1894–1902 *Nucleic Acids Research, 2011, Vol. 39, No. 5* *Published online 10 November 2010*
doi:10.1093/nar/gkq905

Thermodynamic contributions of single internal rA·dA, rC·dC, rG·dG and rU·dT mismatches in RNA/DNA duplexes

Norman E. Watkins Jr*, William J. Kennelly, Mike J. Tsay, Astrid Tuin, Lara Swenson, Hyung-Ran Lee, Svetlana Morosyuk, Donald A. Hicks and John SantaLucia Jr

DNA Software Inc., Ann Arbor, MI 48104, USA

Received May 27, 2010; Revised September 21, 2010; Accepted September 22, 2010

ABSTRACT

The thermodynamic contributions of rA·dA, rC·dC, rG·dG and rU·dT single internal mismatches were measured for 54 RNA/DNA duplexes in a 1 M NaCl buffer using UV absorbance thermal denaturation. Thermodynamic parameters were obtained by fitting absorbance versus temperature profiles using the curve-fitting program *Meltwin*. The weighted average thermodynamic data were fit using singular value decomposition to determine the eight non-unique nearest-neighbor parameters for each internal mismatch. The new parameters predict the ΔG°_{37}, ΔH° and melting temperature (T_m) of duplexes containing these single mismatches within an average of 0.33 kcal/mol, 4.5 kcal/mol and 1.4°C, respectively. The general trend in decreasing stability for the single internal mismatches is rG·dG > rU·dT > rA·dA > rC·dC. The stability trend for the base pairs 5′ of the single internal mismatch is rG·dC > rC·dG > rA·dT > rU·dA. The stability trend for the base pairs 3′ of the single internal mismatch is rC·dG > rG·dC >> rA·dT > rU·dA. These nearest-neighbor values are now a part of a complete set of single internal mismatch thermodynamic parameters for RNA/DNA duplexes that are incorporated into the nucleic acid assay development software programs *Visual oligonucleotide modeling platform* (OMP) and *ThermoBLAST.*

INTRODUCTION

Recently, much attention has been paid to the detection and analysis of messenger RNAs (mRNAs) that contribute to human diseases such as cardiovascular diseases, cancer, immune-deficiency disorders and diabetes (1–7). The investigation of these genes is enabled by such technologies as northern-blot analysis (8), *in situ* hybridization (9), the RNA 'invader' assay (2), reverse transcription (RT)–PCR (10) and expression microarrays (11), all of which depend upon the hybridization of a DNA primer or probe to a target RNA. DNA/RNA hybridization is also important for antisense therapeutic applications (12–14). In the design of primers and probes for molecular diagnostics applications, oligonucleotide sensitivity is an important factor for experimental efficiency (15). The inability to properly account for the effects of nucleotide mismatches often times results in a reduction of assay specificity due to the effect that a mismatch has upon duplex hybridization (16–18). For comparison, consider the relative free energies of mismatches in DNA duplexes: a dG·dT mismatch is stabilizing in all nearest-neighbor contexts, whereas a dC·dT mismatch is *destabilizing* in all nearest-neighbor contexts (17,19). While the effects of internal mismatches in DNA duplexes have been well documented in the literature (16–19), there are very few primer and probe design software programs that account for nucleotide mismatches in DNA/RNA hybrid duplexes. This results in the unwanted amplification of background sequences and false positive or off-target effects. Thus, to design and simulate the most selective DNA primers and probes for RNA/DNA based assays and therapeutics, thermodynamic parameters for internal mismatches are needed.

Thermodynamic parameters for single internal mismatches in RNA/DNA duplexes can also be used to detect and characterize important RNA structural and functional roles (20). Thermodynamic parameters for single internal mismatches in RNA/DNA duplexes can also be used to detect and characterize important RNA structural and functional roles (20–22), and in the interaction of RNA polymerase proteins with the influenza A viral promoter sequence (23).

While a full set of nearest-neighbor thermodynamic parameters have been reported for RNA/DNA

*To whom correspondence should be addressed. Tel: +1 734 222 9080; Fax: +1 734 222 9087; Email: norm@dnasoftware.com

Watson–Crick base pairs (24), there are no such reports for a complete nearest-neighbor thermodynamic parameters for single internal mismatches in RNA/DNA duplexes. This study reports a total of 43 thermodynamic measurements that were combined with 11 measurements from the literature for the single internal mismatches rA•dA, rG•dG, rU•dT and rC•dC in RNA/DNA duplexes. These data were used to solve for 32 nearest-neighbor unknowns (28 degrees of freedom, see 'Materials and Methods' section) using singular value decomposition (SVD) (*25*). These parameters can be readily incorporated into *in silico* design and simulation tools for primers and probes, so that accurate molecular diagnostic assays used for the detection and characterization of these biologically important mismatches in RNA/DNA duplexes can be developed. A full determination is reported of four single internal mismatch base pair contributions to RNA/DNA duplexes, each of which are 2-fold over-determined, except for rC•dC internal mismatches which are 1.6-fold over-determined. These nearest-neighbor values are a sample of the complete RNA/DNA thermodynamic parameter database that contains a comprehensive set of match, internal and terminal mismatches, dangling end nearest-neighbor parameters and salt dependence, all of which are available in the software programs *Visual OMP* (15,26) and *ThermoBLAST* (DNA Software Inc.).

MATERIALS AND METHODS

Oligo synthesis and purification

Sigma-Genosys and Dharmacon Inc. synthesized the DNA and RNA oligonucleotides used in this study using standard phosphoramidite chemistries (27,28). All oligomers were deblocked, HPLC purified, and certified by mass-spectrometry by the suppliers.

Design of sequences

Each duplex in this study (Supplementary Table S1) was designed to have a length between 8 and 11 nts with the mismatch in the central position to yield a melting temperature (T_m) for 10 μM strands to be between 31 and 70°C. Such a design allows for appropriate curve-fitting by ensuring that there are sufficient upper and lower baselines to deduce accurate thermodynamic parameters using the software program *Meltwin* (29,30). Sequences were also designed to minimize the formation of undesired hairpins, slipped duplexes and homodimers (31). For rA•dA internal mismatches, the thermodynamic data sets for 10 new RNA/DNA duplexes determined at DNA Software were combined with four duplexes from the literature (32) to define 14 equations that were used to solve for eight nearest-neighbor unknowns to obtain a non-unique solution with seven free parameters (16) using SVD. Likewise, for rG•dG mismatches 10 new duplexes were combined with four RNA/DNA duplexes from the literature (32). For rU•dT mismatches, 12 new RNA/DNA duplexes were combined with three RNA/DNA duplexes from the literature (32). As there were no data available for RNA/DNA rC•dC internal mismatches in the literature, the SVD equation set was populated exclusively by 11 new duplexes that were determined at DNA Software. Sequences were designed to contain all of the possible nearest-neighbor contexts ensuring uniform representation in deriving thermodynamic parameters that exhibit small standard errors and excellent predictive capabilities.

