1118–1125 *Nucleic Acids Research, 1999, Vol. 27, No. 4*

Thermodynamics of RNA hairpins containing single internal mismatches

May Meroueh and Christine S. Chow*

Department of Chemistry, Wayne State University, Detroit, MI 48202, USA

Received September 24, 1998; Revised and Accepted December 21, 1998

ABSTRACT

Thermodynamic parameters and circular dichroism spectra are presented for RNA hairpins containing single internal mismatches in the stem regions. Three different sequence contexts for the G•U mismatch and two contexts for C•A, G•A, U•U, A•C and U•G mismatches were examined and compared with Watson–Crick base-pair stabilities. The RNA hairpins employed were a microhelix and tetraloop representing the *Escherichia coli* tRNA^{Ala} acceptor stem and sequence variants that have been altered at the naturally occurring G•U mismatch site. UV melting studies were carried out under different conditions to evaluate the effects of sodium ion concentration and pH on the stability of mismatch-containing hairpins. Our main findings are that single internal mismatches exhibit a range of effects on hairpin stability. In these studies, the size and sequence of the loop and stem are shown to influence the overall stability of the RNA, and have a minor effect on the relative mismatch stabilities. The relationship of these results to RNA–ligand interactions involving mismatch base-pairs is discussed.

INTRODUCTION

There is an increasing amount of data to demonstrate that macromolecules, such as proteins and RNA, and small molecules take advantage of structural polymorphisms in RNA to locate and interact with their desired binding sites. Structural motifs that are currently of great interest involve mismatch base-pairs. The G•U wobble pair is one example that was first suggested by Crick in 1966 (1) and has since been observed in many RNAs. Although G•U is the most frequent non-canonical base-pair found in RNA to date (2), other mismatch pairs might also have important biological roles.

Single internal base-pairs are likely to have distinct orientations relative to the adjacent base-pairs, as observed with G•U pairs (3–7). In the case of G•U pairs, their unique stacking arrangements likely influence helix stability (8–10) as well as the ability to be recognized by a vast array of molecules, including proteins (11–16), RNA (17–19), small inorganic complexes (20,21), organic molecules (22,23) and metal ions (24–27). Metal complex recognition also occurs with tandem G•U mismatches (28–30), and G•Us can contribute to ribozyme reactions involving metal ions (31). The G•U wobble pair displays a relatively simple array of hydrogen-bonding groups in the major and minor grooves. Therefore, its recognition by such a large number of different molecules is quite striking considering the level of fidelity and precision that is necessary for proper biochemical functioning of the RNAs. Our goal is to understand how G•U stability in certain sequence and structural contexts can influence ligand recognition. In addition, comparisons of different base mismatch stabilities might also contribute to our understanding of the biological roles of G•U, as well as other base-pair mismatches.

Currently, there is a paucity of information available on the thermodynamic effects of single internal mismatches on duplex or hairpin RNA stability. Early studies demonstrated that G•U stabilities are dependent on sequence context, and stability is greater when the mismatch pair is flanked by G-C base-pairs relative to A-U pairs (8,10,32). Terminal mismatches (9,33) and tandem mismatches (34–40) have been examined in detail, and there have been a few measurements for single G•A (41) and A•A (42) mismatches. More recently, Zhu and Wartell reported that base identity and sequence contexts are important factors in determining mismatch stabilities (43). Their studies employed long RNA helices with single internal mismatches, and stabilities were measured by temperature gradient gel electrophoresis (TGGE). Studies of group II intron ribozyme function also revealed that mismatches can have destabilizing effects on the RNA helix that are both sequence and identity dependent (44).

Several groups have used mismatch substitutions to understand more clearly the role of G•U mismatches in protein recognition or RNA function (15,44–49). For example, *in vivo* selection and sequence analysis revealed the ability of functional 16S rRNA mutants to form a number of potential mismatch base-pairs (e.g., A•C, G•U, A•A and G•G) (46). Similarly, Gabriel *et al.* (45) have shown that the functionally critical G•U site, the so-called 'major determinant', in *Escherichia coli* tRNA^{Ala} can be replaced by certain base mismatches with only minor diminished aminoacylation activity. In this study, we examined RNA stabilities and structures when naturally occurring G•U mismatches were replaced with other pairs, namely C•A, G•A, U•U, A•C, U•G, G-C and A-U. We chose RNA hairpins that are known substrates of alanyl tRNA synthetase (AlaRS) in order to make comparisons between stability and function.

MATERIALS AND METHODS

Preparation of RNAs

The following RNAs, based on the acceptor stem of *E.coli* tRNA^{Ala} and sequence variants, were synthesized chemically on a

*To whom correspondence should be addressed. Tel: +1 313 577 2594; Fax: +1 313 577 8822; Email: csc@chem.wayne.edu

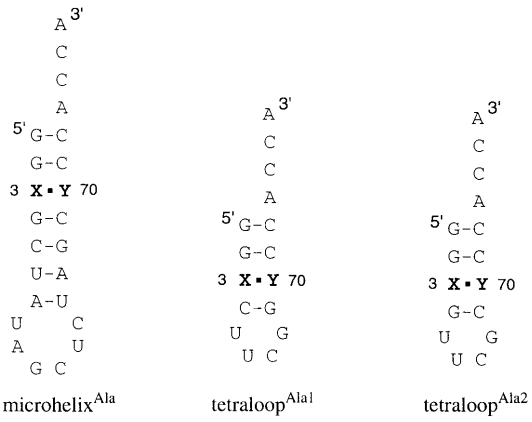

Figure 1. Secondary structure representations of the synthetic microhelices and tetraloops based on the acceptor helix of *E.coli* tRNA^{Ala} (56,57), in which X•Y is G•U, C•A, G•A, U•U, A•C, U•G, G-C or A-U, are shown.

