parameter for 5′ UG/3′GU however is +0.3 kcal/mol (*83*), thus suggesting that once again the stability of GU pairs is highly idiosyncratic. Two small, very broad resonances between 10.5−11 ppm in the imino proton spectrum of 5′**UGC**GAUC**CUG** may be the results of hydrogen bonding in GU pairs. The 5′GU/3′UG motif at the end of the helix 5′**GU**UGCA**GU** has a stability of −1.6 kcal/mol in contrast to the internal nearest neighbor value of 1.3 kcal/mol and value of −1.1 kcal/mol in the context of 5′G**GU**C3′/3′C**UG**G5′. GU pairs were also noted as exceptional cases in the careful *in vitro* experiments designed to test the correlation of miRNA efficiency and specificity with miRNA:mRNA duplex thermodynamic stability (*18*).

Improved Thermodynamic Parameters for Terminal Motifs Can Provide Insight into RNAi Phenomena. Improved thermodynamic parameters will aid the biochemical and thermodynamic analysis of RNAi mechanisms (*17*−*19*). Adenines often occur adjacent to the seed regions for miRNA targets in human and vertebrate genomic analyses (*84*). The new prediction rules for terminal mismatches would predict more favorable thermodynamic stabilities than previous estimates for these adenines stacking on the core duplex in the miRNA:mRNA. The importance of the thermodynamic stability of the duplex formed by the 5′ end of the small RNA to mRNA target and the thermodynamic stability of the mRNA structure has recently been demonstrated in an *in vitro* system using human RISC enzyme (*17*). Table 5 provides a specific example of the importance of accurate thermodynamic parameters for understanding RNA interference specificity. The general correlation between 5′ siRNA seed region thermodynamic stability and translation repression in a carefully controlled *in vitro* assay is reconfirmed, but the exact value of the duplexes changes significantly with new thermodynamic parameters. Including the stability of the 3′ dangling nucleotides on the mRNA sequence makes more favorable all the predicted duplex free energies, although the changes in stability depend on the sequence of the mutation. The improved thermodynamic parameters for internal loops, bulge loops, and terminal motifs make the predicted free energies more favorable by 0.6 to 2.8 kcal/mol, depending on the sequence. The average improvement in stability for these sixteen duplexes is 1.5 kcal/mol, which is 1 order of magnitude in a binding constant at 37 °C. This particular example does not take into account the cost of interrupting any secondary structure in the mRNA target, which may vary with the different mutations. Protein−RNA interactions and the shape of the RNA helix will also

contribute to the stability and specificity of the siRNA:mRNA:RISC complex, but thermodynamic stability of the siRNA:mRNA helix plays an important role. While a complete explanation of the specificity and efficiency of RNAi is still being explored, accurate thermodynamic parameters are a necessary tool for unraveling the mysteries of off-target effects, discovering the fundamental basis for RNAi phenomena, and exploiting the power of RNA to regulate gene expression for the benefit of human health.

ACKNOWLEDGMENT

The research results discussed in this publication were made possible by the OPSR award for project number 080223 from the Oklahoma Center for the Advancement of Science and Technology and project number 080130 from the Pharmaceutical and Research Manufacturers of America Foundation.