Measurement of melting curves

Data were measured using a Beckman DU 650 spectrophotometer with a six-cuvette thermoelectric controller. The melting buffer used in this study was 1.0 M NaCl, 10 mM Na_2HPO_4 and 0.5 mM Na_2EDTA, which was titrated with a 0.1 M HCl solution to obtain a pH of 7.0. The use of high salt concentrations removes length-dependent counterion-condensation effects that can influence the proper determination of nearest-neighbor parameters (33,34). The total concentration of each single-stranded oligonucleotide was calculated from the absorbance reading taken at a wavelength of 260 nM using a NanoDrop 1000 spectrophotometer (Thermo Scientific). Extinction coefficients for RNA and DNA oligonucleotides were predicted using the nearest-neighbor model (16,35,36) using the extinction coefficient calculator provided in *Visual OMP*. The total concentration of each single-stranded oligonucleotide was used to mix equimolar concentrations of the non-self-complementary strands to an initial volume of 180 μl as previously described (36). Serial dilutions of this 180 μl sample were made to obtain six different oligonucleotide concentrations over a 20-fold dilution range so that the observed ultraviolet (UV) absorbance reading was below 2.00 in a 1.0 cm pathlength to stay within the linear region of Beer's Law. The total strand concentration, C_T, was calculated from the average of the two individual strand extinction coefficients and the experimental UV absorbance at 85°C (37). The samples were then allowed to cool to 10°C and the data were collected at a wavelength of 260 nm while heating the sample at a rate of 1.0°C/min, at a read interval of 0.5°C up to a maximum of 95°C (34,37).

Determination of thermodynamic parameters

Absorbance versus temperature profiles were used to obtain a full thermodynamic parameter set for each duplex using *Meltwin* version 3.5 through two different methods: (i) the T_m^{-1} versus $\ln C_T/4$ method and (ii) the Marquardt non-linear least squares curve fit method (38). The Marquardt non-linear least squares method was used to fit absorbance versus temperature curves to calculate the enthalpy change, ΔH°, and the entropy change, ΔS°, by assuming linear sloped upper and lower baselines and by the assumption that the change in heat capacity (ΔC_p°) is zero for the transition equilibrium (37). Recent studies have indicated that the average ΔC_p° for DNA duplexes is ~100 cal/K•mol per base pair (39–41). In this study, however, we have retained the $\Delta C_p^\circ = 0$ assumption because generally UV melting curve data are not accurate enough to justify the use of an additional fitting parameter. In addition, this $\Delta C_p^\circ = 0$ assumption is consistent with previous studies that have used relatively short oligonucleotide duplexes that exhibit little potential for

competing hairpins or single-strand stacking (31,32,37). Therefore the assumption that $\Delta C_p = 0$ appears to be valid, as demonstrated by the two-state behavior (see below) and highly predictive nearest-neighbor parameters reported here. The ΔH° and ΔS° parameters were also calculated from a linear plot of T_m^{-1} versus $\ln C_T/4$ where the slope equals $R/\Delta H^\circ$, where R is the gas constant, and the intercept equals $\Delta S^\circ/\Delta H^\circ$ (36). When enthalpy values from these two methods differed by <15%, the two-state approximation was considered to be valid (16,42). Raw data for the determination of the nearest-neighbor contributions of single internal mismatches in RNA/DNA duplexes are shown in Supplementary Table S1.

Error analysis

Sampling errors were determined using *Meltwin* v. 3.5 and combined into a weighted average as previously described (36). For literature duplexes sampling errors were also combined into a weighted average as a full set of thermodynamic data were reported (32). These results were then combined with the internal mismatch and Watson–Crick dimer errors to compute the square-root of the sum of the squares for duplex ΔG° and ΔH°, which were the error values used in *Mathematica* SVD calculations.

Determination of mismatch contributions

The contribution of an internal mismatch to duplex stability is calculated from the observed error weighted average ΔG°_{37} and ΔH° values of the duplex containing the internal mismatch by subtracting the values for the core duplex (i.e. without a mismatch), and then adding in the Watson–Crick dimer nearest-neighbor free energy and enthalpy values where the mismatch was inserted. For example:

$$\Delta G\begin{pmatrix}\text{AUU}\\ \text{TTA}\end{pmatrix} = \Delta G\begin{pmatrix}\text{CAUCAUUCGCC}\\ \text{GTAGTTAGCGG}\end{pmatrix} - \Delta G\begin{pmatrix}\text{CAUCAUCGCC}\\ \text{GUAGTAGCGG}\end{pmatrix} + \Delta G\begin{pmatrix}\text{AU}\\ \text{TA}\end{pmatrix}. \quad (1)$$

The free energy of the trimer ($\Delta G^\circ_{37\ \text{trimer}}$) used in Equation (1) was used to construct Equation (2).

$$\Delta G^\circ_{37\ \text{trimer}} = \Delta G^\circ_{37\ \text{exp}} - \Delta G^\circ_{37\ \text{core}} + \Delta G^\circ_{37\ \text{NN}} \quad (2)$$

From Equation (2), the free energy of the trimer is decomposed into its dimer nearest-neighbor constituents as shown in Equation (3). For example:

$$\Delta G\begin{pmatrix}\text{AUU}\\ \text{TTA}\end{pmatrix} = \Delta G\begin{pmatrix}\text{AU}\\ \text{TT}\end{pmatrix} + \Delta G\begin{pmatrix}\text{UU}\\ \text{TA}\end{pmatrix} = \Delta G^\circ_{37\ \text{trimer}} \quad (3)$$

A system of equations similar to that shown in Equation (3) were used as input to *Mathematica* version 4.1 and the set of eight nearest-neighbor parameters was solved using SVD (25). A similar method was used to calculate the enthalpic contribution (i.e. ΔH°) for these systems.

