Nearest Neighbor Parameters for Inosine·Uridine Pairs in RNA Duplexes

Daniel J. Wright, Jamie L. Rice, Dawn M. Yanker, and Brent M. Znosko*

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

Received August 18, 2006; Revised Manuscript Received January 12, 2007

ABSTRACT: An enzyme family known as adenosine deaminases that act on RNA (ADARs) catalyzes adenosine deamination in RNA. ADARs act on RNA that is largely double-stranded and convert adenosine to inosine, resulting, in many cases, in an I·U pair. Thermodynamic parameters derived from optical melting studies are reported for a series of 14 oligoribonucleotides containing single I·U pairs adjacent to Watson−Crick pairs. In order to determine unique linearly independent nearest neighbor parameters for I·U pairs, four duplexes containing 3′-terminal I·U pairs and four duplexes containing 5′-terminal I·U pairs have also been thermodynamically characterized. This data was combined with previously published data of seven duplexes containing internal, terminal, or tandem I·U pairs from Strobel et al. [Strobel, S. A., Cech, T. R., Usman, N., and Beigelman, L. (1994) *Biochemistry 33*, 13824−13838] and Serra et al. [Serra, M. J., Smolter, P. E., and Westhof, E. (2004) *Nucleic Acids Res. 32*, 1824−1828]. On average, a duplex with an internal I·U pair is 2.3 kcal/mol less stable than the same duplex with an A-U pair, however, a duplex with a terminal I·U pair is 0.8 kcal/mol more stable than the same duplex with an A-U pair. Although isosteric with a G-U pair, on average, a duplex with an internal I·U pair is 1.9 kcal/mol less stable than the same duplex with a G-U pair, however, a duplex with a terminal I·U pair is 0.9 kcal/mol more stable than the same duplex with a G-U pair. Duplexes with tandem I·U pairs are on average 5.9 and 3.8 kcal/mol less stable than the same duplex with tandem A-U or tandem G-U pairs, respectively. Using the combined thermodynamic data and a complete linear least-squares fitting routine, nearest neighbor parameters for all nearest neighbor combinations of I·U pairs and an additional parameter for terminal I·U pairs have been derived.

We are beginning to understand the thermodynamics, dynamics, and structures of various RNA secondary structure motifs containing standard nucleotides. Understanding motifs containing nonstandard nucleotides is also important, because they too play an important biological role. However, thermodynamics of many nonstandard nucleotides, such as inosine (I[1]), have not yet been systematically investigated. Inosine occurs naturally in the wobble position of the anticodon loop of tRNA and is partly responsible for the degeneracy of the genetic code. Inosine in the anticodon loop can hydrogen bond with adenosine, cytidine, or uridine of the mRNA codon. Inosine is also a common mutation, resulting from oxidative deamination of adenosine (*1*), which usually results in conversion of an A-U pair to an I·U pair.

An enzyme family known as adenosine deaminases that act on RNA (ADARs) catalyzes adenosine deamination in RNA. ADARs have been discovered and characterized in *Xenopus laevis* (*2*), mammals (*3*), birds (*4*), frogs (*5*), fish (*6*), flies (*7*), and worms (*8*). ADARs act on RNA that is largely double-stranded and convert adenosine to inosine. For example, in the mammalian brain, one out of every 17 000 nucleotides in mRNA is inosine (*9*). Inosine is then translated as guanosine and is recognized by most enzymes as guanosine (*1*). Thus, ADARs change the primary sequence information of RNA and increase the diversity of mRNAs and proteins. In addition, because of the ability of its functional groups to form hydrogen bonds, inosine can also change the structure of RNA. For example, in Watson−Crick duplex regions, an adenosine to inosine conversion changes an A-U Watson−Crick pair to an I·U noncanonical pair (Figure 1). Although this may appear as a slight shift in the conformation of the base pair, this shift may have a significant impact on the local helix geometry. Therefore, ADARs could affect any biological process that relies on sequence- or structure-specific interactions with RNA. Incorporation of inosine into RNA has led to a wide variety of functional consequences, including creation of splice sites, sequestering of RNA in the nucleus, and regulation of gene expression in the mammalian brain, to name a few (*1*, and references therein). Recently, studies have shown that conversion of adenosine to inosine prevents double-stranded RNA from entering the RNA interference pathway (*8*, *10*).

Although two studies have investigated thermodynamics of base pairing involving deoxyinosine (*11*, *12*) and many studies have investigated the role of ADARs *in vivo* (*1*, and references therein), relatively few studies have investigated thermodynamics of RNA duplexes containing inosine. Previous thermodynamic studies with RNA containing inosine involve replacement of guanosine with inosine in order to determine the thermodynamic contribution of the amino group of guanosine (*13*−*15*). One study investigated the stability of tandem I·U pairs in RNA (*16*), however, since

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

[1] Abbreviations: ADARs, adenosine deaminases that act on RNA; dI, deoxyinosine; I, inosine; TLC, thin-layer chromatography.

10.1021/bi0616910 CCC: $37.00
Published on Web 03/23/2007

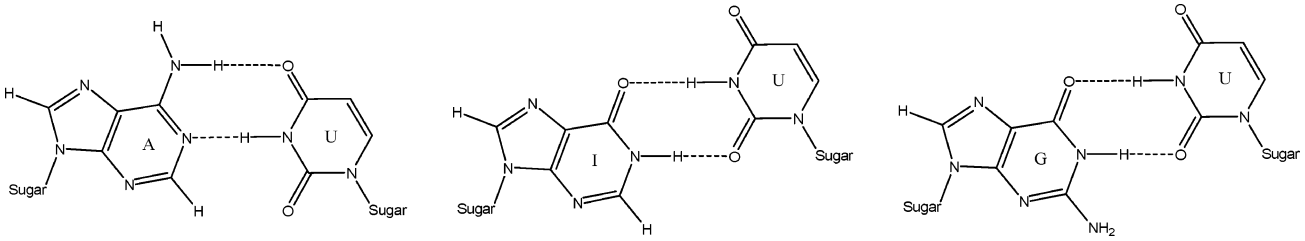

FIGURE 1: Conformation of a Watson−Crick A-U pair (left), likely conformation of an I•U pair (center), and conformation of a wobble G-U pair (right). Dashed lines represent possible hydrogen bonds.

ADARs act on RNA that is largely double-stranded and convert adenosine to inosine, many instances of inosine are likely to involve single I•U pairs located within a Watson−Crick duplex. It has been shown that molecules that are less stable because they contain mismatches (such as I•U pairs) are deaminated much less frequently by ADARs than molecules with long stretches of Watson−Crick pairs (*1*).

Although an I•U base pair can adopt a conformation with the same number of hydrogen bonds as a Watson−Crick A-U pair (Figure 1), Strobel et al. reported that the stability of a single I•U pair located within a Watson−Crick duplex, $\binom{\text{G I C}}{\text{CUG}}$, is 1.8 kcal/mol less stable than an A-U pair (*17*). Similarly, although an I•U base pair can adopt a conformation isosteric with a wobble G-U pair (Figure 1), the I•U pair located within $\binom{\text{G I C}}{\text{CUG}}$ is 0.8 kcal/mol less stable than a G-U pair (*17*). However, a complete, systematic thermodynamic study used to derive nearest neighbor parameters for I•U pairs has not been published for RNA oligonucleotides. Thermodynamic parameters for I•U pairs would provide a reliable prediction of stabilities of duplexes containing I•U pairs, thus leading to comparison of the stabilities and predicted secondary structures of ADAR substrates before and after modification and would help in the design of oligonucleotide probes containing inosine.

Here, we have measured stabilities of a set of RNA oligonucleotide duplexes containing single I•U pairs within a Watson−Crick duplex and adjacent to all possible nearest neighbor combinations. In addition, stabilities of RNA oligonucleotides with 3′- and 5′-terminal I•U pairs adjacent to all possible nearest neighbor combinations have also been measured. This data was combined with data for internal, terminal, and tandem I•U pairs published previously (*16*, *17*). Stabilities of duplexes containing I•U pairs are compared to stabilities of the same duplexes containing A-U and G-U pairs, as predicted by the nearest neighbor model (*18*). A nearest neighbor model to calculate stability of RNA oligonucleotides containing I•U pairs has been derived.