SUPPORTING INFORMATION AVAILABLE

A table of terminal motif enthalpies and entropies, a linear regression analysis spreadsheet, and a comparison of external and internal loops. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature 391*, 806–811.
2. Zamore, P., D., Tuschl, T., Sharp, P. A., and Bartel, D. A. (2000) RNAi:Double-Stranded RNA Directs the ATP-Dependent Cleavage of mRNA at 21 to 23 Nucleotide Intervals. *Cell 101*, 25–33.
3. Napoli, C., Lemieux, C., and Jorgenson, R. (1990) Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in Trans. *Plant Cell 2*, 279–289.
4. van der Krol, A. R., Marcel, L. A., Beld, M., Mol, J. N. M., and Stuitje, A. R. (1990) Flavenoid Genes in Petunia: Addition of a Limited Number of Gene Copies May Lead to a Suppression of Gene Expression. *Plant Cell 2*, 291–299.
5. Kim, D. H., and Rossi, J. J. (2007) Strategies for silencing human disease using RNA interference. *Nat. Rev. Genet. 8*, 173–184.
6. Jackson, A. L., Burchard, J., Schelter, J., Chau, B. N., Cleary, M., Lim, L., and Linsley, P. S. (2006) Widespread siRNA "off-target" transcript silencing mediated by seed region complementarity. *RNA 12*, 1179–1187.
7. Federov, Y., Anderson, E. M., Birmingham, A., Reynolds, A., Karpilow, J., Robinson, K., Leaske, D., Marshall, W. S., and Khvorova, A. (2006) Off-Target Effects by siRNA Can Induce Toxic Phenotype. *RNA 12*, 1188–1196.
8. Mathews, D. H., Burkard, M. E., Freier, S. M., Wyatt, J. R., and Turner, D. H. 1999. Predicting Oligonucleotide Affinity to Nucleic Acid Targets. *RNA 5*, 1458–1469.
9. Long, D., Lee, R., Williams, P., Chan, C. Y., Ambros, V., and Ding, Y. (2007) Potent Effect of Target Structure on microRNA Function. *Nat. Struct. Mol. Biol. 14*, 287–294.
10. Westerhout, E. M., Ooms, M., Vink, M., Das, A. T., and Berkhout, B. (2005) HIV-1 can Escape from RNA Interference by Evolving an Alternative Structure in its RNA Genome. *Nucleic Acids Res. 33*, 796–804.
11. Tafer, H., Ameres, S. L., Obernosterer, G., Gebeshuber, C. A., Schroeder, R., Martinez, J., and Hofacker, I. L. (2008) The Impact of Target Accessibility on the Design of Effective siRNAs. *Nat. Biotechnol. 26*, 578–583.
12. Ding, Y., Chan, C. Y., and Lawrence, C. E. (2004) Sfold Web Server for Statistical Folding and Rational Design of Nucleic Acids. *Nucleic Acids Res. 32*, 135–141.
13. Zuker, M. (2003) Mfold Server for Nucleic Acid Folding and Hybridization Prediction. *Nucleic Acids Res. 31*, 3406–3415.
14. Gultyaev, A. P., van Batenburg, F. H., and Pleij, C. W. (1995) The Computer Simulation of RNA Folding Pathways Using A Genetic Algorithm. *J. Mol. Biol. 250*, 37–51.
15. Mathews, D. H., Disney, M. D., Childs, J. L., 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.
16. Gruber, A., Lorenz, R., Bernhart, S., Neubock, R., and Hofacker, I. (2008) The Vienna RNA Websuite. *Nucleic Acids Res. 36*, W70–74.
17. Ameres, S. L., Martinez, J., and Schroeder, R. (2007) Molecular Basis for Target RNA Recognition and Cleavage by Human RISC. *Cell 130*, 101–112.
18. Doench, J. G., and Sharp, P. A. (2004) Specificity of microRNA Target Selection in Translational Repression. *Genes Dev. 18*, 504–511.
19. Haley, B., and Zamore, P. D. (2004) Kinetic Analysis of the RNAi Enzyme Complex. *Nat. Struct. Biol. 11*, 599–606.
20. Khvorova, A., Reynolds, A., and Jayasena, S. D. (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell 115*, 209–216.
21. Schrwarz, D. S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P. D. (2003) Asymmetry in the assemby of the RNAi enzyme complex. *Cell 115*, 199–208.
22. Palatnik, J. F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J. C., and Weigel, D. (2003) Control of Leaf Morphogenesis by microRNAs. *Nature 425*, 257–263.
23. Kerstetz, M., Iovino, N., Unnerstall, U., Gaul, U., and Segal, E. (2007) The Role of Site Accessibility in microRNA Target Recognition. *Nat. Genet. 39*, 1278–1284.
24. Hofacker, I. (2007) How microRNAs Choose their Targets. *Nat. Genet. 39*, 1191–1192.
25. Shao, Y., Chan, C. Y., Maliyekkel, A., Lawrence, C. E., Roninson, I. B., and Ding, Y. (2007) Effect of Target Secondary Structure on RNAi Efficiency. *RNA 13*, 1631–1640.
26. Lu, Z., and Mathews, D. (2008) Fundamental differences in the equilibrium considerations for siRNA and antisense oligodeoxynucleotide design. *Nucleic Acids Res. 36*, 3738–3745.
27. 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.
28. Schroeder, S. J., Kim, J., and Turner, D. H. (1996) GA and UU mismatches can stabilize 1x2 internal loops of three nucleotides. *Biochemistry 35*, 16105–16109.
29. Schroeder, S. J., and Turner, D. H. (2000) Factors Affecting the Thermodynamics Stability of Small Asymmetric Internal Loops in RNA. *Biochemistry 39*, 9257–9274.
30. Chen, G., and Turner, D. H. (2006) Consecutive GA pairs stabilize medium size internal loops. *Biochemistry 45*, 4025–4043.
31. Chen, G., Znosko, B., Jiao, X., and Turner, D. H. (2004) Factors Affecting the Thermodynamic Stabilities of RNA 3x3 Internal Loops. *Biochemistry 43*, 12865–12876.
32. Chen, G., and Turner, D. H. (2006) Consecutive GA pairs stabilize medium-size internal loops. *Biochemistry 45*, 4025–4043.
33. Davis, A. R., and Znosko, B. (2007) Thermodynamic characterization of single mismatches found in naturally occurring RNA. *Biochemistry 46*, 13425–13436.
34. Badhwar, J., Karri, S., Cass, C. K., Wunderlich, E., and Znosko, B. (2007) Thermodynamic characterization of RNA duplexes containing naturally occurring 1x2 nucleotide internal loops. *Biochemistry 46*, 14715–14724.
35. Santa Lucia, J. J., 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.
36. SantaLucia, J, Jr., 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*, 4313–4321.
37. Serra, M., Barnes, T., Betschart, K., Gutierrez, M., Sprouse, K., Riley, C., Stewart, L., and Temel, R. (1997) Improved parameters for the prediction of RNA hairpin stability. *Biochemistry 36*, 4844–4851.
38. Groebe, D., and Uhlenbeck, O. (1988) Characterization of RNA hairpin loop stability. *Nucleic Acids Res. 16*, 11725–11735.
39. 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.
40. Schroeder, S. J., and Turner, D. H. (2001) Thermodynamic stabilities of internal loops with GU closing pairs in RNA. *Biochemistry 40*, 11509–11517.