Table 1. Non-unique nearest-neighbor thermodynamics of rA•dA, rC•dC, rG•dG and rU•dT single internal mismatches in RNA/DNA duplexes in 1 M NaCl, 10 mM Na_2HPO_4, 0.5 mM Na_2EDTA, pH 7.0

Mismatch dimer	ΔG°_{37} (kcal/mol)	ΔH° (kcal/mol)
rAA/dTA	+1.07 ± 0.08	−4.3 ± 2.0
rUA/dAA	+1.13 ± 0.06	−1.7 ± 1.4
rGA/dCA	+0.51 ± 0.11	−1.9 ± 1.4
rCA/dGA	+0.90 ± 0.12	+5.5 ± 1.6
rAA/dAT	+1.36 ± 0.06	+3.0 ± 1.7
rAU/dAA	+1.85 ± 0.12	+5.6 ± 2.6
rAG/dAC	+0.21 ± 0.09	−10.5 ± 1.4
rAC/dAG	+0.19 ± 0.07	−0.60 ± 1.3
Average value	+0.90	−0.61
rAC/dTC	+1.64 ± 0.08	−8.8 ± 1.8
rUC/dAC	+1.15 ± 0.09	−3.3 ± 3.0
rGC/dCC	+0.96 ± 0.11	−0.1 ± 2.6
rCC/dGC	+1.04 ± 0.10	+10.5 ± 2.1
rCA/dCT	+1.70 ± 0.09	−0.3 ± 2.2
rCU/dCA	+1.88 ± 0.11	+0.8 ± 2.8
rCG/dCC	+0.46 ± 0.08	−11.5 ± 1.7
rCC/dCG	+0.73 ± 0.10	+9.3 ± 2.9
Average value	+1.20	−0.43
rAG/dTG	+0.31 ± 0.08	−3.3 ± 1.9
rUG/dAG	+0.44 ± 0.08	−5.8 ± 2.1
rGG/dCG	−0.58 ± 0.12	−8.0 ± 1.8
rCG/dGG	+0.14 ± 0.08	−8.9 ± 1.4
rGA/dGT	+0.50 ± 0.08	+1.1 ± 2.1
rGU/dGA	+0.97 ± 0.09	−3.7 ± 2.2
rGG/dGC	−0.33 ± 0.09	−16.5 ± 1.7
rGC/dGG	−0.83 ± 0.08	−7.0 ± 1.4
Average value	+0.08	−6.5
rAU/dTT	+0.63 ± 0.07	+0.6 ± 0.2
rUU/dAT	+1.07 ± 0.07	−2.2 ± 0.3
rGU/dCT	+0.18 ± 0.12	−11.6 ± 0.4
rCU/dGT	+0.49 ± 0.07	−0.4 ± 0.2
rUA/dTT	+1.21 ± 0.07	−3.3 ± 0.2
rUU/dTA	+1.03 ± 0.10	+3.0 ± 0.3
rUG/dTC	+0.14 ± 0.07	−13.4 ± 0.5
rUC/dTG	−0.02 ± 0.07	+0.1 ± 0.2
Average value	+0.59	−3.4

RNA/DNA duplexes are chimeric as there are eight dimer nearest-neighbor parameters which are non-unique where the stacking matrix (i.e. the number of unknowns) is eight, but only seven linearly independent trimers are required for a full thermodynamic determination (16). These parameters are for single internal mismatches in RNA/DNA duplexes only and do not apply to terminal mismatches. Errors shown are standard deviations computed by error propagation.

Nearest-neighbor ΔS° values were calculated from the determined values of ΔG°_{37} and ΔH°. Since none of the sequences have terminal mismatches, the stacking matrix in SVD is rank deficient, and thus seven unknowns for each of the single internal mismatches are uniquely determined in this study (16). The equivalence of the dimer and trimer formulations of nearest-neighbor parameters are described in detail in previous work (16,43). The full sets of experimentally determined thermodynamic parameters are listed in Table 1.

RESULTS AND DISCUSSION

Thermodynamic data

The data used in the determination of the thermodynamic contributions presented here are provided in the Supplementary Table S1. This table contains the thermodynamic parameters derived from $1/T_m$ versus $\ln C_T/4$

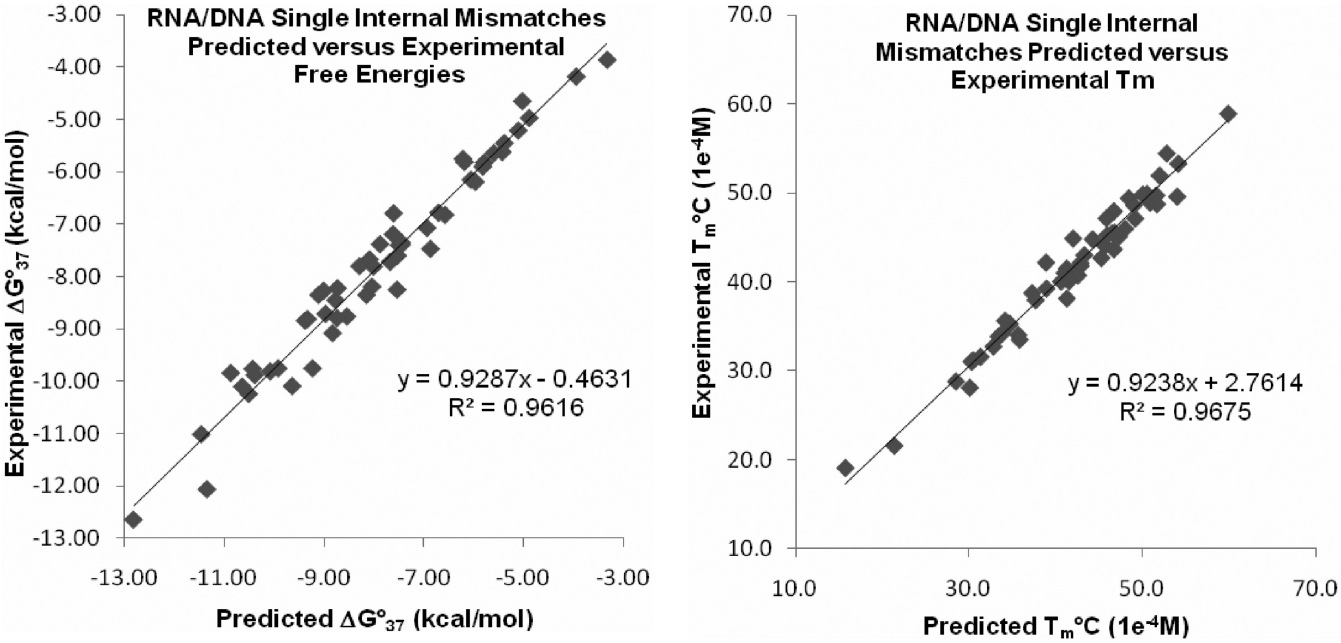

Figure 1. Comparison of experimental versus predicted free energies and T_m's for RNA/DNA duplexes with single internal rA•dA, rC•dC, rG•dG and rU•dT mismatches. Free energies are predicted within 0.34 kcal/mol, on average. Enthalpies and T_m's are predicted to within 3.8 kcal/mol and 1.3°C, on average, respectively.