MATERIALS AND METHODS

Design of Sequences. Duplexes containing internal I•U pairs were designed to have melting temperatures at approximately 50 °C and to have minimal formation of hairpin structures or misaligned duplexes. Terminal G-C pairs were chosen to prevent end fraying during melting experiments. All duplexes studied here containing internal I•U pairs have the same sequence, except for the nearest neighbors directly adjacent to the I•U pairs. The combination of Watson−Crick pairs adjacent to the I•U pairs were chosen to maximize the number of each particular Watson−Crick pair adjacent to the I•U pair. The I•U pair was placed directly in the center of the duplex. Duplexes studied here containing terminal I•U pairs were also designed to have melting temperatures at approximately 50 °C and to have minimal formation of hairpin structures or misaligned duplexes. All duplexes studied here containing terminal I•U pairs have the same sequence except for the nearest neighbor Watson−Crick pair directly adjacent to the I•U pair. Four duplexes, one with each Watson−Crick pair adjacent to the I•U pair, were studied for both 3′- and 5′-terminal I•U pairs.

RNA Synthesis and Purification. Oligonucleotides containing internal I•U pairs were synthesized on controlled pore glass support with an Applied Biosystems 392 DNA/RNA synthesizer at the University of Rochester (Rochester, NY) using the phosphoramidite method (*19*, *20*). Standard support, standard phosphoramidites, and inosine phosphoramidites with 2′-*tert*-butyl dimethyl silyl ether protecting groups were acquired from Glen Research (Baltimore, MD). Glass beads and base-labile protecting groups were removed by treatment with a 3:1 (v/v) ammonia/ethanol solution at 55 °C overnight (*21*). Quik-Sep disposable filter columns from Perkin-Elmer (Norton, OH) were used to filter oligonucleotides from support. Removal of the silyl protecting group was achieved with 9:1 (v/v) triethylamine trihydrofluoride/*N*,*N*-dimethylformamide at 55 °C for 2 h. The sample was precipitated with butanol and redissolved in 5 mM ammonium bicarbonate pH 4.5 or 7.0. This solution was passed through a Waters Sep-Pak C18 chromatography column to remove inorganic salts. Purification of the oligonucleotide was done by preparative thin-layer chromatography (TLC). A large preparative TLC plate (20 cm × 20 cm, 500 μm thick) was used with 55:35:10 (v/v/v) 1-propanol/ammonia/water as the solvent (*22*). The main product band was identified with UV light and scraped from the plate. Distilled water was used to extract RNA from silica. The Sep-Pak procedure was repeated to desalt the sample. Purity of oligonucleotides was checked by analytical TLC similar to that described above. Samples that resulted in multiple bands on the analytical TLC plate were run again on a large, preparative TLC plate. Oligonucleotides containing terminal I•U pairs were ordered from the Keck Lab at Yale University (New Haven, CT) and were purified as described above.

Concentration Calculations and Duplex Formation. The total concentration of each single strand was calculated from the absorbance reading at 260 nm at 80 °C and the extinction coefficient using Beer's law. Samples were diluted so that the absorbance was between 0.2 and 2.0. The absorbance was measured at 80 °C to disrupt any single-strand folding. Extinction coefficients of the single strands were calculated from the molar extinction coefficient of each base at 260 nm and pairwise values for each nearest neighbor combination at 260 nm. Values for standard bases were obtained from

Integrated DNA Technologies (Coralville, IA) (*23*), and values for inosine were obtained from Proligo (Boulder, CO) (*24*). Individual single strand concentrations were used to mix equal molar amounts of non-self-complementary strands to form a duplex containing an I·U pair.

Optical Melting Experiments. Newly formed duplexes were lyophilized and redissolved in 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH = 7.0. A melt scheme was designed that consisted of a three dilution series resulting in nine samples to allow for a concentration range typically >50-fold. Using a heating rate of 1 °C/min on a Beckman-Coulter DU800 spectrophotometer with a Beckman-Coulter high performance temperature controller, curves of absorbance at 260 nm versus temperature were obtained.

Determination of Thermodynamic Parameters for Duplexes. MeltWin (*25*) was used to fit melting curves to a two-state model, assuming linear sloping baselines and temperature-independent $\Delta H°$ and $\Delta S°$ values (*26*, *27*). Additionally, T_M values at different concentrations were used to calculate thermodynamic parameters according to Borer et al. (*28*):

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

The Gibb's free energy change at 37 °C was calculated as

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

MeltWin (*25*) requires the sequence of the duplex in order to calculate the concentration of the duplex from the high temperature absorbance and extinction coefficients. Since MeltWin does not recognize inosine, one of two approaches was taken. Originally, concentrations were calculated manually using inosine extinction coefficients (described above), and these values were inserted into MeltWin. It was then noticed that inosine extinction coefficients were relatively similar to cytidine extinction coefficients. The molar extinction coefficient of the individual inosine base versus the individual cytidine base differs by only 3%. Pairwise values for each nearest neighbor combination containing inosine differed, on average, by 4% from pairwise values for each nearest neighbor combination containing cytidine. We compared thermodynamics derived from analysis of log C_T plots when concentrations were inserted into MeltWin to thermodynamics derived when inosine was replaced with cytidine when entering the sequence into MeltWin; the thermodynamics derived from these two methods differed by <1%. Therefore, for simplicity, most thermodynamics were derived from replacing inosine with cytidine when entering the sequence into MeltWin.

Determination of the Contribution of I·U Pairs to Duplex Thermodynamics. The total free energy change for duplex formation can be approximated by a nearest neighbor model (*18*) that is the sum of energy increments for helix initiation and nearest neighbor interactions between base pairs. For example, for internal I·U pairs,

$$\Delta G°_{37}\begin{pmatrix}\text{GCA ICGC}\\ \text{CGUUGCG}\end{pmatrix} = \Delta G°_{37,i} + \Delta G°_{37}(\text{GCCG}) + \Delta G°_{37}\begin{pmatrix}\text{CA}\\ \text{GU}\end{pmatrix} + \Delta G°_{37}\begin{pmatrix}\text{A I}\\ \text{UU}\end{pmatrix} + \Delta G°_{37}\begin{pmatrix}\text{IC}\\ \text{UG}\end{pmatrix} + \Delta G°_{37}\begin{pmatrix}\text{CG}\\ \text{GC}\end{pmatrix} + \Delta G°_{37}\begin{pmatrix}\text{GC}\\ \text{CG}\end{pmatrix} \quad (3)$$

where $\Delta G°_{37,i}$ is the free energy change for duplex initiation, 4.09 kcal/mol, and remainder of the terms are individual nearest neighbor values (*18*). Therefore, rearranging eq 3 can solve for the contribution of internal inosine nearest neighbors to duplex stability:

$$\Delta G°_{37}\begin{pmatrix}\text{A I}\\ \text{UU}\end{pmatrix} + \Delta G°_{37}\begin{pmatrix}\text{IC}\\ \text{UG}\end{pmatrix} = \Delta G°_{37}\begin{pmatrix}\text{GCA ICGC}\\ \text{CGUUGCG}\end{pmatrix} - \Delta G°_{37,i} - \Delta G°_{37}\begin{pmatrix}\text{GC}\\ \text{CG}\end{pmatrix} - \Delta G°_{37}\begin{pmatrix}\text{CA}\\ \text{GU}\end{pmatrix} - \Delta G°_{37}\begin{pmatrix}\text{CG}\\ \text{GC}\end{pmatrix} - \Delta G°_{37}\begin{pmatrix}\text{GC}\\ \text{CG}\end{pmatrix} \quad (4)$$

where $\Delta G°_{37}\left(\begin{smallmatrix}\text{GCA ICGC}\\ \text{CGUUGCG}\end{smallmatrix}\right)$ is the value determined by optical melting experiments. More specifically,

$$\Delta G°_{37}\begin{pmatrix}\text{A I}\\ \text{UU}\end{pmatrix} + \Delta G°_{37}\begin{pmatrix}\text{IC}\\ \text{UG}\end{pmatrix} = -8.34 - 4.09 - (-3.42) - (-2.11) - (-2.36) - (-3.42) = -1.12 \text{ kcal/mol} \quad (5)$$

