J|A|C|S
ARTICLES
Published on Web 03/14/2002

Internally Mismatched RNA: pH and Solvent Dependence of the Thermal Unfolding of tRNA^{Ala} Acceptor Stem Microhairpins

Ewa Biała† and Peter Strazewski*,‡

Contribution from the Institute of Organic Chemistry, University of Basel, St. Johanns-Ring 19, CH - 4056 Basel, Switzerland

Received May 3, 2001

Abstract: The thermal unfolding of two RNA hairpin systems derived from the aminoacyl accepting arm of *Escherichia coli* tRNA^{Ala} that included all possible single internal mismatches mostly in the third base pair position was measured spectroscopically in 0.1 M NaCl at pH 7.5 and, in part, 5.5. The thermodynamic parameters $\Delta H°$, $\Delta S°$, $\Delta G°$, and T_m of a total of 36 RNA strands were determined through nonlinear curve fitting of the melting profiles (22 tetralooped 22mers and 14 heptalooped 25mers, same stem sequence). Only three of the 22mers, the A·C-containing variants, were shown to be significantly more stable at pH 5.5. A number of remarkable differences—most likely of more general relevance—between the thermodynamics of certain structurally very similar hairpin variants (e.g., G·C versus A·U, G·U versus I·U) at pH 7.5 are discussed with respect to two possible ways of helix stabilization: pronounced hydration versus low entropic penalty. Four selected 22mers were additionally analyzed in 1 M NaCl and in solvent mixtures containing ethanol, ethylene glycol, and dimethylformamide. The wealth of thermodynamic data suggest that the exothermicity $\Delta H°$ and entropic penalty $T·\Delta S°$ of folding are strongly dominated by the rearrangement and formation of hydration layers around the solutes, while it is well-known that the stability of folding results only from the difference ($\Delta G°$) and ratio of both parameters ($T_m = \Delta H°/\Delta S°$).

Introduction

Background. The double-helical region of the aminoacyl acceptor arm of transfer RNAs (tRNAs) bears important recognition elements for a number of tRNA-binding enzymes such as elongation factors (EF)[1] and aminoacyl tRNA synthetases (ARS).[2] The tRNA from *Escherichia coli* specific for alanine (tRNA^{Ala}), for instance, binds to its cognate alanyl tRNA synthetase (AlaRS) through various contacts at the end of the acceptor helix. The major determinant for the alanylation is a G·U wobble base pair at the third position (G3·U70) in the acceptor helix of the tRNA.[3] A number of site-specific tRNA^{Ala} variants bearing mispairs other than G3·U70 were shown to be active and some viable to different degrees when introduced into tRNA^{Ala} knockout strains.[4] The possibility that a differential local deformability of the tRNA^{Ala} acceptor stem—reflected in the thermodynamic parameters of local acceptor stem denaturation—could influence, possibly in an induced fit manner, AlaRS recognition and aminoacylation activity prompted two labs to investigate thermal denaturation of RNA hairpins that resembled the acceptor stem of *E. coli* tRNA^{Ala}.[5] Merouch and Chow[5a] published the full thermodynamic parameters of a subset of 3·70 mismatched variants of three acceptor hairpins differing in stem length and loop sequence (hairpin nucleotide positions according to tRNA numbering). We measured the denaturation profiles of all possible 3·70 variants of two acceptor hairpins differing in loop length and sequence and compared the in vivo activity of the corresponding *E. coli* tRNA^{Ala} 3·70 variants with the denaturation stability of 33 RNA hairpins expressed as midpoint transition temperatures T_m.[5b]

Objectives. Here we present—quite apart from the question of tRNA−protein recognition—the full analysis of the UV-detected denaturation profiles of the above (33 plus three more) RNA haipins, since (i) sufficient data of thermodynamic parameters of *all* possible internal RNA/RNA mismatches within different sequence contexts are still lacking and (ii) because the

* Author for correspondence. E-mail: peter.strazewski@unibas.ch.
† Permanent address: Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, Poland. E-mail: biala@ibch.poznan.pl.
‡ Present address: Laboratoire de Synthèse de Biomolécules, Bâtiment E. Chevreul (5ème étage), Université Claude Bernard-Lyon 1, 43 boulevard du 11 novembre 1918, F-69622 Villeurbanne (Cedex), France.