plots and curve fit data sets for each duplex. The nearest-neighbor thermodynamic parameters for the single internal mismatches rA•dA, rC•dC, rG•dG and rU•dT in RNA/DNA duplexes are shown in Table 1. These parameters were used to calculate the predicted thermodynamics of each of the RNA/DNA duplexes used in this study so that a comparison could be made with the observed experimental thermodynamic values for each of the duplexes listed in Supplementary Table S2. Supplementary Table S2 also lists the standard deviations for each of the individual internal mismatches reported here. The combined average deviation for RNA/DNA duplexes containing the single internal mismatches rA•dA, rC•dC, rG•dG and rU•dT can now be predicted, on average, within 0.33 kcal/mol in free energy change, 4.5 kcal/mol in enthalpy change and 1.4°C in melting temperature when compared to the observed experimental thermodynamics. This good agreement is shown in Figure 1 where the R^2 of the trendline between experimental versus predicted free energy and T_m values are both observed to be 0.96 and 0.97, respectively, which is close to the ideal value of 1.0. The observed agreement between experimental and predicted thermodynamics for a 2-fold over-determination of parameters indicates that the nearest-neighbor model is a valid approximation for single internal mismatches in RNA/DNA hybrid duplexes. This agrees with results obtained previously for mismatches in DNA/DNA and RNA/RNA duplexes (18,42).

Context dependence of single internal mismatches in RNA/DNA duplexes

Trimer stabilities. According to the nearest-neighbor model, the total contribution of a mismatch trimer to an oligonucleotide duplex is equal to the sum of the energy of its nearest-neighbor dimers. For instance, the total energy for the trimer rCGA/dGGU is equal to +0.64 kcal/mol which is equal to the sum of the energies of the nearest-neighbor dimers rCG/dGG and rGA/dGT which equal +0.14 and +0.50 kcal/mol, respectively. As only seven of the eight rG•dG nearest-neighbor parameters are uniquely determined, there are only seven linearly independent trimers that contain a central rG•dG mismatch. The other nine possible rG•dG mismatch trimers can be derived exactly from the seven unique trimers, as previously reported (16). Therefore, the data in Table 1 can be used to predict the thermodynamics of each of the rG•dG, rA•dA, rU•dT and rC•dC single internal mismatches in RNA/DNA duplexes in all 16 different trimer contexts, all of which are read 5′–3′, with Watson–Crick closing pairs as shown in Table 2.

The most stable trimer context is rGGC/dGGC, which contributes −1.40 kcal/mol to duplex free energy at 37°C in comparison to the least stable trimer context rACU/dACT, which destabilizes an RNA/DNA duplex by +3.51 kcal/mol. This 4.9 kcal/mol range of trimer stabilities is the same as that observed for internal mismatches in DNA/DNA duplexes (18) though the RNA/DNA internal mismatch energies are ∼1 kcal/mol less stable than those observed in DNA/DNA duplexes. This result is in contrast to Watson–Crick base pair stabilities in which RNA/DNA nearest-neighbors are more stable than DNA/DNA nearest-neighbors (24,25). This suggests that RNA/DNA hybridization is more specific than DNA/DNA hybridization. Average trimer stabilities in each Watson–Crick closing base pair context were calculated for the sake of comparison. The overall average shown in Table 2 for a particular internal

Table 2. The 16 trimer combinations for each of the reported single internal mismatches in RNA/DNA duplexes. These values were calculated using the parameters in Table 1

	Trimer ΔG°_{37}		Trimer ΔG°_{37}		Trimer ΔG°_{37}		Trimer ΔG°_{37}	Overall average
rG•dG Trimers								
rAGA/dTGT	+0.81	rGGG/dCGC	−0.91	rAGG/dCGT	−0.02	rGGA/dTGC	−0.08	
rUGA/dTGA	+0.94	rCGG/dCGG	−0.19	rUGG/dCGA	+0.11	rCGA/dTGG	+0.64	
rAGU/dAGT	+1.28	rCGC/dGGG	−0.68	rUGC/dGGA	−0.38	rGGU/dAGC	+0.39	
rUGU/dAGA	+1.41	rGGC/dGGC	−1.40	rAGG/dCGT	−0.02	rCGU/dAGG	+1.11	
Average	+1.11		−0.80		−0.08		+0.52	+0.19
rA•dA Trimers								
rAAA/dTAT	+2.43	rGAG/dCAC	+0.72	rAAG/dCAT	+1.28	rGAA/dTAC	+1.87	
rUAA/dTAA	+2.03	rCAG/dCAG	+1.11	rUAG/dCAA	+1.34	rCAA/dTAG	+2.26	
rAAU/dAAT	+2.92	rCAC/dGAG	+1.09	rUAC/dGAA	+1.32	rGAU/dAAC	+2.36	
rUAU/dTAT	+2.98	rGAC/dGAC	+0.70	rAAG/dCAT	+1.28	rCAU/dAAG	+2.75	
Average	+2.59		+0.91		+1.31		+2.31	+1.78
rU•dT Trimers								
rAUA/dTTT	+1.84	rGUG/dCTC	+0.32	rAUG/dCTT	+0.77	rGUA/dTTG	+1.39	
rUUAd/TTA	+2.28	rCUG/dCTG	+0.63	rUUG/dCTA	+1.21	rCUA/dTTG	+1.70	
rAUU/dATT	+1.66	rCUC/dGTG	+0.47	rUUC/dGTA	+1.05	rGUU/dATC	+1.21	
rUUU/dTTT	+2.10	rGUC/dGTC	+0.16	rAUG/dCTT	+0.77	rCUU/dATG	+1.52	
Average	+1.97		+0.40		+0.95		+1.46	+1.20
rC•dC Trimers								
rACA/dTCT	+3.33	rGCG/dCCC	+1.42	rACG/dCCT	+2.09	rGCA/dTCC	+2.66	
rUCA/dTCA	+2.85	rCCG/dCCG	+1.50	rUCG/dCCA	+1.61	rCCA/dTCG	+2.74	
rACU/dACT	+3.51	rCCC/dGCG	+1.77	rUCC/dGCA	+1.88	rGCU/dACC	+2.84	
rUCU/dTCT	+3.03	rGCC/dGCC	+1.69	rACG/dCCT	+2.09	rCCU/dACG	+2.92	
Average	+3.18		+1.60		+1.92		+2.79	+2.37

mismatch is calculated using the average of the averages previously calculated for each of the possible nearest-neighbor contexts. In general, the stability trend for these trimers averaged over all nearest-neighbor contexts in RNA/DNA, as shown in Table 2, is rG•dG > rU•dT > rA•dA > rC•dC, with numerical averages of +0.19, +1.20, +1.78 and +2.37 kcal/mol of free energy, respectively. The range of trimer stabilities for RNA/DNA internal mismatches observed in this study is the same general stability trend that was previously reported for these systems by Sugimoto (32).Hammond *et al*. (44) have developed a model for rationalizing the trends in nearest-neighbor thermodynamics by analyzing the spatial arrangement of amino and carbonyl partial charges. Once tertiary structures are available, it would be interesting to determine if those trends also hold for rA•dA, rC•dC, rG•dG and rU•dT mismatches in RNA/DNA duplexes.