This value, however, is the sum of the contribution from two I·U nearest neighbors. Similar calculations were done for tandem I·U pairs. Calculations with internal and tandem I·U pairs result in a $\Delta G°_{37}$ value that is the sum of multiple nearest neighbor contributions to duplex stability. Similar calculations with duplexes containing terminal I·U pairs result in a $\Delta G°_{37}$ value that is an individual nearest neighbor contribution to duplex stability (as opposed to the $\Delta G°_{37}$ value representing a sum of nearest neighbor contributions as described above for internal or tandem I·U pairs):

$$\Delta G°_{37}\begin{pmatrix}\text{GCGC A I}\\ \text{CGCGUU}\end{pmatrix} = \Delta G°_{37,i} + \Delta G°_{37}\begin{pmatrix}\text{GC}\\ \text{CG}\end{pmatrix} + \Delta G°_{37} + \Delta G°_{37}\begin{pmatrix}\text{CG}\\ \text{GC}\end{pmatrix} + \Delta G°_{37}\begin{pmatrix}\text{GC}\\ \text{CG}\end{pmatrix} + \Delta G°_{37}\begin{pmatrix}\text{CA}\\ \text{GU}\end{pmatrix} + \Delta G°_{37}\begin{pmatrix}\text{A I}\\ \text{UU}\end{pmatrix} \quad (6)$$

where $\Delta G°_{37,i}$ is the free energy change for duplex initiation, 4.09 kcal/mol, and remainder of the terms are individual nearest neighbor values (*18*). Therefore, rearranging eq 6 can solve for the contribution of terminal inosine nearest neighbors to duplex stability:

$$\Delta G°_{37}\begin{pmatrix}\text{A I}\\ \text{UU}\end{pmatrix} = \Delta G°_{37}\begin{pmatrix}\text{GCGCA I}\\ \text{CGCGUU}\end{pmatrix} - \Delta G°_{37,i} - \Delta G°_{37}\begin{pmatrix}\text{GC}\\ \text{CG}\end{pmatrix} - \Delta G°_{37}\begin{pmatrix}\text{CG}\\ \text{GC}\end{pmatrix} - \Delta G°_{37}\begin{pmatrix}\text{GC}\\ \text{CG}\end{pmatrix} - \Delta G°_{37}\begin{pmatrix}\text{CA}\\ \text{GU}\end{pmatrix} \quad (7)$$

where $\Delta G°_{37}\left(\begin{smallmatrix}\text{GCGCA I}\\ \text{CGCGUU}\end{smallmatrix}\right)$ is the value determined by optical melting experiments. More specifically,

$$\Delta G°_{37}\begin{pmatrix}\text{A I}\\ \text{UU}\end{pmatrix} = -9.84 - 4.09 - (-3.42) - (-2.36) - (-3.42) - (-2.11) = -2.62 \text{ kcal/mol} \quad (8)$$

Similar calculations were done for $\Delta H°$ and $\Delta S°$.

Linear Regression and Determination of Linearly Independent I·U Nearest Neighbor Values. Excel's LINEST

function was used to determine the contribution to duplex stability from each individual I·U nearest neighbor. The eight possible I·U/Watson−Crick nearest neighbor combinations, the three tandem I·U nearest neighbor combinations, and a terminal I·U parameter were used as variables. The contribution from a combination of I·U nearest neighbor pairs (as described above) for internal and tandem I·U pairs and the contribution from terminal I·U nearest neighbors were set as the constants. Excel's LINEST function used linear regression to simultaneously solve for each variable. Originally, data was only collected for duplexes containing internal I·U pairs. Linear regression of this data did not result in unique, linearly independent nearest neighbor parameters for $\Delta G°_{37}$. As a result, data was collected on duplexes containing 3′- and 5′-terminal I·U pairs, and data was added from previous studies (*16*, *17*). Linear regression with internal, tandem, and terminal I·U pairs resulted in a set of linearly independent parameters for $\Delta G°_{37}$, $\Delta H°$, and $\Delta S°$.

RESULTS

Thermodynamic Parameters. Thermodynamic parameters of duplex formation derived from T_M^{-1} vs $\log(C_T/4)$ and from fitting the shape of each melting curve to the two-state model are listed in Table 1. The model used to obtain thermodynamic parameters from melting data assumes that the helix to coil transition is two-state. Typically, agreement within 15% for $\Delta H°$ values from curve fitting and from log C_T plots is considered a reasonable indication of the two-state transition. All oligomers listed in Table 1 meet this criterion.

Thermodynamic Comparison to Duplexes Containing A-U and G-U Pairs. Table 2 shows the $\Delta G°_{37}$ of measured duplexes containing I·U pairs compared to nearest neighbor calculations of the same duplexes containing A-U and G-U pairs. On average, duplexes containing internal I·U pairs are 2.4 and 2.0 kcal/mol less stable than equivalent duplexes with internal A-U and internal G-U pairs, respectively. On average, duplexes containing terminal I·U pairs are 0.8 and 0.9 kcal/mol more stable than equivalent duplexes with terminal A-U and terminal G-U pairs, respectively. On average, duplexes containing tandem I·U pairs are 5.9 and 3.8 kcal/mol less stable than equivalent duplexes with tandem A-U and tandem G-U pairs, respectively. Similar calculations and comparisons for enthalpies and entropies are available in Supporting Information (Tables S1 and S2).

Contribution of I·U Nearest Neighbors to Duplex Thermodynamics. The contribution of I·U nearest neighbors to duplex stability is defined in eqs 4, 5, 7, and 8. These contributions are listed in Table S3 (Supporting Information). Similar calculations were done for $\Delta H°$ and $\Delta S°$ and are also listed in Table S3.

I·U Nearest Neighbor Parameters. As described in Materials and Methods, results from optical melting studies can be used to calculate the individual contribution to duplex thermodynamics of each individual I·U nearest neighbor. These nearest neighbor parameters are listed in Table 3. All I·U/Watson−Crick nearest neighbors contribute a negative $\Delta H°$ and a negative $\Delta S°$ to duplex thermodynamics. $\binom{\mathrm{U\ I}}{\mathrm{AU}}$ and $\binom{\mathrm{I\ A}}{\mathrm{UU}}$ contribute a positive $\Delta G°_{37}$ while all other I·U/Watson−Crick combinations contribute a negative $\Delta G°_{37}$ to duplex stability. In addition to the I·U/Watson−Crick nearest neighbor contribution, terminal I·U pairs contribute an additional negative $\Delta H°$, positive $\Delta S°$, and negative $\Delta G°_{37}$. All tandem I·U pairs contribute a positive $\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$. Deviations for the I·U/Watson−Crick nearest neighbors and additional terminal I·U pair are, on average, 2.44 kcal/mol for $\Delta H°$, 7.6 eu for $\Delta S°$, and 0.38 kcal/mol for $\Delta G°_{37}$. It is important to note that a larger deviation is associated with the tandem nearest neighbors due to the much smaller sample size. Deviations for tandem I·U nearest neighbors are, on average, 6.27 kcal/mol for $\Delta H°$, 19.5 eu for $\Delta S°$, and 0.96 kcal/mol for $\Delta G°_{37}$.