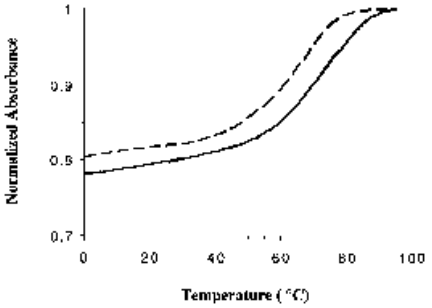

Figure 2. Representative normalized UV melting curves for the G•U microhelix taken at two different NaCl concentrations in 20 mM sodium cacodylate, 0.5 mM EDTA, pH 7.0 are shown (curves from left to right are: dashed line, 0.1 M NaCl, 47 µM RNA; solid line, 1 M NaCl, 75 µM RNA). The melting curves were normalized at 95°C.

Cruachem PS250 DNA/RNA synthesizer: 5′-GGXGCUAUAG-CUCUAGCYCCACCA-3′ (microhelix^{Ala}), 5′-GGXCUUCG-GYCCACCA-3′ (tetraloop^{Ala1}) and 5′-GGXGUUCGCYC-CACCA-3′ (tetraloop^{Ala2}), where the X•Y pairs are G•U, C•A, G•A, U•U, A•C, A-U, G-C or U•G. Either 2′-*O*-Fpmp (50) or 2′-*O*-silyl (51) protected cyanoethyl phosphoramidites from Cruachem were employed. The RNAs were deprotected according to literature procedures (50,51). The oligoribonucleotides were desalted on Sep-Pak C-18 cartridges (Waters) and purified on denaturing polyacrylamide gels (15%, 8 M urea, 42 cm long × 0.8 mm thick). The RNAs were electroeluted from the gel with 1× TBE (90 mM Tris, 90 mM boric acid, 2.5 mM Na_2EDTA, pH 8.3), desalted with Centricon 3′s (Amicon) and dried under vacuum. All of the RNAs were >95% pure as determined by ^{32}P labeling, followed by analysis on high-resolution, 20% denaturing polyacrylamide gels and phosphorimaging.

Melting studies

The melting curves (absorbance versus temperature profiles) were obtained using an Aviv 14DS UV-vis spectrophotometer with a five-cuvette thermoelectric controller as described previously (52). More specifically, three microcuvettes with a 0.1 cm path length and two with a 0.2 cm path length (60 and 120 µl volumes, respectively) were used; thus, five different concentrations over a 100-fold range were examined for each RNA, and each measurement was taken in duplicate or triplicate. The sample compartments were first purged with nitrogen to avoid water condensation. The RNAs (dissolved in 15 mM, 100 mM or 1 M NaCl, 20 mM sodium cacodylate and 0.5 mM Na_2EDTA, pH 7.0 or 5.0) were annealed by raising the temperature to 95°C, followed by cooling to –1.6°C. When the temperature was at 95°C, the absorbances of each sample were measured at 260 nm and used to determine the RNA concentrations. Single-strand extinction coefficients were calculated as described by Richards (53). The absorbances were measured at 260 or 280 nm from 0 to 95°C with a heating rate of 0.8°C/min. The thermodynamic parameters were obtained from the absorbance versus temperature profiles using a Van't Hoff analysis, assuming a two-state model, as described previously (34,54).

Circular dichroism spectroscopy

Circular dichroism (CD) spectra were measured on a Jasco J600 spectropolarimeter from 220 to 330 nm at ambient temperature. The buffer used was 15 mM NaCl, 20 mM sodium cacodylate and 0.5 mM Na_2EDTA, pH 7.0 or 5.0. Taking the RNA strand concentration into consideration, the measured CD spectra were converted to molar ellipticity ($\Delta\varepsilon$) as described by Cantor and Schimmel (55), except that values were expressed in moles of RNA molecules rather than moles of individual residues.

RESULTS AND DISCUSSION

Oligoribonucleotide sequence choice

Figure 1 shows the RNA sequences studied; the X•Y positions were substituted with G•U, U•G, A•C, C•A, G•A, U•U, A-U and G-C pairs. These hairpins mimic the acceptor stem of tRNA^{Ala} with mismatches at positions 3 and 70 (the numbering system is based on full-length tRNA). These RNAs were selected based on the aminoacylation studies of Francklyn and Schimmel (56) and Shi *et al.* (57) in which they were shown to be substrates of AlaRS. The 5′-C(UUCG)G-3′ sequence (tetraloop) is a motif found in rRNAs that provides RNA hairpins with increased thermodynamic stability (58). Gabriel *et al.* (45) employed the same eight variations at position 3•70 of the stem sequence for *in vivo* aminoacylation studies with full-length tRNAs. More recently, Beuning and co-workers (49) reported *in vitro* aminoacylation data and kinetics parameters for all possible 16 base-pairs in duplex RNAs and G•U, G•A, A•C, C•A and G-C base-pairs at nucleotide positions 3•70 in full-length tRNAs.

Thermodynamic parameters for microhelices at pH 7.0

For the eight microhelix^{Ala} RNAs, absorbance versus temperature profiles were obtained at pH 7.0 and analyzed in terms of the melting temperature (T_m), $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$ and $\Delta G°_{60}$ (59). The melting transitions were all monophasic and observed in the range of 55–80°C. Initially, these studies were performed separately in both 0.1 and 1 M NaCl. The normalized plots at single RNA concentrations are shown (Fig. 2) for the G•U microhelix, which formed a hairpin (concentration-independent melting profiles) at 0.1 and 1 M NaCl. For this sequence variant, the RNA was stabilized

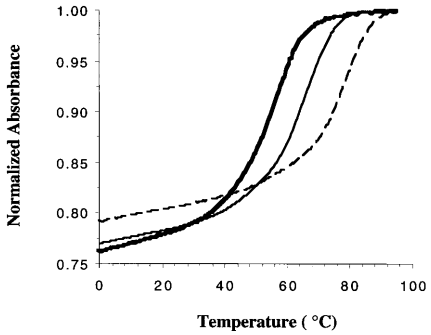

Figure 3. Representative normalized UV melting curves for the U•U, G•U and G-C microhelices obtained in 15 mM NaCl, 20 mM sodium cacodylate, 0.5 mM EDTA, pH 7.0 (from left to right: bold solid line, U•U, 369 µM RNA; solid line, G•U, 211 µM RNA; dashed line, G-C, 449 µM RNA) are shown. The melting curves were normalized at 95°C.

at the higher salt concentration (the T_ms were 71.0 and 76.5°C at 0.1 and 1 M NaCl, respectively). However, under these conditions (0.1 and 1 M NaCl), several mismatch RNAs (specifically, U•U, G•A and G-C variants) formed dimers (duplexes) as indicated by concentration-dependent melting profiles (data not shown). The fact that these RNAs formed duplexes and the G•U variant formed a hairpin under identical buffer conditions makes a comparison of the thermodynamic parameters difficult. Therefore, all further studies were carried out under lower salt conditions.