41. Schroeder, S. J., Burkhard, M. E., and Turner, D. H. (2000) Energetics of Small Internal Loops in RNA. *Biopolymers 52*, 157–167.
42. Kierzek, R., Burkard, M. E., and Turner, D. H. (1999) Thermodynamics of Single Mismatches in RNA Duplexes. *Biochemistry 38*, 14214–14223.
43. 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.
44. Liu, J. D., Zhao, L., and Xia, T. (2008) The dynamic structural basis of differential enhancement of conformational stability by 5′ and 3′ dangling ends in RNA. *Biochemistry 47*, 5962–5975.
45. Ohmichi, T., Nakano, S., Miyoshi, D., and Sugimoto, N. (2002) Long Dangling End Has Large Energetic Contribution to Duplex Stability. *J. Am. Chem. Soc. 124*, 10367–10372.
46. Pasternak, A., Kierzek, E., Pasternak, K., Fratczak, A., Turner, D. H., and Kierzek, R. (2008) The thermodynamics of 3′terminal pyrene and guanosine for the design of isoenergetic 2′-O-methyl-RNA-LNA chimeric oligonucleotide probes of RNA structure. *Biochemistry 47*, 1249–1258.
47. Burkard, M. E., Kierzek, R., and Turner, D. H. (1999) Thermdodynamics of unpaired terminal nucleotides on short RNA helixes correlated with stacking at the helix termini in larger RNAs. *J. Mol. Biol. 290*, 967–982.
48. O'Toole, A. S., Miller, S., Haines, N., Zink, M. C., and Serra, M. J. (2006) Comprehensive Thermodynamic Analysis of 3′ Double Nucleotide Overhangs Neighboring Watson−Crick Terminal Base Pairs. *Nucleic Acids Res. 34*, 3338–3344.
49. O'Toole, A. S., Miller, S., and Serra, M. J. (2005) Stability of 3′ Double Nucleotide Overhangs that Model the 3′ Ends of si RNA. *RNA 11*, 512–516.
50. Borer, P. N. (1975) *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*, 3rd ed., Vol. 1, p 589, CRC Press, Cleveland, OH.
51. 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.
52. 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.
53. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) Stability of ribonucleic double-stranded helices. *J. Mol. Biol. 86*, 843–853.
54. Freier, S. M., Kierzek, R., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) Free energy of GU contributions and other terminal mismatches to helix stability. *Biochemistry 25*, 3209–3213.
55. Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) RNA Structure Prediction. *Annu. Rev. Biophys. Biophys. Biochem. 17*, 167–192.
56. Hickey, Turner, D. H. (1985) Effects of terminal mismatches on RNA Stability:Thermodynamics of Duplex Formation for ACCGGGp, ACCGGAp, and ACCGGCp. *Biochemistry 24*, 3987–3991.
57. Sugimoto, N., Kierkek, R., and Turner, D. H. (1987) Sequence Dependence for the Energetics of Terminal Mismatches in Riboligonucleotides. *Biochemistry 26*, 4559–4562.
58. Xia, T., Santa Lucia, J., Jr., Burkhard, 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*, 14719–35.
59. 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.
60. Kapoor, Sunkar, R., and Zhu, J.-K. (2006) Postranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell 18*, 2051–2065.
61. Vieregg, J., Cheng, W., Bustamante, C., and Tinoco, I, Jr. (2007) Measurement of the effect of monovalent cations on RNA hairpin stability. *J. Am. Chem. Soc. 129*, 14966–14973.
62. Turner, D. H. (2000) Conformational Changes, in *Nucleic Acids: Structures, Properties, and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr., Eds.) pp 281−282, University Science Books, Sausalito, CA.