The new I·U nearest neighbor parameters, together with Watson−Crick nearest neighbor parameters (*18*), were used to predict the stabilities of the duplexes measured (Table 1). Average deviations between the predicted values and the measured values are 5.1%, 4.6%, 5.1%, and 2.8 °C for $\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and T_M, respectively. These values are comparable to those reported for RNA Watson−Crick (3.2%, 6.0%, 6.8%, and 1.3 °C, respectively) (*18*) and dI·T (3.5%, 4.8%, 5.0%, and 1.2 °C, respectively) (*12*) nearest neighbor parameters.

DISCUSSION

Thermodynamic Parameters. Thermodynamic parameters derived from optical melting experiments are listed in Table 1. Melting temperatures of all duplexes were dependent on oligonucleotide concentration, suggesting that a duplex was formed. Although the single strands were designed to form the desired duplexes when mixed, it is possible that some of the single strands could self-associate and form a duplex. For example, 5′GCGICGC3′ was mixed with 3′CGCUGCG5′ to form a duplex with an I·U pair. However, 5′GCGICGC3′ may form a duplex with itself, resulting in a duplex with an I·I pair. Several arguments can be used to show that self-association is unlikely. First, each of the duplex melting curves had a single, sharp transition. If one strand was self-assembling, a contribution from the other strand or the designed duplex would also be expected, resulting in multiple transitions or a broad transition during the optical melting experiment. Second, assuming that inosine provides the same stability as adenosine, the experimental free energy is much closer to the predicted stability for the designed duplex than to the self-assembled duplex. Third, preliminary NMR spectra show one set of sharp peaks (unpublished data), suggesting one conformation and one duplex present. If self-assembly were a problem, multiple sets of peaks from self-assembly of strands, free single strands, and designed duplex may be observed. Fourth, r(GCGICGC)$_2$ was melted (unpublished data) and the resulting thermodynamics were inconsistent with the thermodynamics obtained from melting what is likely the designed duplex. Last, the thermodynamic trends are consistent with the designed duplexes. For example, designed duplexes with I·U pairs adjacent to two G-C pairs are more stable than designed duplexes with I·U pairs next to one G-C pair which is more stable than designed duplexes with I·U pairs next to no G-C pairs. There would probably be an inconsistency in this trend if self-assembly was a problem.

Thermodynamic Comparison to Duplexes Containing A-U and G-U pairs. ADAR enzymes selectively convert adenos-

Table 1: Thermodynamic Parameters for Duplex Formation

	T_M^{-1} vs log C_T plots				average of curve fits				predicted[e]			
Duplex[a]	-ΔH° (kcal/mol)	-ΔS° (eu)	-ΔG°$_{37}$ (kcal/mol)	T_M[b] (°C)	-ΔH° (kcal/mol)	-ΔS° (eu)	-ΔG°$_{37}$ (kcal/mol)	T_M[b] (°C)	-ΔH° (kcal/mol)	-ΔS° (eu)	-ΔG°$_{37}$ (kcal/mol)	T_M[b] (°C)
Internal												
GCAICGC CGUUGCG	71.5 ± 3.3	203.7 ± 10.3	8.34 ± 0.07	45.0	70.5 ± 9.9	200.4 ± 31.9	8.35 ± 0.14	45.2	70.5	199.2	8.66	46.9
GCAIGGC CGUUCCG	74.2 ± 4.3	209.5 ± 13.4	9.20 ± 0.14	48.6	74.6 ± 6.0	210.8 ± 18.8	9.24 ± 0.20	48.7	75.0	210.5	9.75	50.7
GCCIAGC CGGUUCG	70.0 ± 13.7	197.6 ± 42.8	8.69 ± 0.79	46.9	72.1 ± 11.0	204.2 ± 34.5	8.80 ± 0.44	47.1	70.2	199.2	8.43	45.6
GCCICGC CGGUGCG	73.7 ± 4.4	204.2 ± 13.6	10.40 ± 0.15	54.2	70.2 ± 11.4	193.1 ± 35.4	10.32 ± 0.39	54.7	73.7	204.9	10.17	53.0
GCCIGGC CGGUCCG	85.2 ± 4.9	235.4 ± 14.8	12.14 ± 0.31	58.9	83.7 ± 5.6	230.9 ± 16.8	12.09 ± 0.39	59.1	78.3	216.2	11.26	56.9
GCCIUGC CGGUACG	68.7 ± 15.0	195.2 ± 47.7	8.14 ± 0.79	44.5	69.4 ± 9.5	197.5 ± 31.0	8.19 ± 0.31	44.6	77.8	220.5	9.39	48.9
GCGIAGC CGCUUCG	68.8 ± 15.6	196.3 ± 49.5	7.91 ± 0.66	43.3	69.5 ± 9.4	198.4 ± 30.4	7.98 ± 0.25	43.6	65.3	184.4	8.10	44.7
GCGICGC CGCUGCG	72.8 ± 5.0	205.3 ± 15.8	9.08 ± 0.17	48.2	73.4 ± 16.3	207.6 ± 51.7	9.06 ± 0.33	48.0	68.8	190.1	9.84	52.7
GCGIGGC CGCUCCG	72.4 ± 3.1	197.0 ± 9.4	11.33 ± 0.16	59.0	73.5 ± 7.3	200.1 ± 22.3	11.44 ± 0.42	59.2	73.4	201.4	10.93	56.8
GCGIUGC CGCUACG	75.6 ± 3.5	213.9 ± 10.9	9.22 ± 0.11	48.5	75.4 ± 3.7	213.3 ± 11.5	9.25 ± 0.15	48.6	72.9	205.7	9.06	48.3
GCUIAGC CGAUUCG	64.6 ± 9.4	189.2 ± 30.6	5.93 ± 0.36	34.1	65.1 ± 10.0	190.5 ± 33.0	6.02 ± 0.41	34.6	65.4	191.2	6.11	35.0
GCUICGC CGAUGCG	76.1 ± 4.5	218.3 ± 14.2	8.40 ± 0.11	44.8	71.5 ± 6.4	203.6 ± 20.4	8.34 ± 0.12	45.1	68.9	196.9	7.85	43.0
GCUIGGC CGAUCCG	71.2 ± 10.0	200.6 ± 31.1	9.03 ± 0.46	48.3	70.8 ± 8.6	199.0 ± 26.2	9.05 ± 0.47	48.5	73.5	208.2	8.94	47.5
GCUIUGC CGAUACG	75.4 ± 6.3	221.9 ± 20.4	6.58 ± 0.10	37.2	74.9 ± 4.7	220.4 ± 15.6	6.59 ± 0.17	37.3	73.0	212.5	7.07	39.4
GUCUAGIC CIGAUCUG					61.0 ± 1.9[c]	176.6 ± 6.1[c]	6.3 ± 0.06[c]	40.1[c]	67.8	200.4	5.71	36.8
Terminal												
GCGCAI CGCGUU	58.8 ± 6.1	157.7 ± 18.8	9.84 ± 0.33	55.5	58.9 ± 8.9	158.1 ± 27.0	9.90 ± 0.55	55.8	59.0	161.2	8.96	50.5
GCGCCI CGCGGU	64.7 ± 4.9	175.3 ± 14.9	10.35 ± 0.28	56.4	63.6 ± 4.8	171.7 ± 14.4	10.31 ± 0.40	56.6	62.3	166.9	10.47	58.0
GCGCGI CGCGCU	61.6 ± 5.8	168.5 ± 17.8	9.39 ± 0.30	52.1	61.2 ± 5.0	167.2 ± 15.1	9.38 ± 0.36	52.2	57.3	152.1	10.14	57.9
GCGCUI CGCGAU	50.9 ± 3.5	137.6 ± 11.0	8.16 ± 0.10	47.3	48.6 ± 4.9	130.3 ± 15.4	8.19 ± 0.18	47.9	57.4	158.9	8.15	46.0
IAGCGC UUCGCG	53.1 ± 3.8	144.7 ± 12.0	8.21 ± 0.13	47.1	52.2 ± 10.6	141.6 ± 32.7	8.24 ± 0.48	47.5	55.6	153.0	8.09	46.1
ICGCGC UGCGCG	63.3 ± 4.4	171.2 ± 13.6	10.14 ± 0.23	55.8	62.4 ± 3.1	168.6 ± 9.0	10.09 ± 0.28	55.8	59.1	158.7	9.83	55.5
IGGCGC UCCGCG	64.8 ± 4.8	176.2 ± 14.8	10.20 ± 0.27	55.6	63.3 ± 4.7	171.4 ± 14.3	10.16 ± 0.35	55.8	63.6	170.0	10.92	59.9
IUGCGC UACGCG	67.1 ± 4.8	181.9 ± 14.8	10.64 ± 0.28	57.3	66.4 ± 7.0	179.5 ± 21.4	10.67 ± 0.40	57.6	63.1	174.3	9.05	50.1
ICCGGU UGGCCI					55.7 ± 1.7[c]	152.1 ± 5.2[c]	8.5 ± 0.09[c]	53.5[c]	57.1	155.0	9.08	56.3
IGAGGG UCUCCCGG					54.8 ± 1.1[c]	146.2 ± 3.6[c]	9.5 ± 0.02[c]	54.5[c]	58.8	158.4	9.61	54.5
Tandem												
GUCUIGAC CAGIUCUG					65.1 ± 0.6[c]	193.0 ± 1.6[c]	5.3 ± 0.18[c]	34.9[c]	62.4	185.6	4.87	33.0
GGCIUGCC CCGUICGG	59.9[d]	171.2[d]	6.8[d]	43.0[d]	59.1[d]	168.9[d]	6.7[d]	42.4[d]	59.9	171.2	6.80	43.0
GGCUIGCC CCGIUCGG	68.6[d]	193.6[d]	8.6[d]	50.7[d]	73.8[d]	211.4[d]	8.2[d]	47.9[d]	71.3	200.8	9.05	52.2
GGCIIGGC CCGUUCCG	72.9[d]	207.4[d]	8.6[d]	45.9[d]	79.4[d]	227.7[d]	8.8[d]	46.1[d]	68.8	207.4	8.60	27.9