At low salt concentrations (15 mM NaCl), the helix-to-coil transitions for all RNAs examined were independent of RNA concentration over a range of ~5–500 µM, indicating hairpin rather than duplex formation. Figure 3 shows typical normalized plots from the melting curves of the U•U, G•U and G-C microhelices at 15 mM NaCl. The thermodynamic parameters for the microhelix variants at pH 7.0 are listed in Table 1. The G•U, C•A and A•C variants were each synthesized two or three times and independent melting curves were obtained. The thermodynamic parameters were within the assumed error limits (3% for ΔG°_{60}, 5% for ΔG°_{37}, 7% for ΔH° and 8% for ΔS°; values not shown) (54,59). As expected, the G-C and A-U variants were more stable than the mismatch RNAs. Of the mismatch-containing RNAs, the U•G, G•U and G•A variants demonstrated intermediate stability and A•C, C•A and U•U were the least stable.

Overall, the observed order of stability of pairs at pH 7.0 is G-C > A-U > U•G > G•U > G•A > A•C > C•A > U•U (in order of decreasing T_m) for the microhelices. The trend for ΔG°_{60} is identical, whereas minor deviations are seen with the ΔG°_{37} trends. ΔG° values close to the melting temperature (T_m) are probably more reliable than ΔG°_{37} (59). The ΔG°_{60} values for the A•C and C•A variants are within 0.2 kcal/mol, suggesting that the nearest-neighbor sequences have a minimal influence on stability of these mismatch pairs. The ΔG°_{60} values for the G•U and U•G variants differ by ~0.4 kcal/mol, indicating that these base mismatches are influenced to a small extent by the neighboring base-pairs, consistent with previous studies (10,32). Due to a lack of data for nearest-neighbor parameters for single mismatches in RNA other than G•U (60), the thermodynamic parameters for the microhelix hairpins were not well predicted by the MFOLD program (61), even with an appropriate salt correction (62). In the MFOLD program, all mismatches other than G•U are assumed to destabilize the helix by the same energy value (0.8 kcal/mol). From this study, it is apparent that single mismatches have different effects on helix stability depending on the sequence contexts, and individual parameters need to be derived for each mismatch as done with DNA (54,63,64). The single mismatches also behave differently than tandem mismatches or terminal mismatches. The relative trends for single mismatch stabilities in the microhelix RNAs follow those reported by Zhu and Wartell (43), in which G•U or U•G pairs are more stable than other single mismatches, and mismatches containing a purine base are more stable than those with two pyrimidine bases. Those authors also found that base identity, nearest neighbors and next-nearest neighbors can influence base-pair stability even though they examined single mismatches in different sequence contexts than reported here.

Table 1. Thermodynamics of microhelix^{Ala} RNAs, 5′-GGXGCUAUAGCUCUAGCYCCACCA-3′, with X•Y mismatches or Watson–Crick base-pairs

X•Y pair	ΔG°_{60} (kcal/mol)[a]		ΔG°_{37} (kcal/mol)[a]		ΔH° (kcal/mol)[a]		ΔS° (e.u.)[a]		T_m (°C)	
	pH 7.0	pH 5.0	pH 7.0	pH 5.0	pH 7.0	pH 5.0	pH 7.0	pH 5.0	pH 7.0	pH 5.0
G-C	–3.0		–6.5		–53.6		–152.0		79.8	
A-U	–2.2		–6.3		–60.9		–176.1		72.9	
U•G	–1.3	–1.4	–5.1	–5.2	–57.5	–58.4	–168.8	171.2	67.3	67.7
G•U	–0.9		–4.1		–47.0		–138.3		66.5	
G•A	–0.7		–3.7		–45.5		–134.5		64.8	
A•C	0.1	(–5.9)[b]	–2.6	(–9.4)[b]	–38.2	(–56.6)[b]	–114.9	(–152.1)[b]	59.3	59.2
C•A	0.3	(–7.1)[b]	–2.3	(–10.3)[b]	–36.8	(–52.2)[b]	–111.3	(–135.3)[b]	57.9	60.7
U•U	0.5		–2.4		–42.3		–128.6		55.7	

[a]Conservative estimates of standard errors for ΔG°_{60}, ΔG°_{37}, ΔH° and ΔS° are 3, 5, 7 and 8%, respectively (54,59).
[b]Best fits in parentheses were obtained by assuming duplex formation (the melting profiles were concentration dependent). The buffer conditions were 20 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 5.0 or 7.0 and 15 mM NaCl.

Thermodynamic parameters for microhelices at pH 5

Previous studies indicated that changes in mismatch structures in RNA can occur at lower pH (41,46,65). At pH ~5.0, A•C mismatches can form A^+•C wobble pairs that are isomorphous with G•U pairs (46,65). In order to understand more clearly the role of possible protonation sites in the mismatch base-pairs of microhelix RNAs, we examined the pH dependence of the thermodynamics for three mismatches, G•U, A•C and C•A. As shown in Table 1, there was no pH effect on the parameters for the G•U variant. In contrast, however, the protonated A•C and C•A (A^+•C and C•A^+) variants both formed stable duplexes under these conditions (thus, the values shown in parentheses in Table 1 cannot be compared to the values obtained from the hairpin data). At pH 7.0, formation of only one hydrogen bond between A•C or C•A is likely, therefore this mismatch is expected to destabilize RNA structure. At pH 5.0, the purine base can be protonated, thus leading to the formation of a second hydrogen bond. In this case, the A•C and C•A mismatches were stabilized at lower pH and the microhelix RNAs preferentially formed dimer structures. Our data are consistent with studies by Lee *et al*. (46) in which the stabilities of A•C and G•U pairs were compared at pH 5.0 and 7.0. In that study, the protonated A•C pair was more stable than the unprotonated pair, whereas G•U showed no pH dependence on stability.