63. Freier, S. M., Alkema, D., Sinclair, A., Neilson, T., and Turner, D. H. (1985) Contributions of Dangling End Stacking and Terminal Base Pair Formation to the Stabilities of XGGCp, XCCGGp, XGGCCYp, and XCCGGYp Helices. *Biochemistry 24*, 4533–4539.
64. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry*, W.H. Freeman and Company, New York.
65. Kollman, P. (2000) in *Nucleic Acids: Structures, Properties, and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr., Eds.) p 282, University Science Books, Sausalito, CA.
66. Turner, D. H. (2000) Conformational Changes, in *Nucleic Acids: Structures, Properties, and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr., Eds.) p 282, University Science Books, Sausalito, CA.
67. Serra, M., Axenson, T., and Turner, D. H. (1994) A model for the stability of RNA hairpins based on a study of sequence dependence of stability for hairpins of six nucleotides. *Biochemistry 33*, 14289–14296.
68. Seol, Y., Skinner, G. M., and Visscher, K. (2004) Elastic properties of single-stranded charged homopolymeric ribonucleotide. *Phy. Rev. Lett. 10*, 118102.
69. Seol, Y., Skinner, G. M., and Visscher, K. (2007) Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking. *Phys. Rev. Lett. 98*, 158103–158107.
70. Hyeon, C., Dima, R. I., and Thirumalai, D. (2006) Size, shape, and flexibility of RNA Structures. *J. Chem. Phys. 125*, 194905–194915.
71. Stannard, B. S., and Felsenfeld, G. (1975) The conformation of polyriboadenylic acid at low temeprature and neutral pH. A single-stranded rodlike structure. *Biopolymers 14*, 299–307.
72. Freier, S. M., Hill, K., Dewey, T., Marky, L., Breslauer, K., and Turner, D. H. (1981) Solvent Effects on the Kinetics and Thermodynamics of Stacking in Poly(cytidylic acid). *Biochemistry 20*, 1419–1426.
73. Peritz, A., Kierzek, 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.
74. Chen, S.-J., and Dill, K. A. (2000) RNA folding energy landscapes. *Proc. Natl. Acad. Sci. U.S.A. 97*, 646–652.
75. Cao, S., and Chen, S.-J. (2005) Predicting RNA folding thermodynamics with a reduced representation model. *RNA 11*, 1884–1897.
76. Zhang, J., Lin, M., Chen, R., Wang, W., and Liang, J. (2008) Discrete state model and accurate estimation of loop entropy of RNA secondary structures. *J. Chem. Phys. 128*, 125107–125117.
77. Kierzek, R., Burkard, M. E., and Turner, D. H. (1999) Thermodynamics of Single Mismatches in RNA Duplexes. *Biochemistry 38*, 14214–14223.
78. Burkard, M. E., and Turner, D. H. (2000) NMR structures of r(GCAGGCGUGC)2 and determinants of stability for single guanosine-guanosine base pairs. *Biochemistry 39*, 11748–11762.
79. Wu, M., and Turner, D. H. (1996) Solution structure of r(GCG-GACGC)2 by two-dimensional NMR and the iterative relaxation matrix approach. *Biochemistry 35*, 9677–9689.
80. McDowell, J., 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.
81. McDowell, J., 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.
82. Tolbert, B. S., Kennedy, S. D., Schroeder, S. J., Krugh, T. R., and Turner, D. H. (2007) NMR structures of (rGCUGAGGCU)2 and (rGCGGAUGCU)2: probing the structural features that shape the thermodynamic stability of GA pairs. *Biochemistry 46*, 1511–1522.
83. He, L., Kierzek, R., SantaLucia, J., Jr., Walter, A., and Turner, D. H. (1991) Nearest-neighbor parameters for GU mismatches. *Biochemistry 30*, 11124–11132.
84. Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell 120*, 15–20.
85. Zuker, M. (2003) Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction. *Nucleic Acids Res. 31*, 3406–3415.

BI801594K