[a] Solutions are 1 M NaCl, 20 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH = 7.0. The top strand of each duplex is written 5′ to 3′ and the bottom strand is written 3′ to 5′. [b] Calculated for 10^{-4} M oligonucleotide concentration. [c] Reference *17*. [d] Reference *16*. [e] Predicted values based on Watson-Crick (*18*) and inosine (Table 3) nearest neighbor parameters.

Table 2: Free Energy Comparison between Duplexes Containing I·U, A-U, and G-U Pairs

Duplex[a]	$-\Delta G^{\circ}_{37}$ (kcal/mol)[b]	-NN A-U (kcal/mol)[c]	ΔA-U (kcal/mol)[d]	-NN G-U (kcal/mol)[e]	ΔG-U (kcal/mol)[f]	Duplex[a]	$-\Delta G^{\circ}_{37}$ (kcal/mol)[b]	-NN A-U (kcal/mol)[c]	ΔA-U (kcal/mol)[d]	-NN G-U (kcal/mol)[e]	ΔG-U (kcal/mol)[f]
Internal						*Terminal*					
GCAICGC CGUUGCG	8.34	10.39	2.05	10.28	1.94	GCGCAI CGCGUU	9.84	7.70	-2.14	7.32	-2.52
GCAIGGC CGUUCCG	9.20	11.13	1.93	10.78	1.58	GCGCCI CGCGGU	10.35	10.03	-0.32	9.33	-1.02
GCCIAGC CGGUUCG	8.69	11.13	2.44	10.77	2.08	GCGCGI CGCGCU	9.39	9.37	-0.02	8.55	-0.84
GCCICGC CGGUGCG	10.40	12.72	2.32	12.29	1.89	GCGCUI CGCGAU	8.16	8.07	-0.09	7.74	-0.42
GCCIGGC CGGUCCG	12.14	13.46	1.32	12.79	0.65	IAGCGC UUCGCG	8.21	7.67	-0.54	8.01	-0.20
GCCIUGC CGGUACG	8.14	11.33	3.19	10.89	2.75	ICGCGC UGCGCG	10.14	9.26	-0.88	9.53	-0.61
GCGIAGC CGCUUCG	7.91	10.47	2.56	9.99	2.08	IGGCGC UCCGCG	10.20	10.00	-0.20	10.03	-0.17
GCGICGC CGCUGCG	9.08	12.06	2.98	11.51	2.43	IUGCGC UACGCG	10.64	7.87	-2.77	8.13	-2.51
GCGIGGC CGCUCCG	11.33	12.80	1.47	12.01	0.68	ICCGGU UGGCCI	8.5[g]	7.94	-0.56	8.48	-0.02
GCGIUGC CGCUACG	9.22	10.67	1.45	10.11	0.89	IGAGGG UCUCCCGG	9.5[g]	8.69	-0.81	8.72	-0.98
GCUIAGC CGAUUCG	5.93	9.17	3.24	9.18	3.25	Average			-0.83		-0.93
GCUICGC CGAUGCG	8.40	10.76	2.36	10.70	2.30	*Tandem* GUCUIGAC CAGIUCUG	5.3[g]	10.15	4.85	8.58	3.28
GCUIGGC CGAUCCG	9.03	11.50	2.47	11.20	2.17	GGCIUGCC CCGUICGG	6.8[h]	14.16	7.36	10.37	3.57
GCUIUGC CGAUACG	6.58	9.37	2.79	9.30	2.72	GGCUIGCC CCGIUCGG	8.6[h]	14.33	5.73	12.76	4.16
GUCUAGIC CIGAUCUG	6.3[g]	10.15	3.85	9.05	2.75	GGCIIGGC CCGUUCCG	8.6[h]	14.39	5.79	12.86	4.26
Average			2.43		2.01	Average			5.93		3.82

[a] For each duplex, the top strand is written 5′ to 3′ and the bottom strand is written 3′ to 5′. [b] The measured $-\Delta G^{\circ}_{37}$ of the listed duplex, using values from the log C_T plots, if available. [c] $-\Delta G^{\circ}_{37}$ calculated using the nearest neighbor model (*18*) for the duplex if the I·U pair was an A-U pair. [c] The difference in $-\Delta G^{\circ}_{37}$ between the −NN A-U value and the measured $-\Delta G^{\circ}_{37}$. [e] $-\Delta G^{\circ}_{37}$ calculated using the nearest neighbor model (*18*) for the duplex if the I·U pair was a G-U pair. [f] The difference in $-\Delta G^{\circ}_{37}$ between the −NN G-U value and the measured $-\Delta G^{\circ}_{37}$. [g] Reference *17*. [h] Reference *16*.

ine to inosine within mostly double-stranded RNA, which results in the conversion of A-U to I·U pairs. After this conversion, RNA duplexes become more sensitive to single-strand specific nucleases (*29*) suggesting that I·U pairs are less stable than A-U pairs. A previous study reported that the r(GUCUAGIC)$_2$ duplex is 1.8 kcal/mol less stable than the r(GUCUAGAC)$_2$ duplex (*17*). This sequence, however, was the only sequence studied that contained an internal I·U pair. In order to determine if this destabilization is general for sequences with internal I·U pairs, the stabilities of 14 RNA duplexes containing internal I·U pairs were measured, with each sequence containing a unique combination of I·U nearest neighbors.