Thermodynamic parameters for tetraloops at pH 7.0

The initial stability studies demonstrated the ability of the microhelix RNAs to form alternative structures (e.g., duplexes), therefore the loop sequence of the $tRNA^{Ala}$ microhelix was altered to avoid problems of self-complementarity. Two $tetraloop^{Ala}$ RNAs were synthesized with two different 4 bp stem sequences: 5′-GGXC-3′,3′-CCYG-5′ for $tetraloop^{Ala1}$ and 5′-GGXG-3′,3′-CCYC-5′ for $tetraloop^{Ala2}$, where X•Y was variable (G•U, U•G, G•A, A•C, C•A, U•U, A-U or G-C). The $tetraloop^{Ala}$ RNAs (where X•Y = G•U) were reported by Shi *et al*. (57) to be substrates of AlaRS and to have a high stability (T_ms = 74.5 and 73.3°C for $tetraloop^{Ala1}$ and $tetraloop^{Ala2}$, respectively, at 100 mM NaCl). As expected, the RNA hairpins with (UUCG) tetraloop sequences were more stable than the corresponding microhelices, consistent with the high stability of other hairpins containing this loop sequence (58). The G•U and mismatch variants of $tetraloop^{Ala}$ all formed hairpin structures at 15 mM NaCl, and had overall destabilizing effects on the hairpin relative to Watson–Crick base-pairs.

The thermodynamic parameters for the tetraloop RNAs are listed in Table 2. Examples of the normalized plots are shown in Figure 4A–D for the G•U, U•U, G•A, C•A and A•C variants. The observed order of stability (decreasing T_m) of internal X•Y pairs at pH 7.0 is G-C > A-U > G•U ≈ U•G > C•A > A•C ≈ G•A > U•U for the tetraloop RNAs. As with the microhelix RNAs, the trends with ΔG°_{60} and T_m are identical, and minor deviations are observed with the ΔG°_{37} trends. The G-C variant was too stable to obtain thermodynamic parameters (T_m > 85°C). Also similar to the microhelix RNAs, predicted thermodynamic values for tetraloop stability determined from MFOLD do not correlate well with the measured values. In addition, different trends are observed for this set of RNAs compared to the microhelix RNAs, providing further evidence that neighboring sequences and loop sizes can influence mismatch stability. It should also be noted that the stem sequence of $tetraloop^{Ala1}$ is symmetric (5′GXC-3′,3′-CYG-5′), yet the destabilizing effects of the C•A base-pair compared to the A•C pair are not the same (Table 2). The ΔG°_{60} values for the U•G and G•U variants differ by ~0.3 kcal/mol, and the C•A and A•C stabilities differ by ~0.6 kcal/mol, demonstrating a minor influence of next-nearest neighbors on stability of base-pair mismatches. In the case of the $tetraloop^{Ala1}$ RNAs, the stability of the G•U pair has less sequence and orientation dependence than the A•C pair. The tetraloop data agree with trends reported by Zhu and Wartell (43), in which the G•U pair is the most stable, the purine-containing mismatches (G•A, C•A and A•C) have intermediate stability, and pyrimidine–pyrimidine (U•U) mismatches are the least stable.

Table 2. Thermodynamics of $tetraloop^{Ala}$ RNAs

X•Y pair	ΔG°_{60} (kcal/mol)[a]		ΔG°_{37} (kcal/mol)[a]		ΔH° (kcal/mol)[a]		ΔS° (e.u.)[a]		T_m (°C)	
	pH 7.0	pH 5.0	pH 7.0	pH 5.0	pH 7.0	pH 5.0	pH 7.0	pH 5.0	pH 7.0	pH 5.0
G-C	ND		ND		ND		ND		>85	
A-U	–3.9		–7.6		–58.5		–164.0		83.2	
G•U	–2.1		–5.7		–53.2		–153.3		74.2	
U•G	–1.8	–1.7	–4.4	–4.4	–40.8	–41.5	–117.2	–119.6	74.9	73.9
C•A	–1.3	–3.2	–4.6	–6.7	–51.1	–53.0	–149.6	–149.5	68.1	81.5
A•C	–0.7	–2.6	–4.1	–6.2	–48.9	–54.5	–144.6	–155.9	65.0	76.7
G•A	–0.7	–1.3	–3.7	–4.5	–44.4	–47.4	–131.2	–138.4	65.2	69.5
U•U	0.2		–3.2		–48.3		–145.5		58.7	
G•U	–1.2		–5.0		–55.9		–164.1		67.5	

The upper portion of table shows data for $tetraloop^{Ala1}$ RNAs, 5′-GGXCUUCGGYCCACCA-3′, with X•Y mismatches or Watson–Crick base-pairs. The last row presents data for the $tetraloop^{Ala2}$ RNA, 5′-GGXGUUCGCYCCACCA-3′.

[a]Conservative estimates of standard errors for ΔG°_{60}, ΔG°_{37}, ΔH° and ΔS° are 3, 5, 7 and 8%, respectively (54,59). The buffer conditions were 20 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 5.0 or 7.0 and 15 mM NaCl.

ND, not determined.