(1) (a) Nissen, P.; Kjeldgaard, M.; Thirup, S.; Polekhina, G.; Reshetnikova, L.; Clark, B. F.; Nyborg, J. *Science* **1995**, *270*, 1464−72. (b) Nissen, P.; Thirup, S.; Kjeldgaard, M.; Nyborg, J. *Structure Fold Des.* **1999**, *7*, 143−56.
(2) (a) Rould, M. A.; Perona, J. J.; Söll, D.; Steitz, T. A. *Science* **1989**, *246*, 1135−42. (b) Ruff, M.; Krishnaswamy, S.; Boeglin, M.; Poterszman, A.; Mitschler, A.; Podjarny, A.; Rees, B.; Thierry, J. C.; Moras, D. *Science* **1991**, *252*, 1682−9. (c) Sankaranarayanan, R.; Dock-Bregeon, A. C.; Romby, P.; Caillet, J.; Springer, M.; Rees, B.; Ehresmann, C.; Ehresmann, B.; Moras, D. *Cell* **1999**, *97*, 371−81. (d) Delagoutte, B.; Keith, G.; Moras, D.; Cavarelli, J. *Acta Crystallogr. D: Biol. Crystallogr.* **2000**, *56*, 492−4.
(3) (a) McClain, W. H.; Foss, K. *Science* **1988**, *240*, 793−796. (b) Hou, Y.-M.; Schimmel, P. *Nature* **1988**, *333*, 140−145. (c) McClain, W. H.; Chen, Y.-M.; Foss, K.; Schneider, J. *Science* **1988**, *242*, 1681−1684. (d) McClain, W. H.; Foss, K.; Jenkins, R. A.; Schneider, J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9272−9276. (e) Musier-Forsyth, K.; Schimmel, P. *Nature* **1992**, *357*, 513−15.
(4) McClain, W. H.; Jou, Y. Y.; Bhattacharya, S.; Gabriel, K.; Schneider, J. *J. Mol. Biol.* **1999**, *290*, 391−409.
(5) (a) Meroueh, M.; Chow, C. S. *Nucleic Acids Res.* **1999**, *27*, 1118−25. (b) Strazewski, P.; Biała, E.; Gabriel, K.; McClain, W. H. *RNA* **1999**, *5*, 1490−4.

10.1021/ja0161305 CCC: $22.00

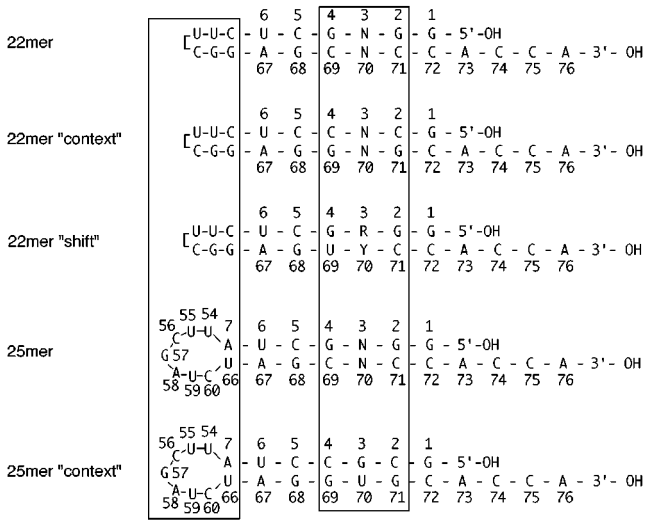

Figure 1. Studied tRNA^{Ala} acceptor hairpins. Nucleotide numbering according to the position within a tRNA. N: any ribonucleotide. R Y: Watson−Crick pair, R = purine, Y = pyrimidine.

relationship between the enthalpic and the entropic contributions to RNA unfolding thermodynamics sheds an interesting light on the enthalpy−entropy and exothermicity−stability relationship of nucleic acid double-helix formation. This latter point is dealt with in the article following this one. We elucidated from the recorded and normalized melting profiles the enthalpy and entropy changes, $\Delta H°$ and $\Delta S°$, respectively, that accompany the thermal unfolding by means of curve fitting using a T_m-independent method, i.e., using the equilibrium equation that describes a unimolecular two-state transition, $\alpha(T) = 1/(1 + e^{(\Delta H° - T\cdot\Delta S°)/RT})$, with α being the mole fraction 0−1 of paired versus unpaired RNA molecules, $\Delta H°$ and $\Delta S°$ being the to be optimized parameters, and R the universal gas constant. We calculated the free energy of pairing, $\Delta G°_T = \Delta H° - T\cdot\Delta S°$, and the mid-transition temperatures $T_{\Delta G°=0} = T_m$(unimolec.) = $\Delta H°/\Delta S°$. From the observed pH and solvent dependencies of the thermodynamic parameters we conclude that, while the *folding stability* is best described by $\Delta G°_T$ and T_m, the *extent and strength of solvation* of the folded structures finds its signature in the accompanying exothermicity $\Delta H°$ and entropic penalty $T\cdot\Delta S°$.

Results

The analyzed acceptor hairpins (22mers and 25mers depicted in Figure 1) bear the same 6 bp double-helical stem motif and 4 nt single-stranded 3′-overhang sequence but differ from each other in their loop sequence and length, the seventh loop-closing base pair, and base−base combination at position 3·70 (framed in the figure). Three hairpins also differ from the other 31 in the flanking positions 2·71 and 4·69—the parent G·C pairs have been switched to C·G pairs; the corresponding hairpins were denoted as "context" and another two bear a G·U mismatch in position 4·69 and are denoted "shift" in the tables to follow. The 22 22mers combine all possible base−base combinations in position 3·70 including two inosine·uridine combinations (I3·U70 and U3·I70), two "context" strands ("G3·U70 context" and "C3·A70 context"), and two "shift" strands ("A3·U70-G4·U70 shift" and "G3·C70-G3·U69 shift"). The 14 25mers lack G·G, U·I, A·G, A·A, and U·C in position 3·70, as well as the corresponding "C3·A70 context" and both "shift" strands.