As expected from ADAR nuclease studies (*29*) and from the value reported in Strobel et al. (*17*), all 14 duplexes containing an internal I·U pair are significantly less stable than predicted for the same duplex containing an internal A-U pair (Table 2). On average, a duplex containing an internal I·U pair is 2.4 kcal/mol less stable than predicted for the same duplex containing an internal A-U pair. An I·U pair can adopt a conformation with two hydrogen bonds and a slight shift of the uridine in the I·U in comparison to the uridine in an A-U pair (Figure 1). It is possible that the electrostatics of the inosine ring system result in hydrogen bonds (Figure 1) that are weaker than those in an A-U pair. Similarly, electrostatics of the inosine ring may result in a

Table 3: Nearest Neighbor Parameters for I·U Pairs

Nearest Neighbors[a]	Number of Occurrences[b]	ΔH° (kcal/mol)	ΔS° (eu)	ΔG°_{37} (kcal/mol)
IU UI	1	17.00 ± 7.11	43.3 ± 22.2	3.58 ± 1.09
II UU	1	9.53 ± 5.75	8.8 ± 17.9	2.66 ± 0.88
UI IU	2	8.41 ± 5.94	19.9 ± 18.5	2.23 ± 0.91
IA UU	4	-8.22 ± 2.87	-27.9 ± 8.9	0.43 ± 0.44
UI AU	5	-10.08 ± 2.55	-33.8 ± 7.9	0.37 ± 0.39
AI UU	3	-11.68 ± 3.05	-36.3 ± 9.5	-0.41 ± 0.47
IU UA	4	-15.83 ± 2.87	-49.4 ± 8.9	-0.50 ± 0.44
CI GU	8	-11.99 ± 2.55	-36.2 ± 8.0	-0.77 ± 0.39
IC UG	9	-11.56 ± 1.92	-34.0 ± 6.0	-1.03 ± 0.30
IG UC	11	-13.38 ± 2.40	-39.3 ± 7.5	-1.22 ± 0.37
GI CU	7	-9.81 ± 2.16	-27.4 ± 6.7	-1.34 ± 0.33
Terminal I·U	11	-0.08 ± 1.63	4.0 ± 5.1	-1.33 ± 0.25

[a] For each nearest neighbor pair, the top sequence is written 5′ to 3′ and the bottom sequence is written 3′ to 5′. [b] The number of times that nearest neighbor pair appears in the sequences studied.

ring system that results in less favorable stacking interactions. The slight shift in conformation in going from an A-U pair to an I·U pair may reduce favorable stacking interactions, may situate the carbonyl oxygen above or below another partial negative charge, or may cause an unfavorable helix distortion. Differences in hydration may also play a role. As shown in Figure 1, different functional groups in the major and minor grooves of an A-U pair and an I·U pair are available for interaction with water molecules. It is also possible that the I·U pair is in a different conformation than that shown in Figure 1 or in multiple conformations, resulting in fewer hydrogen bonds and weaker stacking interactions.

Since I·U pairs can form a conformation isosteric with the wobble G-U conformation (Figure 1), the stabilities of duplexes containing internal I·U pairs were compared to nearest neighbor calculations for the same duplexes containing G-U pairs (Table 2). The results show, on average, that duplexes containing an I·U pair in the middle of a Watson−Crick duplex are 2.0 kcal/mol less stable than the same duplex containing an internal G-U pair. Due to the fact that I·U pairs can form a conformation isosteric with the wobble G-U pair (Figure 1), it was surprising to see such a large difference in stability between the two types of duplexes. These results showing that an internal I·U pair is less stable than an internal G-U pair are inconsistent with pK_a data. For example, the pK_a of guanosine is 9.2, but the pK_a of inosine is 8.8 (*13*). This suggests that the imino proton of guanosine is a weaker hydrogen bond donor than the imino proton of inosine. It is possible, however, that the exocyclic amino group of guanosine improves stacking interactions of the G-U pair. Additionally, as suggested previously by Serra et al. (*16*), the I·U pair may experience an increase in base pair opening in comparison to a G-U pair due to the absence of the 2-amino group in the minor groove leading to a decrease in the stability of water networks around the I·U base pair. It is also possible that the I·U pair is not in the conformation in Figure 1 or is in a variety of conformations.

Similar comparisons can be made between duplexes with terminal I·U pairs and the same duplexes containing terminal A-U and terminal G-U pairs. Interestingly, however, the results are quite different from internal pairs. Duplexes containing internal I·U pairs are approximately 2.2 kcal/mol *less* stable than the same duplexes containing A-U or G-U pairs, however, duplexes containing terminal I·U pairs are approximately 0.88 kcal/mol *more* stable than the same duplexes containing A-U or G-U pairs (Table 2). More specifically, duplexes containing terminal I·U pairs are 0.8 and 0.9 kcal/mol more stable than the same duplexes containing A-U and G-U pairs, respectively. A terminal I·U pair may have more conformational flexibility than a terminal A-U or terminal G-U pair and provide a favorable entropic term. Also, conformational flexibility may allow the terminal I·U pair to overcome destabilizing factors associated with an internal I·U pair and allow the terminal I·U pair to stabilize a duplex more than a terminal A-U or terminal G-U pair. It is interesting to note that the stability of a terminal I·U pair may be utilized in nature; it may play a role in the interaction between the tRNA anticodon loop and the mRNA codon where a terminal inosine pair can form in the wobble position.

Although the thermodynamics were not measured as part of this study, it is important to mention the calculations done to compare duplexes containing tandem I·U pairs to the same duplexes containing tandem A-U and tandem G-U pairs. On average, duplexes containing tandem I·U pairs are 5.9 and 3.8 kcal/mol less stable than the same duplexes containing tandem A-U and tandem G-U pairs, respectively. The destabilizing factors discussed above for single, internal I·U pairs are likely compounded by placing two I·U pairs adjacent to each other, resulting in a significantly less stable duplex in comparison to duplexes containing tandem A-U and tandem G-U pairs.

Thermodynamic Contributions of an I·U Pair to the Stability of a Duplex. As expected, summation of the I·U nearest neighbor contributions for duplexes containing a nontandem, internal I·U pair stabilized 13 out of 15 duplexes (Table S3). The two duplexes that were not stabilized by the summation of the I·U nearest neighbor contributions contained an A-U pair on both sides of the I·U pair.

The nearest neighbor combination, $\binom{\text{G I C}}{\text{CUG}}$, was studied by Strobel et al. (*17*) and in this study. The summation of the I·U nearest neighbor free energy contribution (Table S3) for the sequence studied here is −1.6 kcal/mol, while the summation of the I·U nearest neighbor free energy contributions per I·U pair for the sequence studied by Strobel et al.

(*17*) is −2.7 kcal/mol. There are three differences between the duplex studied here and the duplex studied previously that may help explain this 1.1 kcal/mol difference. First, non-nearest neighbors may be affecting the stability. Second, in this study, the I·U pair is located in the center of the duplex while in the Strobel et al. duplex (*17*), the I·U pair is adjacent to the terminal pair. A terminal base pair adjacent to an I·U pair may have more conformational flexibility to accommodate the destabilizing factors contributed by the I·U pair in comparison to Watson−Crick pairs located in the center of a Watson−Crick duplex. Third, the Strobel et al. duplex (*17*) contains two $\binom{\text{G I C}}{\text{CUG}}$ stretches separated by two A-U pairs. This duplex may be more flexible and "breathe" due to the lack of G-C stretches and may be better able to accommodate an I·U pair in comparison to the duplex studied here which contains only one $\binom{\text{G I C}}{\text{CUG}}$ stretch in the center of an otherwise all G-C duplex.

Another nearest neighbor combination was studied by both Strobel et al. (*17*) and Serra et al. (*16*), $\binom{\text{CU I G}}{\text{C I UC}}$. The summation of the I·U nearest neighbor free energy contribution (Table S1) for the sequence studied by Strobel et al. (*17*) is −0.6 kcal/mol, while the summation of the I·U nearest neighbor free energy contribution for the sequence studied by Serra et al. (*16*) is 0.2 kcal/mol. In both of these duplexes, the I·U pairs are placed in the center of the duplex and the duplexes are of the same length. The only difference between the two duplexes is the base pair two pairs away from the I·U pair. Since the duplexes are self-complementary, there are a total of two base pairs different between the two strands. Clearly, the difference in the summation of the I·U nearest neighbor free energy contribution is a result of non-nearest neighbor effects.