Closing base pair trends. The nearest-neighbor free energy and enthalpy contributions for single nucleotide internal mismatches for both 5′-and 3′-contexts have been compared using Table 1 and Equation (4).

$$\text{Average NN } \Delta G^\circ_{37} \text{ with } 5'\frac{\text{rA}}{\text{T}}\text{pair} = \frac{G^\circ_{37}\begin{pmatrix}\text{rAA}\\ \text{TA}\end{pmatrix}+G^\circ_{37}\begin{pmatrix}\text{rAG}\\ \text{TG}\end{pmatrix}+G^\circ_{37}\begin{pmatrix}\text{rAC}\\ \text{TC}\end{pmatrix}+G^\circ_{37}\begin{pmatrix}\text{rAU}\\ \text{TT}\end{pmatrix}}{4}. \quad (4)$$

The stability trend for the 5′-closing Watson–Crick pair is: rG•dC > rC•dG > rA•dT > rU•dA, with average ΔG°_{37} of +0.27, +0.64, +0.91 and + 0.95 kcal/mol, respectively. This trend follows the expected property that G•C pairs are more stable than A/T pairs due to their differences in hydrogen bonding and stacking. The stability trend for the 3′-closing Watson–Crick pair is: rC–dG > rG–dC >> rA–dT > rU–dA, with average ΔG°_{37} of +0.02, +0.12, +1.19, +1.43 kcal/mol, respectively, which is a consistent result in that G•C pairs are more stable than A•T pairs because of the formation of an extra hydrogen bond as well as different stacking interactions. Equation (4) was used to calculate the average stabilities for the internal mismatches in Table 1 so that they may be compared. For the single internal mismatches reported here, the average stability trend is rG•dG > rU•dT > rA•dA > rC•dC, with average ΔG°_{37} of +0.08, +0.59, +0.90, and +1.20 kcal/mol, respectively. These data support the hypothesis that a nearest-neighbor context dependence exists for single internal mismatches in RNA/DNA duplexes, which is consistent with literature reports for single internal mismatches in RNA and DNA duplexes (18,45).

For example, Sugimoto *et al*. (32) hypothesized that the nearest-neighbor context dependence of rG•dG internal mismatch stabilities was relatively greater for this mismatch than for rU•dT, and rA•dA mismatches. A plausible explanation for this observation can be found in oligonucleotide structural studies of rG•dG single internal mismatches, with neighboring GC base pairs, have been observed to form a rG(anti)•dG(syn) conformation, stabilized by two hydrogen-bonds (18), as shown in Figure 2. The added stability is due to enhanced base stacking for two purines, especially since this conformation does not perturb the helical backbone structure of the duplex (18,32). In other structural studies, it has been observed that when dG•dG mismatches have at

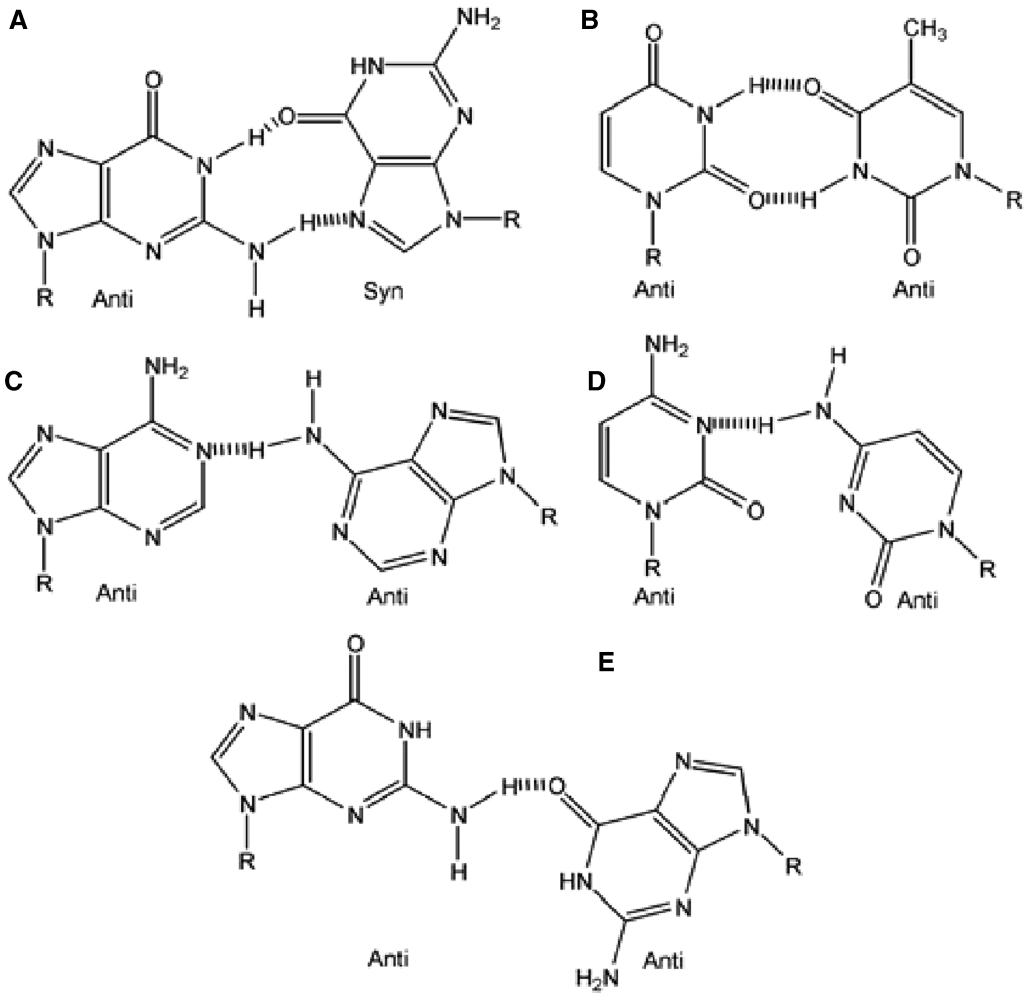

Figure 2. Proposed structures for single internal mismatches in RNA/DNA duplexes. The structures are (**A**) rG(anti)•dG(syn), (**B**) rU•dT, (**C**) rA•dA, (**D**) rC•dC, (**E**) rG (anti)•dG(anti).

least one A•T nearest-neighbor, then backbone distortion may allow the mismatch to form the less stable dG(anti)•dG(anti) conformation that may be stabilized by a single hydrogen bond (46,47). Therefore, the nearest-neighbor parameters for rG•dG internal mismatches in RNA/DNA duplexes are a combination of these two nearest-neighbor dependent conformations. Similar observations were made by Peyret *et al.* (18) in the evaluation of dG•dG mismatches in DNA duplexes.