The most obvious differences in the melting profiles between the 22mers and the 25mers are the low- and high-temperature baselines.[6] The more stable 22mers closed by a C(UUCG)G tetraloop produced nondrifting straight baselines over a sufficiently large low-temperature region indicative, albeit not proof, of a two-state equilibrium between the fully denatured strand and the hairpin structure. The 25mers showed linearly drifting low- and high-temperature baselines (with varying slopes mainly of the high-temperature baselines), suggesting a significant population of additional possibly preorganized[7a] folding intermediates in the larger and more flexible 7 nt loop. Hence, the normalization of the 22mer data from absorbance at 260 nm, $A_{260}(T)$, to mole fraction $\alpha(T)$ involved merely two additional parameters a and b for the linear scaling of the experimental datapoints: $A_{260}(T) = a + b\cdot\alpha(T)$. The profiles of the 25mers, in contrast, could only be converted into two-state $\alpha(T)$ profiles by means of two preceding linear regressions of both the high- and the low-temperature baseline region separately, the difference of which was then used for the scaling of $A_{260}(T)$ to $\alpha(T)$.[7b]

The A_{260} profiles of the 22mers and 25mers in 0.1 M NaCl/10 mM phosphate-buffered solutions (3−6 parallel runs per strand) were concentration-independent over a 45-fold concentration range.[8] The C3·G70 and G3·C70 25mers showed a small temperature-dependent second transition at approximately 30−40° below the main transition contributing approximatively 15% (C3·G70) and 5% (G3·C70) to the total hyperchromicity, thus not posing any problems for the curve fitting onto the main transition.[6] The T_m values calculated from $\Delta H°/\Delta S°$ (Table 1) are within the standard deviation of ±0.5° all fully consistent with the values T_m read out as temperatures at half the observed relative hyperchromicity of the main transition as published earlier.[5b]

pH Dependence. All oligoribonucleotides were measured at physiological pH 7.5 and ionic strength 0.1 M NaCl. Several base pair combinations can in principle be stabilized by an additional proton, C·A$^+$ wobble, C·C$^+$ wobble, U·C$^+$ wobble, and G(syn)·A$^+$(anti) Hoogsteen pairing modes,[9] and would then show pH-dependent pairing thermodynamics. In the 22mer series the C·A, C·A context, A·C, C·C, U·C, C·U, G·A, A·G, A·A, and G·U 3·70 variants were additionally measured at pH 5.5.

Out of those candidates, only the C3·A70 context, C3·A70, and A3·C70 variants exhibited a significant pH dependence in their pairing strength: $\Delta\Delta G°_{25°C}$(pH 7.5−5.5) = 3.6, 3.8, and 2.1 kcal/mol; ΔT_m(pH 7.5−5.5) = 6.3°, 5.9°, and 5.3°,

(6) Examples of melting profiles depicted in ref 5b; more examples are in the Supporting Information to this work.

(7) (a) Vesnaver, G.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 3569−73. (b) Marky, L. A.; Breslauer, K. J. *Biopolymers* **1987**, *26*, 1601−20.

(8) In agreement with a large number of studies on UUCG-containing RNA sequences showing that hairpins are the predominant secondary structure even at mM strand concentrations: Williams, D. J.; Hall, K. B. *J. Mol. Biol.* **2000**, *297*, (a) 1045−61; (b) 251−65. (c) Williams, D. J.; Hall, K. B. *Biophys. J.* **1999**, *76*, 3192−205. (d) Chang, K. Y.; Varani, G.; Bhattacharya, S.; Choi, H.; McClain, W. H. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11764−9. (e) Ramos, A.; Varani, G. *Nucleic Acids Res.* **1997**, *25*, 2083−90. (f) Reif, B.; Wittmann, V.; Schwalbe, H.; Griesinger, C.; Wörner, K.; Jahn-Hofmann, K.; Engels, J. W.; Bermel, W. *Helv. Chim. Acta* **1997**, *80*, 1952−1971. (g) Miller, J. L.; Kollman, P. A. *J. Mol. Biol.* **1997**, *270*, 436−450. (h) Kanyo, J. E.; Duhamel, J.; Lu, P. *Nucleic Acids Res.* **1996**, *24*, 4015−4022. (i) Molinaro M, Tinoco, I., Jr. *Nucleic Acids Res.* **1995**, *23*, 3056−63. (j) Antao, V. P.; Lai, S. Y.; Tinoco, I., Jr. *Nucleic Acids Res.* **1991**, *19*, 5901−05. (k) Cheong, C., Varani, G., Tinoco, I., Jr. *Nature* **1990**, *346*, 680−682.

Table 1. Midpoint Transition Temperatures and Free Energies of Folding of All 22mers and 25mers