I·U Nearest Neighbor Parameters. Nearest neighbor parameters for I·U/Watson−Crick neighbors, tandem I·U neighbors, and terminal I·U pairs are shown in Table 3. Similar to the nearest neighbor parameters for Watson−Crick/Watson−Crick (*18*) and G-U/Watson−Crick nearest neighbors (*30*), all combinations of I·U/Watson−Crick nearest neighbors contribute a negative term to the total $\Delta H°$ and $\Delta S°$. Unlike all Watson−Crick/Watson−Crick and G-U/Watson−Crick nearest neighbor combinations, not all possible I·U/Watson−Crick nearest neighbor combinations contribute a negative (favorable) term to total $\Delta G°_{37}$. Both $\binom{\text{I A}}{\text{U U}}$ and $\binom{\text{U I}}{\text{A U}}$ nearest neighbors contribute a positive (unfavorable) term to total $\Delta G°_{37}$. Similar to terminal A-U and terminal G-U pairs, terminal I·U pairs contribute an additional positive term to the total $\Delta S°$ value. However, unlike terminal A-U and terminal G-U pairs, terminal I·U pairs contribute a negative term to total $\Delta H°$ and $\Delta G°_{37}$ values. Tandem I·U pairs contribute a positive term to total $\Delta G°_{37}$ and a positive term to total $\Delta H°$, and two out of three tandem I·U pairs contribute a positive term to total $\Delta S°$.

The neighboring Watson−Crick pairs have a large influence over the stability of I·U pairs. The stability trend for the base pair 5′ of the I·U pair is $\binom{\text{G I}}{\text{CU}} > \binom{\text{C I}}{\text{GU}} > \binom{\text{A I}}{\text{UU}} > \binom{\text{U I}}{\text{AU}}$. The extra hydrogen bond in a G-C pair versus an A-U pair may help explain why $\binom{\text{C I}}{\text{CU}}$ and $\binom{\text{C I}}{\text{GU}}$ are more stable than $\binom{\text{A I}}{\text{UU}}$ and $\binom{\text{U I}}{\text{AU}}$. In order to explain the difference between $\binom{\text{G I}}{\text{CU}}$ versus $\binom{\text{C I}}{\text{GU}}$ and $\binom{\text{A I}}{\text{UU}}$ versus $\binom{\text{U I}}{\text{AU}}$, differences in hydration, stacking, and helix distortion must be considered. The trend observed here is similar to that for the base pair 5′ of a dI·T pair ($\binom{\text{G I}}{\text{C T}} \approx \binom{\text{C I}}{\text{G T}} > \binom{\text{A I}}{\text{T T}} > \binom{\text{T I}}{\text{A T}}$) (*12*). The order of the last two pairs is switched for the base pair 5′ of an A-U pair ($\binom{\text{GA}}{\text{CU}} > \binom{\text{CA}}{\text{GU}} > \binom{\text{UA}}{\text{AU}} > \binom{\text{AA}}{\text{UU}}$) (*18*) and for the base pair 5′ of a G-U pair ($\binom{\text{GG}}{\text{CU}} > \binom{\text{CG}}{\text{GU}} > \binom{\text{UG}}{\text{AU}} > \binom{\text{AG}}{\text{UU}}$) (*30*). It is unclear what causes the last two pairs to switch positions when comparing inosine and deoxyinosine to canonical pairs.

The stability trend for the base pair 3′ of the I·U pair is $\binom{\text{I G}}{\text{UC}} > \binom{\text{I C}}{\text{UG}} > \binom{\text{I U}}{\text{UA}} > \binom{\text{I A}}{\text{UU}}$. The extra hydrogen bond in a G-C pair versus an A-U pair may help explain why $\binom{\text{I G}}{\text{UC}}$ and $\binom{\text{I C}}{\text{UG}}$ are more stable than $\binom{\text{I U}}{\text{UA}}$ and $\binom{\text{I A}}{\text{UU}}$. In order to explain the difference between $\binom{\text{I G}}{\text{UC}}$ and $\binom{\text{I C}}{\text{UC}}$ and the difference between $\binom{\text{I U}}{\text{UA}}$ and $\binom{\text{I A}}{\text{UU}}$, differences in hydration, stacking, and helix distortion must be considered. The order of the first two pairs in this trend is switched for the trend observed for the base pair 3′ of an A-U pair ($\binom{\text{AC}}{\text{UG}} > \binom{\text{AG}}{\text{UC}} > \binom{\text{AU}}{\text{UA}} > \binom{\text{AA}}{\text{UU}}$) (*18*) and for the base 3′ of a G-U pair ($\binom{\text{GC}}{\text{UG}} > \binom{\text{GG}}{\text{UC}} > \binom{\text{GU}}{\text{UA}} \approx \binom{\text{GA}}{\text{UU}}$) (*30*). It is unclear why a switch in the trends is observed. The trend seen for the base 3′ of a dI-T pair is quite different ($\binom{\text{IC}}{\text{TG}} > \binom{\text{IT}}{\text{TA}} > \binom{\text{IG}}{\text{TC}} > \binom{\text{IA}}{\text{TT}}$) (*12*). Different helical parameters for DNA helices versus RNA helices may account for some of these differences. These data provide benchmarks for testing computational chemistry approaches to predicting the sequence dependence of helical stability.

Biological Implications. The stability of an RNA duplex containing an internal I·U can be calculated using the new I·U nearest neighbor model. For example, the stability of $\binom{\text{GCA I CGC}}{\text{CGUUGCG}}$ would be calculated as follows using previously determined Watson−Crick nearest neighbors (*18*) and the new inosine nearest neighbors in Table 3:

$$\Delta G°_{37}\binom{\text{GCA I CGC}}{\text{CGUUGCG}} = \Delta G°_{37,\text{i}} + \Delta G°_{37}\binom{\text{GC}}{\text{CG}} + \Delta G°_{37}\binom{\text{CA}}{\text{GU}} + \Delta G°_{37}\binom{\text{A I}}{\text{UU}} + \Delta G°_{37}\binom{\text{I C}}{\text{UG}} + \Delta G°_{37}\binom{\text{CG}}{\text{GC}} + \Delta G°_{37}\binom{\text{GC}}{\text{CG}} \quad (9)$$

$$\Delta G°_{37}\binom{\text{GCA I CGC}}{\text{CGUUGCG}} = 4.09 + (-3.42) + (-2.11) + (-0.41) + (-1.03) + (-2.36) + (-3.42) = -8.66 \text{ kcal/mol} \quad (10)$$

Similarly, the stability of $\binom{\text{GCGCA I}}{\text{CGCGUU}}$ would be calculated as follows:

$$\Delta G°_{37}\binom{\text{GCGCA I}}{\text{CGCGUU}} = \Delta G°_{37,\text{i}} + \Delta G°_{37}\binom{\text{GC}}{\text{CG}} + \Delta G°_{37}\binom{\text{CG}}{\text{GC}} + \Delta G°_{37}\binom{\text{GC}}{\text{CG}} + \Delta G°_{37}\binom{\text{CA}}{\text{GU}} + \Delta G°_{37}\binom{\text{A I}}{\text{UU}} + \Delta G°_{37,\text{terminal IU}} \quad (11)$$

$$\Delta G°_{37}\binom{\text{GCGCA I}}{\text{CGCGUU}} = 4.09 + (-3.42) + (-2.36) + (-3.42) + (-2.11) + (-0.41) + (-1.33) = -8.96 \text{ kcal/mol} \quad (12)$$

The I·U nearest neighbor parameters derived here should allow for the comparison of the stability of RNA structures before and after adenosine deamination. For example, using nearest neighbor parameters, the $\binom{\text{CAG}}{\text{GUC}}$ at the Q/R site of GluR-B (*31*) is predicted to contribute −5.53 kcal/mol toward

duplex stability. When the adenosine is converted to inosine resulting in $\binom{C\ IG}{GUC}$, these nearest neighbors contribute only −1.99 kcal/mol toward duplex stability, resulting in a destabilization of 3.5 kcal/mol upon adenosine deamination to inosine. Also, when an adenosine upstream of the Q/R site in GluR-6 (*31*) is converted to inosine, it is predicted that the duplex becomes 1.9 kcal/mol less stable. Similarly, upon adenosine to inosine conversion at the B, C, and D site of the 2C subtype of serotonin receptor (*31*), it is predicted that the duplex would become 2.9, 1.1, and 2.3 kcal/mol less stable, respectively. Along with a destabilization, conversion of adenosine to inosine at these sites may also change the local dynamics and structure resulting in changes in biomolecular recognition.