Comparison of single internal mismatches in RNA/DNA with DNA/DNA duplexes

The nearest-neighbor thermodynamics of internal dA•dA, dC•dC, dG•dG and dT•dT single internal mismatches in DNA/DNA duplexes have previously been reported (18) and the most and least stable of the trimer contexts for single internal mismatches are dGGC/dGGC and dACT/dACT, which contribute from −2.22 to +2.66 kcal/mol, respectively. Trimer stability upper and lower limits in RNA/DNA duplexes, when compared to DNA/DNA duplexes, have the same contexts of rGGC/dGCC and rACU/dACT which have free energies of −1.40 and +3.51 kcal/mol, respectively. This result shows that these trimers are ∼0.84 kcal/mol more stable in DNA/DNA duplexes than in RNA/DNA duplexes. This can be rationalized by considering the combined effects of nearest-neighbor identity, base stacking and hydrogen bond interactions. The effect of nearest-neighbor identity upon an internal mismatch in RNA/DNA duplexes has already been discussed. For the contribution of stacking energies to RNA/DNA duplexes, one must consider the differences in RNA and DNA oligonucleotide geometries. RNA duplexes exhibit an A-type backbone conformation characterized by a C3′-endo sugar conformation with strong stacking (48). In contrast, DNA duplexes exhibit a B-form backbone conformation characterized by a C2′-endo sugar conformation and weaker stacking than observed for RNA (35,37). However, when these two strand types are annealed together in a hybrid duplex, structural studies have shown that RNA sugars usually remain in the A-form C3′-endo conformation, and the DNA sugars usually assume an equilibrium between the C2′ and the C3′-endo conformations (49,50).

The observed conformational equilibrium in DNA sugar puckers may interfere with the contribution of the base stacking energies to the hybrid duplex, which are affected by the oligonucleotide backbone distortions required to accommodate an internal mismatch within the hybrid duplex. Such backbone distortions are required so that the bases in an internal mismatch can obtain the optimal geometry for hydrogen bond formation to stabilize the duplex.

Peyret *et al.* (18) reported that, on average, the order of stabilities for internal dG•dG, dT•dT, dA•dA and dC•dC single internal mismatches in DNA/DNA duplex trimers are −0.23, +0.43, +0.48 and +0.97 kcal/mol, respectively. This is similar to the trend observed in this study for RNA/DNA duplexes for +0.08, +0.59, +0.90 and +1.20 kcal/mol for rG•dG, rU•dT, rA•dA and rC•dC internal mismatches, respectively. The rationalization for this trend for DNA/DNA duplexes was made with consideration for base stacking and hydrogen bonding (Figure 2). The mismatch base pair dG•dG is most stable because it has two purines that stack well and it also forms two H-bonds in the G(anti)–G(syn) conformation. The dT•dT mismatch can form a wobble pair with two hydrogen bonds, but it stacks poorly due to the two pyrimidine bases. The dA•dA mismatch forms only a single hydrogen bond, but forms strong stacking interactions. The dC•dC mismatch is least stable of all mismatches because the mismatch stacks weakly with its nearest-neighbors and the mismatch structure can only form a single hydrogen bond at pH 7. At low pH, rC•dC can form a protonated pair with two hydrogen bonds as described previously (18). It appears that the same explanation can be applied to the stability trend for mismatches in RNA/DNA duplexes since the stability trends for both DNA/DNA and RNA/DNA duplex internal mismatches are the same. Overall, we hypothesize that the difference in duplex stabilites between the two systems results from the reduction of base stacking interactions due to the DNA strands rapid sugar pucker equilibrium between the C2′ and C3′-endo conformations. This 'unstacked' conformation, particularly for rA•dA mismatches, is supported by previous work done on RNA/DNA duplexes (32), which most likely also applies to rC•dC internal mismatches as well, given that these mismatches are the least stable internal mismatch reported here.

Comparison of single internal mismatches in RNA/DNA with RNA/RNA duplexes

The nearest-neighbor thermodynamics of internal rC•rC single internal mismatches in RNA/RNA duplexes has not been reported in the literature. To our knowledge there is only one complete set of rG•rU nearest-neighbor thermodynamic parameters in RNA duplexes (42), though some raw data for other single internal mismatches have been reported (45). Nonetheless, single mismatch trimer stabilities in RNA duplexes can be calculated and the average stability is: rGGC/rGGC, rGUC/rGUC and rGAC/rGAG, with average contributions to duplex stability of −2.29, −1.03 and +0.59 kcal/mol, respectively (45). Comparatively, the RNA/DNA trimers rGGC/dGGC, rGUC/dGTC and rGAC/dGAC, as shown in Table 2, contribute −1.40, +0.16 and +0.70 kcal/mol, respectively, to the duplex. These results show that these trimer contexts in RNA/RNA duplexes are 0.89, 0.87 and 0.11 kcal/mol more stable, respectively, than the comparable trimers in RNA/DNA duplexes. The nearest-neighbor parameters suggest that the difference in free energies in RNA/RNA compared to RNA/DNA duplexes is due to stacking effects. This hypothesis has been supported in the literature by Kierzek *et al.* (45). For instance, the average effect of the adjacent G•C or A•T nearest-neighbor identity upon rU•rU internal mismatches in RNA duplexes was observed to be 1.6 kcal/mol (45), compared to an average difference of 0.79 kcal/mol for RNA/DNA duplexes. Again, the rationalization for this observation was based upon the added stability of G•C pairs compared to A•T pairs due to the difference in the numbers of hydrogen bonds and stacking interactions. A difference in nearest-neighbor orientation was also observed for the trimer contexts rGUC/rGUC and rCUG/rCUG in RNA duplexes where the former context was reported to be 1 kcal/mol more favorable than the latter context. This observation is also consistent in RNA/DNA duplexes for the trimer contexts rGUC/dGTC and rCUG/dCTG where the former is observed to be 0.47 kcal/mol more stable than the latter context, which is within experimental error of the reported data for RNA duplexes. Kierzek *et al.* went on to discuss similarities between the contributions of the stacking and hydrogen bonding energies of the internal mismatches rG•rG and rU•rU to those reported for DNA duplexes (18,45). Because of the similarities for these contributions to RNA/DNA and DNA/DNA duplexes, it appears that hypothesis is applicable for the comparison of RNA/DNA to RNA/RNA duplexes. For example, the comparison of imino proton NMR spectra between rG•rG and rU•rU and dG•dG and dT•dT internal mismatches suggested that both mismatches have two hydrogen bonds in their respective duplex types. The order of the stability for internal mismatches in RNA/RNA duplexes is rG•rG > rU•rU > rA•rA, which is also similar to the stability trends for these mismatches in both RNA/DNA and DNA/DNA duplexes.