no.	3·70 nt, total nt	T_m [°C] (±1.0°)[a]	$\Delta G°_{25°C}$ [kcal/mol] (±0.8 kcal/mol)[a]
1	G·C 22	100.0	−13.4
2	C·G 22	96.9	−13.6
3	A·U 22	91.6	−14.9
4	U·A 22	90.7	−13.4
5	G·C-G·U shift 22	89.7	−12.8
6	C·A^+ context 22	88.0	−15.0
7	G·U 22	87.6	−11.1
8	G·U context 22	86.9	−14.6
9	A^+·C 22	86.4	−13.0
10	C·A^+ 22	86.1	−13.4
11	I·U 22	84.1	−12.8
12	U·G 22	84.1	−11.9
13	U·I 22	83.3	−14.5
14	G·G 22	83.3	−10.0
15	C·A context 22	81.7	−11.4
16	G·A 22	81.5	−9.4
17	C·A 22	80.8	−9.6
18	C·U 22	80.5	−9.1
19	A·C 22	80.5	−10.9
20	A·U-G·U shift 22	79.9	−11.2
21	A·G 22	79.9	−9.3
22	U·C 22	78.9	−8.3
23	C·C 22	78.5	−9.3
24	U·U 22	77.7	−7.9
25	A·A 22	77.3	−8.1
26	G·C 25	85.4	−11.5
27	C·G 25	84.2	−13.1
28	U·A 25	77.8	−12.6
29	A·U 25	76.4	−12.0
30	G·U 25	72.9	−10.5
31	U·G 25	70.5	−10.0
32	G·U context 25	70.2	−10.3
33	I·U 25	68.1	−10.0
34	G·A 25	64.5	−8.0
35	A·C 25	63.7	−8.2
36	C·A 25	63.0	−8.3
37	C·U 25	61.4	−8.1
38	U·U 25	60.3	−7.6
39	C·C 25	59.5	−6.1

[a] Approximate 90% confidence range. Denaturation profiles and confidence statistics: see Supporting Information.

respectively (Table 1). The corresponding changes in exothermicity, $\Delta\Delta H°$(pH 7.5−5.5) = 14.4, 18.0, and 6.7 ± 5 kcal/mol, in entropic penalty, 298 K · $\Delta\Delta S°$(pH 7.5−5.5) = 10.8, 14.2, and 4.6 ± 5 kcal/mol, respectively, reveal subtle sequence context effects. Meroueh and Chow's[5a] shorter four base pair stem-UUCG tetraloop hairpin A3·C70 variant bearing the same nearest neighboring base pairs as ours exhibited very similar differences in $\Delta G°_{25°C}$, $\Delta H°$, and $T·\Delta S°$ between pH 7.0 and 5.0 (2.2, 5.6, and 3.4 kcal/mol, respectively). Their C3·A70 tetraloop variant, however, embedded in a different nearest neighborship, (G-C-C)·(C-A-G), produced much smaller corresponding differences (1.9, 1.9, and 0.0 kcal/mol, respectively). The above data are consistent with previous results that N(1) of adenine in A·C pairs is protonated at pH 5.0−5.5 and unprotonated at pH 7.0−7.5.[9b,c,d,g,r,s] Presumably because of the difference of 0.5 pH unit under acidic conditions, Meroueh and Chow's respective ΔT_m values for their A3·C70 and C3·A70 variants are much higher than ours, 11.7° (A3·C70) and 13.4° (C3·A70), suggesting that protonation of adenine in A·C mismatches may not be complete in all sequence contexts at pH 5.5. All other tested mismatch variants showed melting profiles at pH 5.5 that were virtually superimposable when compared to those at pH 7.5 (not shown).

Pairing Thermodynamics. The general rank order of stabilities *in terms of* T_m, the most exactly measurable parameter, largely confirms the previous studies that analyzed internally mismatched RNA hairpins,[5a] RNA/DNA duplices,[10a,e,f] and short[10b,c] and long[10d] RNA/RNA duplices: Watson−Crick pairs > purine·pyrimidine mismatches (pur·pyr) > pur·pur > pyr·pyr with few exceptions. Among the pur·pyr combinations we see G·C > C·G > A·U ≈ U·A > C·A^+ ≈ G·U ≈ A^+·C > (U·G >)I·U > U·I(> U·G) > A·C ≈ C·A; thus, in the sequence context studied, (pur-pur-pur)·(pyr-pyr-pyr) tracts often appear more stable than the alternating (pur-pyr-pur)·(pyr-pur-pyr) "context" isomers, albeit with pH- and sequence-dependent differences. In the pur·pur variants we see G·G > G·A > A·G > A·A, in the pyr·pyr variants C·U > U·C > C·C ≈ U·U, with some differences only weak or insignificant. The order of hairpin stability as expressed in free energies of folding $\Delta G°_{25°C}$ roughly reflect the order of T_m values, however—due to varyingly steep transitions of the melting profiles[6]—with notable exceptions that will be addressed later (entries 1, 2, 6, 7, 14, 19, 20, and 26 in Table 1).

This order changes quite drastically when the variants are ranked according to their exothermicities of pairing as expressed by the calculated folding enthalpies $\Delta H°$ (Table 2). The most exothermic variants are not necessarily the most stable, fully Watson−Crick paired hairpins (see entries 1−12 and 26−31, Table 2). The entropic penalty of folding concomitantly compensates to subtly different degrees for the large exothermicities, which explains why the free energies of folding $\Delta G°_{25°C}$ are comparatively small (Table 1) and why the order of stabilities is different from that of exothermicities.