From the results of this study, it is obvious that occurrences of inosine in nature cannot be simply approximated as guanosine or adenosine. More studies are required to fully characterize the thermodynamics, structure, and dynamics of RNA containing inosine.

ACKNOWLEDGMENT

The authors would like to thank Doug Turner for insightful discussions and the Turner Lab for the use of their RNA synthesizer. The authors would also like to thank Dave Mathews for his help with linear regression. Partial funding for this project was provided by the St. Louis University College of Arts and Sciences, St. Louis University Department of Chemistry, a St. Louis University Summer Research Award (BMZ), and a Research Corporation Cottrell College Science Award CC6547 (BMZ).

SUPPORTING INFORMATION AVAILABLE

Tables showing the enthalpy comparison between duplexes containing I•U, A-U, and G-U pairs; the entropy comparison between duplexes containing I•U, A-U, and G-U pairs; and the I•U contribution to thermodynamics. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Bass, B. L. (2002) RNA editing by adenosine deaminases that act on RNA, *Annu. Rev. Biochem. 71*, 817−846.
2. Rebagliati, M. R., and Melton, D. A. (1987) Antisense RNA injections in fertilized frog eggs reveal an RNA duplex unwinding activity, *Cell 48*, 599−605.
3. Kim, U., Wang, Y., Sanford, T., Zeng, Y., and Nishikura, K. (1994) Molecular-cloning of cDNA for double-stranded-RNA adenosine-deaminase, a candidate enzyme for nuclear-RNA editing, *Proc. Natl. Acad. Sci. U.S.A. 91*, 11457−11461.
4. Herbert, A., Lowenhaupt, K., Spitzner, J., and Rich, A. (1995) Chicken double-stranded-RNA adenosine-deaminase has apparent specificity for Z-DNA, *Proc. Natl. Acad. Sci. U.S.A. 92*, 7550−7554.
5. Hough, R. F., and Bass, B. L. (1997) Analysis of Xenopus dsRNA adenosine deaminase cDNAs reveals similarities to DNA methyltransferases, *RNA 3*, 356−370.
6. Slavov, D., Crnogorac-Jurcevic, T., Clark, M., and Gardiner, K. (2000) Comparative analysis of the DRADA A-to-I RNA editing gene from mammals, pufferfish and zebrafish, *Gene 250*, 53−60.
7. Palladino, M. J., Keegan, L. P., O'Connell, M. A., and Reenan, R. A. (2000) dADAR, a Drosophila double-stranded RNA-specific adenosine deaminase is highly developmentally regulated and is itself a target for RNA editing, *RNA 6*, 1004−1018.
8. Tonkin, L. A., Saccomanno, L., Morse, D. P., Brodigan, T., Krause, M., and Bass, B. L. (2002) RNA editing by ADARs is important for normal behavior in Caenorhabditis elegans, *EMBO J. 21*, 6025−6035.
9. Paul, M. S., and Bass, B. L. (1998) Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA, *EMBO J. 17*, 1120−1127.
10. Knight, S. W., and Bass, B. L. (2002) The role of RNA editing by ADARs in RNAi, *Mol. Cell 10*, 809−817.
11. Martin, F. H., Castro, M. M., Aboulela, F., & Tinoco, I. (1985) Base-pairing involving deoxyinosine-Implications for probe design, *Nucleic Acids Res. 13*, 8927−8938.
12. Watkins, N. E., and SantaLucia, J. Jr. (2005) Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes, *Nucleic Acids Res. 33*, 6258−6267.
13. SantaLucia, J., Kierzek, R., and Turner, D. H. (1991) Functional-group substitutions as probes of hydrogen-bonding between GA mismatches in RNA internal loops, *J. Am. Chem. Soc. 113*, 4313−4322.
14. Broda, M., Kierzek, E., Gdaniec, Z., Kulinski, T., and Kierzek, R. (2005) Thermodynamic stability of RNA structures formed by CNG trinucleotide repeats. Implication for prediction of RNA structure, *Biochemistry 44*, 10873−10882.
15. Turner, D. H., Sugimoto, N., Kierzek, R., and Dreiker, S. D. (1987) Free energy increments for hydrogen bonds in nucleic acid base pairs, *J. Am. Chem. Soc. 109*, 3783−3785.
16. Serra, M. J., Smolter, P. E., and Westhof, E. (2004) Pronounced instability of tandem IU base pairs in RNA, *Nucleic Acids Res. 32*, 1824−1828.
17. Strobel, S. A., Cech, T. R., Usman, N., and Beigelman, L. (1994) The 2,6-diaminopurine riboside-5-methylisocytidine wobble base pair: An isoenergetic substitution for the study of GU pairs in RNA, *Biochemistry 33*, 13824−13835.
18. Xia, T., SantaLucia, J. Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs, *Biochemistry 37*, 14719−14735.
19. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. J. (1987) Automated chemical synthesis of long oligoribonucleotides using 2′-O-silylated ribonucleoside 3′-O-phosphoramidites on a controlled-pore glass support-Synthesis of a 43-nucleotide sequence similar to the 3′-half molecule of an Escherichia-coli formylmethionine transfer-RNA, *J. Am. Chem. Soc. 109*, 7845−7854.
20. Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., and Usman, N. (1995) Synthesis, deprotection, analysis and purification of RNA and ribozymes, *Nucleic Acids Res. 23*, 2677−2684.
21. Stawinski, J., Stromberg, R., Thelin, M., and Westman, E. (1988) Evaluation of the use of the tert-butyldimethylsilyl group for 2′-protection in RNA-synthesis via the H-phosphonate approach, *Nucleosides Nucleotides 7*, 779−782.
22. Chou, S. H., Flynn, P., and Reid, B. (1989) Solid-phase synthesis and high-resolution NMR-studies of two synthetic double-helical RNA dodecamers-r(CGCGAAUUCGCG) and r(CGCGUAUACGCG), *Biochemistry 28*, 2422−2435.
23. Devor, E. J., and Behlke, M. A. (2005). Integrated DNA Technologies, http://www.idtdna.com/support/technical/TechnicalBulletinPDF/Oligonucleotide_Yield_Resuspension_and_Storage.pdf (accessed Oct 6, 2004).
24. Sigma-Proligo, http://www.proligo.com/pro_primprobes/PP_08−8._UVAbsorbance.html (accessed Oct 4, 2004).
25. McDowell, J. A. (1996) MeltWin v. 3.5: Melt Curve Processing Program.
26. Petersheim, M., and Turner, D. H. (1983) Base-stacking and base-pairing contributions to helix stability: Thermodynamics of double-helix formation with CCGG, CCGGp, mCCGGAp, ACCGGp, CCGGUp, and ACCGGUp, *Biochemistry 22*, 256−263.
27. 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.
28. Borer, P. N., Dengler, B., Tinoco, I., and Uhlenbeck, O. (1974) Stability of ribonucleic-acid double-stranded helices, *J. Mol. Biol. 86*, 843−853.

29. Bass, B. L., and Weintraub, H. (1988) An unwinding activity that covalently modifies its double-stranded-RNA substrate, *Cell 55*, 1089−1098.
30. 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.
31. Rueter, S. M., and Emeson, R. B. (1998) in *Modification and editing of RNA* (Grosjean, H., and Benne, R., Eds.) pp 343−361, ASM Press, Washington, D.C.

BI0616910