Significance of data with respect to probe and primer design

Stability trends for the internal mismatches rG•dG, rU•dT, rA•dA and rC•dC have been observed to have the same stability trends for RNA/RNA, RNA/DNA and DNA/DNA duplexes, which is consistent with the stability trend observations for the matched nearest-neighbor determination for RNA/DNA duplexes (24). However, the differences in magnitude of the energy contributions to their relative duplex stabilities are large enough that a full nearest-neighbor thermodynamic determination, including all mismatches and dangling ends, was warranted. To allow for optimal DNA primer and probe design and hybridization simulation to RNA targets *in silico*, the parameters reported here have been added

to the programs *Visual OMP* and *ThermoBLAST* at DNA Software Inc. (www.dnasoftware.com). *Visual OMP* uses established methods to extrapolate (from 1 M to as low as 50 mM monovalent salt) the thermodynamic data presented here, which were collected in 1 M NaCl salt buffer, to physiologically relevant salt concentrations (51) so that any RNA target can be folded under a variety of temperature and salt conditions to select the best primer or probe designs for various molecular biology applications. These designs can then be simulated so that monomer, homodimer and heterodimer species can be predicted at physiological salt conditions. *Thermo BLAST* is used to detect unwanted mishybridization events in any given set of RNA sequences in order to improve primer or probe specificity by reducing 'off-target' or false positive effects in any assay design. This work has shown that simple approximations of RNA/DNA hybridization using RNA/RNA or DNA/DNA parameters will not lead to an optimal design for the nucleic acid based hybridization technologies discussed in the introduction. For instance, the RNA/RNA sequence GCG<u>G</u>CGC/GCG<u>G</u>CGC has an experimental ΔG°_{37} of -9.33 kcal/mol and a T_m of 50.0°C at a total strand concentration of 100 μM, whereas the same sequence, calculated from RNA/DNA nearest-neighbor parameters is -4.41 kcal/mol and 44.9°C, a total difference of 4.92 kcal/mol and 5.1°C. The knowledge of RNA/DNA nearest-neighbor parameters will allow for the design of optimal DNA primers and probes that bind to target RNAs.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

US Department of Health and Human Services National Institutes of Health SBIR grants HG003255 and HG003923. Funding for open access charge: DNA Software, Inc.

Conflict of interest statement. None declared.

REFERENCES

1. Dasu,M.R., Devaraj,S., Park,S. and Jialal,I. (2010) Increased toll-like receptor (TLR) activation and TLR ligands in recently diagnosed type 2 diabetic subjects. *Diabetes Care*, **33**, 861–868.
2. Eis,P.S., Olson,M.C., Takova,T., Curtis,M.L., Olson,S.M., Vener,T.I., Ip,H.S., Vedvik,K.L., Bartholomay,C.T., Allawi,H.T. *et al*. (2001) An invasive cleavage assay for direct quantitation of specific RNAs. *Nat. Biotechnol.*, **19**, 673–676.
3. Fleischer,A., Duhamel,M., Lopez-Fernandez,L.A., Munoz,M., Rebollo,M.P., Alvarez-Franco,F. and Rebollo,A. (2007) Cascade of transcriptional induction and repression during IL-2 deprivation-induced apoptosis. *Immunol. Lett.*, **112**, 9–29.
4. Murray,E.L. and Schoenberg,D.R. (2008) Application of the invader RNA assay to the polarity of vertebrate mRNA decay. *Methods Mol. Biol.*, **419**, 259–276.
5. Neelima,P.S. and Rao,A.J. (2008) Gene expression profiling during Forskolin induced differentiation of BeWo cells by differential display RT-PCR. *Mol. Cell Endocrinol.*, **281**, 37–46.
6. Streit,S., Michalski,C.W., Erkan,M., Kleeff,J. and Friess,H. (2009) Northern blot analysis for detection and quantification of RNA in pancreatic cancer cells and tissues. *Nat. Protoc.*, **4**, 37–43.
7. Yu,X.H., Zhang,H., Wang,Y.H., Liu,L.J., Teng,Y. and Liu,P. (2009) Time-dependent reduction of glutamine synthetase in retina of diabetic rats. *Exp. Eye Res.*, **89**, 967–971.
8. Alwine,J.C., Kemp,D.J. and Stark,G.R. (1977) Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc. Natl Acad. Sci. USA*, **74**, 5350–5354.
9. Jin,L. and Lloyd,R.V. (1997) In situ hybridization: methods and applications. *J. Clin. Lab. Anal.*, **11**, 2–9.
10. Orlando,C., Pinzani,P. and Pazzagli,M. (1998) Developments in quantitative PCR. *Clin. Chem. Lab. Med.*, **36**, 255–269.
11. Kulesh,D.A., Clive,D.R., Zarlenga,D.S. and Greene,J.J. (1987) Identification of interferon-modulated proliferation-related cDNA sequences. *Proc. Natl Acad. Sci. USA*, **84**, 8453–8457.
12. Fluiter,K., Mook,O.R., Vreijling,J., Langkjaer,N., Hojland,T., Wengel,J. and Baas,F. (2009) Filling the gap in LNA antisense oligo gapmers: the effects of unlocked nucleic acid (UNA) and 4′-C-hydroxymethyl-DNA modifications on RNase H recruitment and efficacy of an LNA gapmer. *Mol. Biosyst.*, **5**, 838–843.
13. Koizumi,M. (2007) True antisense oligonucleotides with modified nucleotides restricted in the N-conformation. *Curr. Top Med. Chem.*, **7**, 661–665.
14. Dmochowski,I.J. and Tang,X. (2007) Taking control of gene expression with light-activated oligonucleotides. *Biotechniques*, **43**, 161, 163, 165 passim.
15. SantaLucia,J. Jr (2007) Physical principles and visual-OMP software for optimal PCR design. *Methods Mol. Biol.*, **402**, 3–34.
16. Allawi,H.T. and SantaLucia,J. Jr (1997) Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry*, **36**, 10581–10594.
17. Allawi,H.T. and SantaLucia,J. Jr (1998) Thermodynamics of internal C.T mismatches in DNA. *Nucleic Acids Res.*, **26**, 2694–2701.
18. Peyret,N., Seneviratne,P.A., Allawi,H.T. and SantaLucia,J. Jr (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. *Biochemistry*, **38**, 3468–3477.
19. Allawi,H.T. and SantaLucia,J. Jr (1998) NMR solution structure of a DNA dodecamer containing single G.T mismatches. *Nucleic Acids Res.*, **26**, 4925–4934.
20. Davis,A.R. and Znosko,B.M. (2007) Thermodynamic characterization of single mismatches found in naturally occurring RNA. *Biochemistry*, **46**, 13425–13436.
21. Saito,H. and Richardson,C.C. (1981) Processing of mRNA by ribonuclease III regulates expression of gene 1.2 of bacteriophage T7. *Cell*, **27**, 533–542.
22. Saito,M. and Nakamura,T. (2005) Two point mutations identified in emmer wheat generate null Wx-A1 alleles. *Theor. Appl. Genet.*, **110**, 276–282.
23. Bae,S.H. and Choi,B.S. (2001) Structure of influenza a virus promoter and its implications for viral RNA synthesis. *Scientific World J.*, **1**, 812–814.
24. Sugimoto,N., Nakano,S., Katoh,M., Matsumura,A., Nakamuta,H., Ohmichi,T., Yoneyama,M. and Sasaki,M. (1995) Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry*, **34**, 11211–11216.
25. SantaLucia,J. Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA*, **95**, 1460–1465.
26. Pont-Kingdon,G., Margraf,R.L., Sumner,K., Millson,A., Lyon,E. and Schutz,E. (2008) Design and application of noncontinuously binding probes used for haplotyping and genotyping. *Clin. Chem.*, **54**, 990–999.
27. Kierzek,R., Caruthers,M.H., Longfellow,C.E., Swinton,D., Turner,D.H. and Freier,S.M. (1986) Polymer-supported RNA synthesis and its application to test the nearest-neighbor model for duplex stability. *Biochemistry*, **25**, 7840–7846.
28. Marugg,J.E., Piel,N., McLaughlin,L.W., Tromp,M., Veeneman,G.H., van der Marel,G.A. and van Boom,J.H. (1984) Polymer supported DNA synthesis using hydroxybenzotriazole