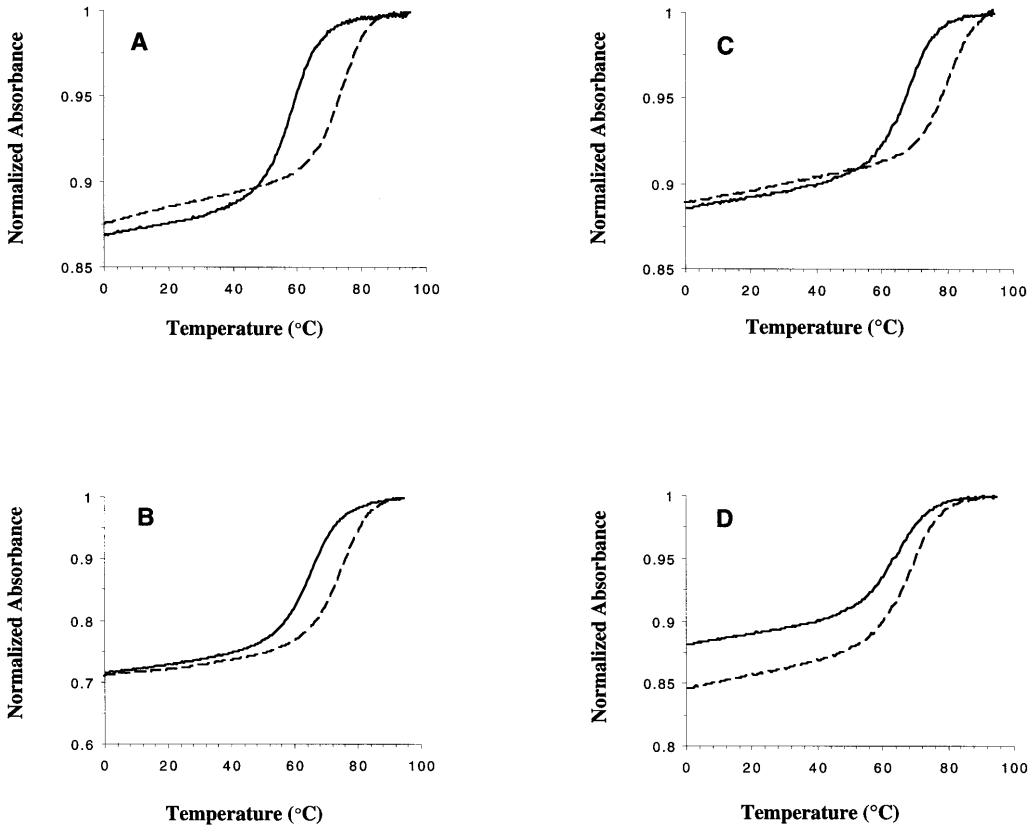

Figure 4. Normalized UV melting curves for the tetraloop RNAs in 15 mM NaCl, 20 mM sodium cacodylate, 0.5 mM EDTA. (**A**) Curves for the U•U and G•U tetraloops obtained at pH 7.0 (solid line, U•U, 55 µM RNA; dashed line, G•U, 63 µM RNA) are shown. Curves for the A•C (**B**), C•A (**C**) and G•A (**D**) tetraloops obtained at pH 7.0 or 5.0 (solid lines, pH 7.0; dashed lines, pH 5.0) are presented. The melting curves were normalized at 95°C.

As shown in Table 2, tetraloop^{Ala2} stability is decreased by 0.9 kcal/mol for ΔG°_{60} relative to tetraloop^{Ala1}. These two RNAs have the same first three base-pairs, but differ in the closing base-pair (5′-GC-3′ for tetraloop^{Ala2} versus 5′-CG-3′ for tetraloop^{Ala1}). Corroborating results were observed by Antao *et al.* (66) in which a 5′-C(UUCG)G-3′ tetraloop was more stable than a 5′-G(UUCG)C-3′ tetraloop (T_ms = 71.7 and 60.1°C, respectively), although this effect can be dependent on the sequences that are adjacent to the closing base-pair (67). Both G•U-containing tetraloop^{Ala} RNAs were more stable than the G•U microhelix^{Ala} variant, confirming the importance of the loop size and closing base-pair for overall RNA stability, as observed previously by Serra *et al.* (33,67,68).

Thermodynamic parameters for tetraloops at pH 5

The effects of pH on stability or structure can be more easily assessed with the tetraloops than with the microhelices because they form hairpins at both pH 7.0 and 5.0. The protonated A variants (G•A$^+$, A$^+$•C and C•A$^+$) were all stabilized relative to the unprotonated RNAs ($\Delta\Delta G^\circ_{60}$, pH 7.0–5.0 = 0.6–1.9 kcal/mol; see Table 2) (normalized plots are shown in Fig. 4), whereas pH had no effect on the U•G variant. Similarly, Morse and Draper (41) demonstrated that G•A-containing RNA duplexes can be stabilized at lower pH by 0.5–6 kcal/mol (ΔG°_{37}), depending on the sequence contexts. They also provided evidence for protonated G•A pairs in RNA helices which led to unusual duplex structures, and the effects were sequence dependent. G•A$^+$ pairs are formed only in certain sequence contexts, therefore the effects on ΔG°_{37} or ΔG°_{60} will also be context dependent. Our thermodynamic data are suggestive of a protonated structure for the G•A tetraloop^{Ala1} variant at pH 5.0. The results presented in this section also reveal that mismatch stabilities for the tetraloops need to be reordered at pH 5.0 so that C•A$^+$ and A$^+$•C are > G•U.

CD spectra of tetraloops at pH 7.0 and 5.0

The CD spectra of tetraloop^{Ala1} RNAs were measured in order to compare the effects of the various mismatches on the overall folded structure. The spectra of the mismatch RNAs and Watson–Crick variants all have maxima centered around 264 nm and minima near 240 nm, similar to other A-RNAs (Fig. 5A–C) (data not shown for the Watson–Crick variants). The CD spectra of the mismatch RNA tetraloops exhibit subtle differences in the maximum and minimum wavelengths, peak heights (molar ellipticity) and crossover points, suggesting that the various base-pair mismatches have only a minor influence on the overall RNA structure. This difference is most apparent between the G•U and U•G variants (Fig. 5A). The A•C and C•A tetraloop variants have essentially identical CD spectra at pH 5.0 (Fig. 5B), as well as at pH 7.0 (data not shown). The G•A variant, also measured at pH 5.0 (Fig. 5C), exhibits a slightly different spectrum than the

other mismatches with a smaller positive Δε value at the peak maximum of 220 nm and different crossover points. At pH 7.0 (data not shown), the G•A variant has a similar CD spectrum, but with smaller positive Δε values at both 265 and 220 nm. Together, the CD results suggest that the single base-pair position at 3•70 has only a minor influence on the overall RNA structure. The G•U and U•G variants are located in symmetric sequence environments, but could have different base stacking arrangements or helix winding angles based on their slightly different CD spectra. Conversely, the C•A and A•C variants differ in stabilities by ~0.6 kcal/mol at pH 5.0 and 7.0, but exhibit essentially overlapping CD spectra at both pH values. Recent NMR studies revealed a deviation between an A-form RNA helix with a G3•C70 pair and one with a G3•U70 pair in the context of a $tRNA^{Ala}$ acceptor stem duplex (69,70). In contrast, a C•A mismatch at the same 3•70 position of the RNA duplex destacks in the opposite direction as the G•U pair (71). Although the NMR structures revealed local structural changes associated with the mismatches, the global RNA structures were essentially A-form (69–71).