Changes in solvent composition or salt concentration have a much greater effect on exothermicity and entropic penalty than

(9) Crystal structure of C·A pairs in DNA: (a) Hunter, W.; Brown, T.; Anand, N. N.; Kennard, O. *Nature* **1986**, *320*, 552−555. NMR study on C·A pairs in DNA: (b) Wang, C.; Gao, H.; Gaffney, B. L.; Jones, R. A. *J. Am. Chem. Soc.* **1991**, *113*, 5486-5488. (c) Puglisi, J. D.; Wyatt, J. R.; Tinoco, I., Jr. *Biochemistry* **1990**, *29*, 4215−4226. (d) Kalnik, M. W.; Kouchakdjian, M.; Li, B. F. L.; Swann, P. F.; Patel, D. J. *Biochemistry* **1988**, *27*, 100−108. Crystal structure of C·U pairs in RNA: (e) Holbrook, S. R.; Cheong, C.; Tinoco, I., Jr.; Kim, S. H. *Nature*, **1991**, *353*, 578−581. (f) Cruse, W. B. T.; Saludjian, P.; Biała, E.; Strazewski, P.; Prangé, T.; Kennard, O. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4160−4164. NMR study on tandem C·A, C·C, C·U and other mispairs in RNA: (g) SantaLucia, J., Jr.; Kierzek, R.; Turner, D. H. *Biochemistry* **1991**, *30*, 8242−8251. Crystal structures of G·A pairs in DNA: (h) Kennard, O. *J. Biomol. Struct. Dyn.* **1985**, *3*, 205−26. (i) Prive, G. G.; Heinemann, U.; Chandrasegaran, S.; Kan, L. S.; Kopka, M. L.; Dickerson, R. E. *Science* **1987**, *238*, 498−504. (j) Brown, T.; Leonard, A.; Booth, E. D.; Chambers, J. *J. Mol. Biol.* **1989**, *207*, 455−457. (k) Webster, G. D.; Sanderson, M. R.; Skelly, J. V., Neidle, S.; Swann, P. F.; Li, B. F.; Tickle, I. J. *Proc. Natl. Acad. Sci. U S A*, **1990**, *87*, 6693−7. NMR study on G·A pairs in DNA: (l) Nikonowicz, E. P.; Meadows, R. P.; Fagan, P.; Gorenstein, D. G. *Biochemistry* **1991**, *30*, 1323−1334. (m) Li, Y.; Zon, G.; Wilson, W. D. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 26−30. (n) Katahira, M.; Sato, H.; Mishima, K.; Uesugi, S.; Fujii, S. *Nucleic Acids Res.* **1993**, *21*, 5418−5424. NMR study on G·A pairs in RNA: (o) Katahira, M.; Kanagawa, M.; Sato, H.; Uesugi, S.; Fujii, S.; Kohno, T.; Maeda, T. *Nucleic Acids Res.* **1994**, *22*, 2752−2759. pH dependence of pur·pur mispairs in RNA: (p) Morse, S. E.; Draper, D. E. *Nucleic Acids Res.* **1995**, *23*, 302−6. pH dependence of C·C coaxial dsRNA mismatches: (q) Kim, J.; Walter, A. E.; Turner, D. H. *Biochemistry* **1996**, *35*, 13753−61. Review on RNA mismatches: (r) Limmer, S. *Prog. Nucleic Acid Res. Mol. Biol.* **1997**, *57*, 1−39 and ref 10b. Thermodynamics of A·C mismatched DNA: (s) Allawi, H. T.; SantaLucia, J., Jr. *Biochemistry* **1998**, *37*, 9435−44.

(10) (a) Sugimoto, N.; Nakano, M.; Nakano, S. *Biochemistry* **2000**, *39*, 11270−81. (b) Kierzek, R.; Burkard, M. E.; Turner, D. H. *Biochemistry* **1999**, *38*, 14214−23. (c) Bevilacqua, J. M.; Bevilacqua, P. C. *Biochemistry* **1998**, *37*, 15877−84. (d) Zhu, J.; Wartell, R. M. *Biochemistry* **1997**, *36*, 15326−35. (e) Sugimoto, N.; Nakano, S.; Katoh, M.; Matsumura, A.; Nakamuta, H.; Ohmichi, T.; Yoneyama, M.; Sasaki, M. *Biochemistry* **1995**, *34*, 11211−6. (f) Freier, S. M.; Kierzek, R.; Jaeger, J. A.; Sugimoto, N.; Caruthers, M. H.; Neilson, T.; Turner, D. H. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 9373−7. (g) Xia, T.; SantaLucia, J., Jr.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X.; Cox, C.; Turner, D. H. *Biochemistry* **1998**, *37*, 14719−35.

Table 2. Enthalpy and Entropy of Folding of All 22mers and 25mers.