activated phosphotriester intermediates. *Nucleic Acids Res*., **12**, 8639–8651.
29. Olmsted,M.C., Anderson,C.F. and Record,M.T. Jr (1989) Monte Carlo description of oligoelectrolyte properties of DNA oligomers: range of the end effect and the approach of molecular and thermodynamic properties to the polyelectrolyte limits. *Proc. Natl Acad. Sci. USA*, **86**, 7766–7770.
30. 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.
31. SantaLucia,J. Jr and Hicks,D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct*., **33**, 415–440.
32. Sugimoto,N., Nakano,M. and Nakano,S. (2000) Thermodynamics-structure relationship of single mismatches in RNA/DNA duplexes. *Biochemistry*, **39**, 11270–11281.
33. Manning,G.S. (1977) A field-dissociation relation for polyelectrolytes with an application to field-induced conformational changes of polynucleotides. *Biophys. Chem*., **7**, 189–192.
34. SantaLucia,J. Jr (2000) *The Use of Spectrosocopic Techniques in the Study of DNA Stability*. University Press, Oxford, NY.
35. Bommarito,S., Peyret,N. and SantaLucia,J. Jr (2000) Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res*., **28**, 1929–1934.
36. Watkins,N.E. Jr and SantaLucia,J. Jr (2005) Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes. *Nucleic Acids Res*., **33**, 6258–6267.
37. SantaLucia,J. Jr, Allawi,H.T. and Seneviratne,P.A. (1996) Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, **35**, 3555–3562.
38. 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.
39. Holbrook,J.A., Capp,M.W., Saecker,R.M. and Record,M.T. Jr (1999) Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: interpretation in terms of coupled processes of formation and association of single-stranded helices. *Biochemistry*, **38**, 8409–8422.
40. Rouzina,I. and Bloomfield,V.A. (2001) Force-induced melting of the DNA double helix. 2. Effect of solution conditions. *Biophys. J*., **80**, 894–900.
41. Chaires,J.B. (1997) Possible origin of differences between van't Hoff and calorimetric enthalpy estimates. *Biophys. Chem*., **64**, 15–23.
42. 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. USA*, **83**, 9373–9377.
43. Hall,T.S., Pancoska,P., Riccelli,P.V., Mandell,K. and Benight,A.S. (2001) Sequence context and thermodynamic stability of a single base pair mismatch in short deoxyoligonucleotide duplexes. *J. Am. Chem. Soc*., **123**, 11811–11812.
44. Hammond,N.B., Tolbert,B.S., Kierzek,R., Turner,D.H. and Kennedy,S.D. (2010) RNA internal loops with tandem AG pairs: the structure of the 5′GAGU/3′UGAG loop can be dramatically different from others, including 5′AAGU/3′UGAA. *Biochemistry*, **49**, 5817–5827.
45. Kierzek,R., Burkard,M.E. and Turner,D.H. (1999) Thermodynamics of single mismatches in RNA duplexes. *Biochemistry*, **38**, 14214–14223.
46. Borden,K.L., Jenkins,T.C., Skelly,J.V., Brown,T. and Lane,A.N. (1992) Conformational properties of the G.G mismatch in d(CGC GAATTGGCG)2 determined by NMR. *Biochemistry*, **31**, 5411–5422.
47. Faibis,V., Cognet,J.A., Boulard,Y., Sowers,L.C. and Fazakerley,G.V. (1996) Solution structure of two mismatches G.G and I.I in the K-ras gene context by nuclear magnetic resonance and molecular dynamics. *Biochemistry*, **35**, 14452–14464.
48. 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 XGGCCp, XCCGGp, XGGCCYp, and XCCGGYp helixes. *Biochemistry*, **24**, 4533–4539.
49. Gyi,J.I., Lane,A.N., Conn,G.L. and Brown,T. (1998) Solution structures of DNA.RNA hybrids with purine-rich and pyrimidine-rich strands: comparison with the homologous DNA and RNA duplexes. *Biochemistry*, **37**, 73–80.
50. Gyi,J.I., Conn,G.L., Lane,A.N. and Brown,T. (1996) Comparison of the thermodynamic stabilities and solution conformations of DNA.RNA hybrids containing purine-rich and pyrimidine-rich strands with DNA and RNA duplexes. *Biochemistry*, **35**, 12538–12548.
51. Petruska,J. and Goodman,M.F. (1995) Enthalpy-entropy compensation in DNA melting thermodynamics. *J. Biol. Chem*., **270**, 746–750.