Conclusions

Several conclusions can be made from these studies regarding the relative stabilities of single internal mismatches in different RNA sequence contexts. First, all of the RNA mismatches examined in three different sequence contexts were destabilizing to the RNA hairpin relative to Watson–Crick base-pairs. Comparison of our thermodynamic parameters with predicted values is difficult, however, because the predictions are based on free energy values for tandem or terminal mismatches at 1 M NaCl (60). Due to a lack of available data, Serra and Turner assumed that the free energy contributions of single internal mismatches are independent of their own identity or sequence contexts (60). Zhu and Wartell (43) suggested a modification of this rule in order to accommodate data from TGGE studies. Our results are consistent with those of Zhu and Wartell, but different sequence contexts were examined in our studies (5′-GCG-3′,3′-CYC-5′ and 5′-GXC-3′,3′-CYG-5′). RNAs containing G•U and U•G mismatches are the most stable relative to other mismatch variants. Mismatches with at least one purine base (e.g., C•A, A•C, G•A) are more stable than two pyrimidine bases (e.g., U•U). Different trends in stability are observed, however, between the microhelix and tetraloop variants, suggesting that stability is influenced by other factors such as loop size and sequence or neighboring sequences. Different influences on mismatch stability by the unpaired 3′ terminal ACCA are also possible (25), but not examined in this study. Small differences in free energy values for the G•U versus U•G and A•C versus C•A variants in $tetraloop^{Ala1}$, which has a symmetric sequence surrounding the mismatch site (5′-GXC-3′,3′-CYG-5′), suggest the importance of next-nearest neighbors in determining stability. Although these results are corroborated by several recent reports (43,44), the need for improved thermodynamics that predict single internal mismatch stabilities in different sequence and structural contexts is still apparent.

The G•U variants of the chosen RNAs are known substrates for AlaRS (56,57). Several groups have debated the importance of a helix distortion in the cognate tRNA for AlaRS recognition and employed a series of base mismatches to test such a model (45,49,72). Here, we have considered the possible differences in RNA structure when a G•U site is substituted with other base-pair mismatches. The CD spectra shown in Figure 5 indicate that G•U, G•A, C•A, A•C and U•U tetraloop RNAs exist in solution as A-form helices, but exhibit subtle differences in peak maxima, minima and crossover points. The most notable differences are observed with the G•U and G•A variants, suggesting possible differences in local base-stacking arrangements or other helix parameters. Chemical probing experiments with a rigid rhodium complex without any potential for hydrogen bonding (20) indicate that the G•U structure is unique relative to other mismatches and the local G•U structure is maintained in both the microhelix and tetraloop RNAs (M.Meroueh and C.Chow, unpublished results). Based on these results, we suggest that local

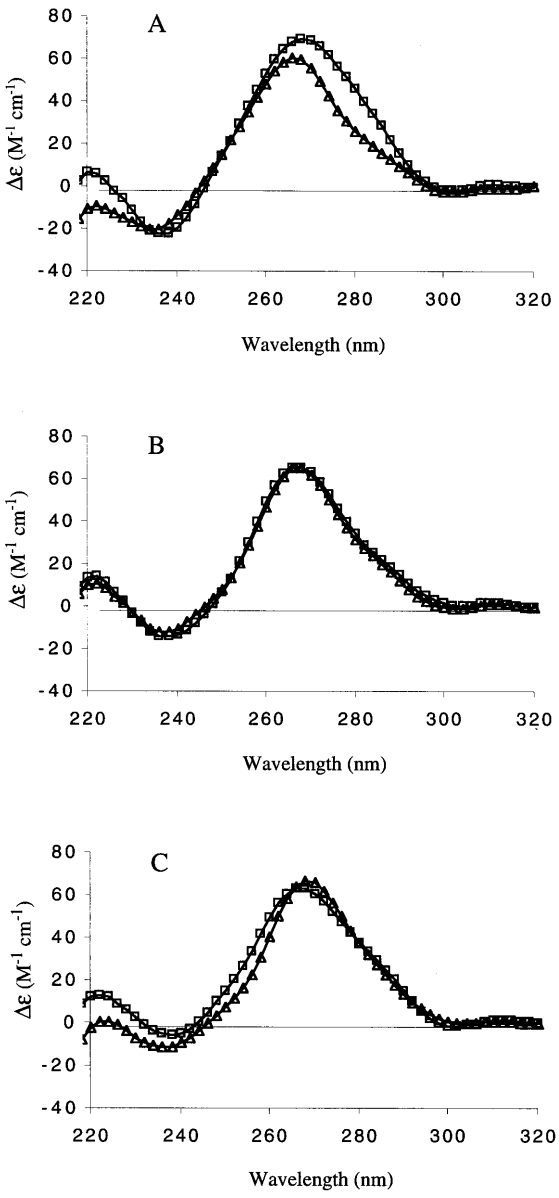

Figure 5. CD spectra of the mismatch $tetraloop^{Ala}$ RNAs. The molar ellipticities are normalized to RNA concentrations (3.0×10^{-6} M in molecules of RNA). Each spectra is an average of four scans. (**A**) Spectra of the G•U (□) and U•G (Δ) variants (pH 7.0) with overlays of (**B**) A•C (□) and C•A (Δ) (pH 5.0) and (**C**) U•U (□) (pH 7.0) and G•A (Δ) (pH 5.0).

stability might also play a role in the recognition of mismatch-containing tRNAs by AlaRS.