no.	3·70 nt, total nt	$\Delta H°$ [kcal/mol] (±4.5 kcal/mol)[a]	$\Delta S°$ [cal/(mol·K)] (±12 cal/(mol·K))[a]
1	U·I 22	−88.6	−248.6
2	C·A^+ context 22	−86.2	−238.6
3	G·U context 22	−84.8	−235.4
4	A·U 22	−81.3	−222.9
5	C·A^+ 22	−79.0	−219.8
6	I·U 22	−77.3	−216.4
7	A^+·C 22	−76.3	−212.1
8	U·A 22	−74.3	−204.1
9	A·U-G·U shift 22	−72.1	−204.1
10	C·A context 22	−71.8	−202.2
11	U·G 22	−71.8	−200.9
12	G·C-G·U shift 22	−71.2	−196.4
13	C·G 22	−69.9	−188.8
14	A·C 22	−69.6	−196.8
15	G·C 22	−66.7	−178.7
16	G·U 22	−64.0	−177.4
17	C·A 22	−61.0	−172.3
18	G·G 22	−61.0	−171.1
19	C·C 22	−61.0	−173.4
20	A·G 22	−60.0	−169.9
21	G·A 22	−59.1	−166.5
22	C·U 22	−57.8	−163.3
23	A·A 22	−54.3	−155.0
24	U·C 22	−53.9	−153.1
25	U·U 22	−52.6	−150.0
26	U·A 25	−84.4	−240.6
27	A·U 25	−81.7	−233.9
28	I·U 25	−79.5	−233.1
29	C·G 25	−79.4	−222.1
30	G·U context 25	−78.0	−227.1
31	G·U 25	−77.5	−224.0
32	U·G 25	−75.2	−219.0
33	C·U 25	−74.4	−222.3
34	C·A 25	−73.2	−217.8
35	U·U 25	−71.7	−215.0
36	A·C 25	−71.5	−212.2
37	G·C 25	−68.3	−190.5
38	G·A 25	−68.3	−202.2
39	C·C 25	−58.7	−176.4

[a] Approximate 90% confidence range. Denaturation profiles and confidence statistics: see Supporting Information.

they have on relative pairing stabilities. Four 3·70 variants of the 22mer tRNA acceptor hairpins, U3·I70, U3·G70, A3·G70, and U3·U70, were chosen for a systematic study in different solvent mixtures: (a) 1 M NaCl, (b) 0.1 M NaCl, (c) ethanol/0.1 M NaCl 1:9, (d) ethylene glycol/0.1 M NaCl 3:7, and (e) dimethylformamide/0.1 M NaCl 3:7, all aqueous parts buffered with 10 mM Na-phosphate to pH 7.5. In agreement with solvent-dependence studies on DNA[11a,b] and salt-dependence studies on RNA,[11c] we observe that salt stabilizes nucleobase pairing in terms of T_m, while organic aprotic polar additives destabilize more than protic additives (Table 3).

The variants show within ±0.8 kcal/mol and ±1.0° 90% confidence a well-distributed range of pairing stabilities in all solvent mixtures with varying free energy differences between the most and least stable hairpin: $\Delta\Delta G°_{25°C} = 3.5$ kcal/mol (a); 4.3 kcal/mol (e); 5.4 kcal/mol (d); 5.8 kcal/mol (c); 6.6 kcal/mol (b). The rank order with respect to the free energies $\Delta G°_{25°C}$ and exothermicities $\Delta H°$ are in all solvent mixtures as shown, U·I > U·G > A·G > U·U, with one exception in 10% ethanol, U·U > A·G (entries 11 and 12). The corresponding range of mid-transition temperatures ΔT_m does not quite reflect the varying relative stabilities $\Delta\Delta G°_{25°C}$ but stays between 5.5° and 7.8°: $\Delta T_m = 5.5°$ (e); 5.6° (b); 6.6° (d); 7.6° (c); 7.8° (a). The relative stabilities as expressed in T_m follow the general rule U·G > U·I > A·G > U·U with one exception in 0.1 M NaCl, U·I ≥ U·G (entries 5 and 6), and another in 30% dimethyl formamide, U·U ≥ A·G (entries 19 and 20).

Table 3. Solvent- and Salt-dependent Thermodynamics of Four Selected 22mers.

no.	3·70 nt, total nt	$\Delta H°$ [kcal/mol] (±4.5 kcal/mol)	$\Delta S°$ [cal/(mol·K)] (±12 cal/(mol·K))	T_m [°C] (±1.0°)	$\Delta G°_{25°C}$ [kcal/mol] (±0.8 kcal/mol)
1[a]	U·I 22	−74.0	−206.0	86.0	−12.6
2	U·G 22	−65.2	−179.7	89.7	−11.6
3	A·G 22	−60.1	−168.7	82.9	−9.8
4	U·U 22	−57.0	−160.6	81.9	−9.1
5[b]	U·I 22	−88.6	−248.6	83.3	−14.5
6	U·G 22	−71.5	−200.9	82.8	−11.6
7	A·G 22	−60.0	−169.9	79.9	−9.3
8	U·U 22	−52.6	−150.0	77.7	−7.9
9[c]	U·I 22	−78.3	−223.2	77.9	−11.8
10	U·G[22]	−71.7	−203.2	79.6	−11.1
11	A·G 22	−43.3	−125.2	72.9	−6.0
12	U·U 22	−55.1	−159.6	72.0	−7.5
13[d]	U·I 22	−80.4	−232.4	72.6	−11.1
14	U·G 22	−72.1	−207.5	74.6	−10.3
15	A·G 22	−54.9	−160.9	68.1	−6.9
16	U·U 22	−45.6	−133.7	68.0	−5.7
17[e]	U·I 22	−54.7	−161.3	65.6	−6.6
18	U·G 22	−62.5	−184.2	66.1	−7.6
19	A·G 22	−38.9	−117.0	59.4	−4.0
20	U·U 22	−31.1	− 93.1	60.6	−3.3

[a] Entries 1−4 in 1 M aqueous NaCl. [b] Entries 5−8 in 0.1 M aqueous NaCl. [c] Entries 9−12 in 10% (v/v) ethanol/0.1 M aqueous NaCl. [d] Entries 13−16 in 30% (v/v) ethylene glycol/0.1 M aqueous NaCl. [e] Entries 17−20 in 30% (v/v) dimethyl formamide/0.1 M aqueous NaCl, denaturation profiles at 285 nm (all others at 260 nm). All aqueous parts buffered with 10 mM Na-phosphate, pH 7.5.