We have shown here that specific base mismatches lead to RNA instability, which is influenced by sequence contexts. The mismatch instability might also lead to the formation of alternative structures upon protein binding to the RNA. Two independent NMR structures of G•U-containing duplexes revealed a local helical distortion at the G•U site (69,70). The G•U pair can destabilize the tRNA acceptor stem, and this might be important for forming an optimal AlaRS active site geometry by an induced-fit mechanism (70). The instability observed with other single mismatches (non-Watson–Crick or non-G•U) could play a role in such a mechanism. Furthermore, Vogtherr *et al*. (71) showed by NMR that C•A and G•U base-pairs, although both recognition elements for AlaRS, are destacked in opposite directions in $tRNA^{Ala}$ duplexes. These results suggest that the enzyme might be recognizing a general helix distortion, or 'locally enhanced flexibility', of the mismatch, rather than a static structure (71). The Watson–Crick or U•G pairs might be too stable to distort and present specific functional groups to the enzyme active site. In contrast, the less stable mismatches might enable the acceptor stem to adapt its orientation in order to present complementary functionalities to AlaRS and allow efficient aminoacylation.

Detailed X-ray or NMR studies with mismatch RNAs and co-crystals with proteins or small molecules will be necessary in order to gain a complete understanding of the relative contributions of local conformational changes to stability changes and ligand specificities. We have shown here that considerations of the RNA sequences and specific experimental conditions, particularly pH and sodium ion concentrations, will be important in such studies with RNAs containing single internal mismatches.

ACKNOWLEDGEMENTS

We are grateful to Professor John SantaLucia, Dr Hatim Allawi and Shikha Varma for their advice and many helpful discussions, and for allowing us to use their UV-visible spectrophotometer. This work was supported through startup funds and an Interdisciplinary Research Stimulation Grant from Wayne State University.

REFERENCES

1 Crick,F.H.C. (1966) *J. Mol. Biol.*, **19**, 548–555.
2 Gutell,R.R., Larsen,N. and Woese,C.R. (1994) *Microb. Rev.*, **58**, 10–26.
3 Gautheret,D., Konings,D. and Gutell,R.R. (1995) *RNA*, **1**, 807–814.
4 Kim,S.H., Sussman,J.L., Suddath,F.L., Quigley,G.J., McPherson,A., Wang,A.H.J., Seeman,N.C. and Rich,A. (1974) *Proc. Natl Acad. Sci. USA*, **71**, 4970–4974.
5 Kim,S.H., Suddath,F.L., Quigley,G.J., McPherson,A., Sussman,J.L., Wang,A.H.J., Seeman,N.C. and Rich,A. (1974) *Science*, **185**, 435–440.
6 Westhof,E., Dumas,P. and Moras,D. (1985) *J. Mol. Biol.*, **184**, 119–145.
7 Mizuno,H. and Sundaralingam,M. (1978) *Nucleic Acids Res.*, **5**, 4451–4461.
8 Alkema,D., Hader,P.A., Bell,R.A. and Neilson,T. (1982) *Biochemistry*, **21**, 2109–2117.
9 Freier,S.M., Kierzek,R., Caruthers,M.H., Neilson,T. and Turner,D.H. (1986) *Biochemistry*, **25**, 3209–3213.
10 He,L., Kierzek,R., SantaLucia,J.,Jr, Walter,A.E. and Turner,D.H. (1991) *Biochemistry*, **30**, 11124–11132.
11 McClain,W.H. and Foss,K. (1988) *Science*, **240**, 793–796.
12 Hou,Y.-M. and Schimmel,P. (1988) *Nature*, **333**, 140–145.
13 Li,H., Dalal,S., Kohler,J., Vilardell,J. and White,S.A. (1995) *J. Mol. Biol.*, **250**, 447–459.
14 White,S.A. and Li,H. (1996) *RNA*, **2**, 226–234.
15 Bénard,L., Mathy,N., Grunberg-Manago,M., Ehresmann,B., Ehresmann,C. and Portier,C. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 2564–2567.
16 Mizutani,T., Tanabe,K. and Yamada,K. (1998) *FEBS Lett.*, **429**, 189–193.
17 Pyle,A.M., Moran,S., Strobel,S.A., Chapman,T., Turner,D.H. and Cech,T.R. (1994) *Biochemistry*, **33**, 13856–13863.
18 Strobel,S.A. and Cech,T.R. (1995) *Science*, **267**, 675–679.
19 Abramovitz,D.L., Friedman,R.A. and Pyle,A.M. (1996) *Science*, **271**, 1410–1413.
20 Chow,C.S. and Barton,J.K. (1992) *Biochemistry*, **31**, 5423–5429.
21 Hickerson,R.P., Watkins-Sims,C.D., Burrows,C.J., Atkins,J.F., Gesteland,R.F. and Felden,B. (1998) *J. Mol. Biol.*, **279**, 577–587.
22 Burgstaller,P. and Famulok,M. (1997) *J. Am. Chem. Soc.*, **119**, 1137–1138.
23 Burgstaller,P., Hermann,T., Huber,C., Westhof,E. and Famulok,M. (1997) *Nucleic Acids Res.*, **25**, 4018–4027.
24 Ott,G., Arnold,L. and Limmer,S. (1993) *Nucleic Acids Res.*, **21**, 5859–5864.
25 Limmer,S., Hofmann,H.-P., Ott,G. and Sprinzl,M. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 6199–6202.
26 Auffinger,P. and Westhof,E. (1997) *J. Mol. Biol.*, **269**, 326–341.
27 Allain,F.H.-T. and Varani,G. (1995) *Nucleic Acids Res.*, **23**, 341–350.
28 Cate,J.H. and Doudna,J.A. (1996) *Structure*, **4**, 1221–1229.
29 Correll,C.C., Freeborn,B., Moore,P.B. and Steitz,T.A. (1997) *Cell*, **91**, 705–712.
30 Kieft,J.S. and Tinoco,I.,Jr (1997) *Structure*, **5**, 713–721.
31 Suga,H., Cowan,J.A. and Szostak,J.W. (1998) *Biochemistry*, **37**, 10118–10125.
32 Sugimoto,N., Kierzek,R., Freier,S.M. and Turner,D.H. (1986) *Biochemistry*, **25**, 5755–5759.
33 Serra,M.J., Lyttle,M.H., Axenson,T.J., Schadt,C.A. and Turner,D.H. (1993) *Nucleic Acids Res.*, **21**, 3845–3849.
34 McDowell,J.A. and Turner,D.H. (1996) *Biochemistry*, **35**, 14077–14089.
35 McDowell,J.A., He,L., Chen,X. and Turner,D.H. (1997) *Biochemistry*, **36**, 8030–8038.
36 SantaLucia,J.,Jr, Kierzek,R. and Turner,D.H. (1990) *Biochemistry*, **29**, 8813–8819.
37 SantaLucia,J.,Jr, Kierzek,R. and Turner,D.H. (1991) *Biochemistry*, **30**, 8242–8251.
38 Walter,A.E., Wu,M. and Turner,D.H. (1994) *Biochemistry*, **33**, 11349–11354.
39 Wu,M., McDowell,J.A. and Turner,D.H. (1995) *Biochemistry*, **34**, 3204–3211.
40 Xia,T., McDowell,J.A. and Turner,D.H. (1997) *Biochemistry*, **36**, 12486–12497.
41 Morse,S.E. and Draper,D.E. (1995) *Nucleic Acids Res.*, **23**, 302–306.
42 Peritz,A.E., Kierzek,R., Sugimoto,N. and Turner,D.H. (1991) *Biochemistry*, **30**, 6428–6436.
43 Zhu,J. and Wartell,R.M. (1997) *Biochemistry*, **36**, 15326–15335.
44 Xiang,Q., Qin,P.Z., Michels,W.J., Freeland,K. and Pyle,A.M. (1998) *Biochemistry*, **37**, 3839–3849.
45 Gabriel,K., Schneider,J. and McClain,W.H. (1996) *Science*, **271**, 195–197.
46 Lee,K., Varma,S., SantaLucia,J.,Jr and Cunningham,P.R. (1997) *J. Mol. Biol.*, **269**, 732–743.
47 Musier-Forsyth,K., Shi,J.-P., Henderson,B., Bald,R., Fürste,J.P., Erdmann,V.A. and Schimmel,P. (1995) *J. Am. Chem. Soc.*, **117**, 7253–7254.
48 Musier-Forsyth,K., Usman,N., Scaringe,S., Doudna,J., Green,R. and Schimmel,P. (1991) *Science*, **253**, 784–786.
49 Beuning,P.J., Yang,F., Schimmel,P. and Musier-Forsyth,K. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 10150–10154.
50 Capaldi,D.C. and Reese,C.B. (1994) *Nucleic Acids Res.*, **22**, 2209–2216.
51 Usman,N., Ogilvie,K.K., Jiang,M.-Y. and Cedergren,R.J. (1987) *J. Am. Chem. Soc.*, **109**, 7845–7854.
52 SantaLucia,J.,Jr, Allawi,H.T. and Seneviratne,P.A. (1996) *Biochemistry*, **35**, 3555–3562.
53 Richards,E.G. (1975) In Fasman,G.D. (ed.), *Handbook of Biochemistry and Molecular Biology. Nucleic Acids*, CRC Press, Cleveland, OH, pp. 596–599.
54 Allawi,H.T. and SantaLucia,J.,Jr (1997) *Biochemistry*, **36**, 10581–10594.
55 Cantor,C.R. and Schimmel,P.R. (1980) *Biophysical Chemistry. Part II: Techniques for the Study of Biological Structure and Function*. W.H.Freeman and Company, San Francisco, CA, pp. 412–413.
56 Francklyn,C. and Schimmel,P. (1989) *Nature*, **337**, 478–481.
57 Shi,J.-P., Martinis,S.A. and Schimmel,P. (1992) *Biochemistry*, **31**, 4931–4936.
58 Tuerk,C., Gauss,P., Thermes,C., Groebe,D.R., Gayle,M., Guild,N., Stormo,G., D'Aubenton-Carafa,Y., Uhlenbeck,O.C., Tinoco,I.,Jr, Brody,E.N. and Gold,L. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 1364–1368.
59 SantaLucia,J.,Jr, Kierzek,R. and Turner,D.H. (1992) *Science*, **256**, 217–219.