The relative exothermicities and entropies of folding, $\Delta H°$ and $\Delta S°$, are apparently strongly influenced by the solvent composition, as would be expected from a dominance of (differential) solvent contributions to the overall thermodynamics as compared to contributions from the same set of solutes. The most pronounced $\Delta H°$ and $\Delta S°$ values were obtained in 0.1 M NaCl, our "standard solvent" (entries 5−8), the addition of protic and aprotic cosolvents to 0.1 M NaCl gradually lowered $\Delta H°$ and $\Delta S°$ (entries 9−20) without much changing the difference in thermodynamics among the four variants: $\Delta\Delta H° = 31-36$ kcal/mol, $\Delta\Delta S° = 91-99$ cal/(mol·K) (solvent systems b, c, d, e). However, a 10-fold increase in salt concentration reduced this difference to about one-half: $\Delta\Delta H° = 17.0$ kcal/mol, $\Delta\Delta S° = 45.4$ cal/(mol·K) (solvent system a). The wobble pairs U·I and U·G appear to be the mispairs that are most sensitive to changes in monovalent salt concentration. They are the only variants among the studied four that, with the caveat of unequal significance, exhibit a weaker exothermicity under high salt conditions.

Discussion

Stabilizing Protons. Consistent with a large number of investigations, A·C mispairs belong to the weak mismatches unless they are stabilized by a proton at pH 5.5 or below. The protonated form $[A·C]^+$ stabilizes the double-helix at least as

(11) (a) Breslauer, K. J.; Bodnar, C. M.; McCarthy, J. E. *Biophys. Chem.* **1978**, *9*, 71−78, and cited references therein. (b) DePrisco Albergo, D.; Turner, D. H. *Biochemistry* **1981**, *20*, 1413−18. (c) Williams, D. J.; Hall, K. *Biochemistry* **1996**, *35*, 14665−70.

well as the leader of mispairs G·U does (Table 1, entries 6−10 and 12). A comparison between the variants C3·A70 context and A3·C70 bearing the same nearest neighbors shows that the former hairpin folds more exothermically (Table 2, entries 10 and 14), especially when protonated (Table 2, entries 2 and 7), and may be more stable (Table 1, entries 6/9 and 15/19) than the latter. Apparently, a mispaired adenine within an A-RNA conformation prefers to be placed in a more purine-rich and/or a longer strand, which clearly confirms that sequence context effects may extend over more than the merely nearest neighorship interactions.[10b]

Cytosines in single C·U or C·C mismatches within an otherwise fully Watson−Crick-paired A-RNA environment are not sufficiently basic to be stabilized by a proton at pH 5.5. They remain relatively unstable because—unlike tandem repeats of $C\cdot C^+$ pairs,[9g] $C\cdot C^+$ pairs between single-stranded overhangs coaxially stacked to each other,[9q] or poly$(C\cdot C)^+$ tracts[12]—narrowing the base−base distance needed for protonation of only one C·C mismatch in an A-RNA environment is too costly; an intrahelical bridging water molecule must be assumed instead.[9e,f,g]

Isolated A·A, G·G, and G·A mispairs in dsRNA 9mers in an alternating (GC)N(GCGC) A-RNA environment were shown to be only weakly or insensitive to the lowering of the pH from 7.0 to 5.0,[9p] so are the pur·pur mispairs in the less alternating (GG)N(GCU) neighborship within a tRNA acceptor stem.[5a] G·A mispairs are known to adopt alternative pairing modes without proton stabilization: G(anti)·A(syn) Hoogsteen, G(anti)·A(anti) Watson−Crick-like, or sheared G(anti)·A(anti).[9h−o] The G(anti)·A(anti) mispairs seem unlikely because they have so far been found only within consecutive G·A double mismatches. However, a dynamic equilibrium between the neutral forms exhibiting indistinguishable thermodynamics cannot be excluded.

Stability Predictions. The general hierarchy of RNA mismatch stabilities to date does not change much with this investigation. Some missing information on the thermodynamics of certain RNA/RNA mispairs in particular sequence contexts may help to further fine-tune the RNA secondary structure prediction algorithm MFOLD.[13] The current version 3.1 (based on data from 1 M NaCl, pH 7.0 conditions) overestimates the free energies of folding at 37 °C of our 22mers by an average of 4.8 kcal/mol {min. 0.1 kcal/mol (U3·I70)/max. 7.4 kcal/mol (G3·C70)} and of our 25mers by an average of 1.3 kcal/mol {min. −1.5 kcal/mol (I3·U70)/max. 3.6 kcal/mol (G3·C70)}, after a sequence-independent salt correction[11c] of − 1.0 kcal/mol had been applied, by an average 3.8 and 0.3 kcal/mol, respectively. If, in addition, MFOLD's global extra bonus for UUCG hairpins of 3.0 kcal/mol, which takes into account "extra stability and potential tertiary interactions", is reduced to 2.1 kcal/mol, a global UUCG bonus value ascribed to "extra stability only", an average 2.9 kcal/mol overestimation of the stability of our 22mers results. The relative order of variant hairpin stabilities in both loop frameworks matches in large part the ones depicted in Table 1, i.e., according to the measured T_m values. On a relative scale, the most severely misestimated free energies are the I3·U70 and U3·I70 variants, which were calculated with the least deviations from our experimental values (1.4 to −1.0 kcal/mol without salt correction), but also calculated as least stable of all (probably due to the lack of literature data). Likewise, the G3·A70 (not A3·G70) variants tend to show up too close to the pyr·pyr mismatches; we measured stabilities of G3·A70 closer to the ones of C3·A70 and A3·C70 (see Supporting Information for details).