60 Serra,M.J. and Turner,D.H. (1995) *Methods Enzymol*., **259**, 242–261.
61 Zuker,M., Jaeger,J.A. and Turner,D.H. (1991) *Nucleic Acids Res*., **19**, 2707–2714.
62 Williams,D.J. and Hall,K.B. (1996) *Biochemistry*, **35**, 14665–14670.
63 Allawi,H.T. and SantaLucia,J.,Jr (1998) *Biochemistry*, **37**, 9435–9444.
64 Allawi,H.T. and SantaLucia,J.,Jr (1998) *Nucleic Acids Res*., **26**, 2694–2701.
65 Puglisi,J.D., Wyatt,J.R. and Tinoco,I.,Jr (1990) *Biochemistry*, **29**, 4215–4226.
66 Antao,V.P., Lai,S.Y. and Tinoco,I.,Jr (1991) *Nucleic Acids Res*., **19**, 5901–5905.
67 Serra,M.J., Barnes,T.W., Betschart,K., Gutierrez,M.J., Sprouse,K.J., Riley,C.K., Stewart,L. and Temel,R.E. (1997) *Biochemistry*, **36**, 4844–4851.
68 Serra,M.J., Axenson,T.J. and Turner,D.H. (1994) *Biochemistry*, **33**, 14289–14296.
69 Limmer,S., Reif,B., Ott,G., Arnold,L. and Sprinzl,M. (1996) *FEBS Lett*., **385**, 15–20.
70 Ramos,A. and Varani,G. (1997) *Nucleic Acids Res*., **25**, 2083–2090.
71 Vogtherr,M., Schübel,H. and Limmer,S. (1998) *FEBS Lett*., **429**, 21–26.
72 McClain,W.H., Chen,Y.-M., Foss,K. and Schneider,J. (1988) *Science*, **242**, 1681–1684.