In short, the average *absolute* hairpin stabilities predicted by MFOLD v3.1 seem to be satisfactory for "ordinary loops". The global UUCG bonus should be stem sequence-dependently reduced so as to lower the stabilities of UUCG tetralooped hairpins and, among those, the all-Watson−Crick variants most. The *relative* stability differences are predicted—within the experimental uncertainties—correctly for most single internal mismatches except for I·U (U·I) wobble pairs and perhaps (pur-G-pur)·(pyr-A-pyr) sequences.

Effects of Hydration. The solvent- and salt-dependence results depicted in Table 3 suggest that changes in folding enthalpy and entropy are strongly influenced and may even be clearly dominated by changes in the hydration layer of the solutes. Volumetric and acoustic measurements on double-stranded B-DNA solutions lead to the conclusion that, in "mixed" sequences, the "first hydration layer" or 24—at AT- or GC-rich regions up to 37—hydration molecules *per nucleotide* could be approved "volumetrically effective".[14a,b] Acoustic measurements detected, as judged by higher apparent molar adiabatic compressibilities, a slightly lower hydration of double-stranded A-RNA,[14c] as expected.[15] From a minute analysis of calorimatric data on thermal nucleic acid unfolding, Breslauer and co-workers[14d] deduced "that on average about 70% of the transition enthalpy results from duplex nonspecific general interactions."

The number of favorable hydration water hydrogen bonds formed upon RNA folding, most likely a major producer of overall folding exothermicity[16a] at the cost of less favorable water-RNA contacts, translates into an upper limit of hydration molecules that could be termed "calorimetrically effective". The exothermicities measured here, 53−88 kcal/mol (Table 2), equal approximately 10−22 extra hydrogen bonds formed (13−16 hydrogen bonds for the all-Watson−Crick hairpins, $\Delta H^\circ \approx 5$ kcal/mol per hydrogen bond) and, thus, correspond to the net formation of roughly 5−11 tightly bound hydration waters *per folded hairpin*. Confusingly, this number of extra hydrogen bonds is approximately the same as the number, albeit not necessarily strength, of hydrogen bonds formed between the paired bases. Yet, in the denatured state the number of hydrogen bonds between RNA bases and water must be at least the same if not more.[16a,b] The small and easily adjustable hydration water molecules do not contribute directly to the overall stability of solute folding because the exothermicity of solute hydration is fully compensated by the inevitable entropic penalty of solute

(12) (a) Moyzis, R.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3874−3878. (b) Kanaori, K.; Shibayama, N.; Gohda, K.; Tajima, K.; Makino, K. *Nucleic Acids Res.* **2001**, *29*, 831−840.
(13) (a) Zuker, M.; Mathews, D. H.; Turner, D. H. *RNA Biochemistry and Biotechnology*; Barciszewski, J.; Clark, B. F. C., Eds.; NATO ASI Series, Kluwer Academic Publishers: 1999; pp 11−43. (b) Mathews, D. H.; Sabina, J.; Zuker, M.; Turner, D. H. *J. Mol. Biol*. **1999**, *288*, 911−940. (c) http://bioinfo.math.rpi.edu/%7Ezukerm/rna.
(14) (a) Chalikian, T. V.; Plum, G. E.; Sarvazyan, A. P.; Breslauer, K. J. *Biochemistry* **1994**, *33*, 8629−40. (b) Chalikian, T. V.; Sarvazyan, A. P.; Breslauer, K. J. *Biophys. Chem.* **1994**, *51*, 89−109. (c) Chalikian, T. V.; Völker, J.; Srinivasan, A. R.; Olsson, W. K.; Breslauer, K. J. *Biopolymers* **1999**, *50*, 459−471. (d) Chalikian, T. V.; Völker, J.; Plum, G. E.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 7853−8.
(15) Saenger, W.; Hunter, W. N.; Kennard, O. *Nature* **1986**, *324*, 385−8.
(16) (a) Cooper, A. *Biophys. Chem.* **2000**, *85*, 25−39. (b) Williams, D. H. *Aldrichimica Acta* **1991**, *24*, 71−80, see also addendum by Williams, D. H. *Aldrichimica Acta* **1992**, *25*, 9.

hydration.[17] Therefore, a major part, if not all, of the sum of net favorable interactions, dipole−dipole, hydrophobic, hydrogen, and salt bridges, that originate from the covalent structure of the folded solute manifests itself not in $\Delta H°$ but only in the free energy of folding $\Delta G°_T$.

The sensitivity of A-RNA to differential hydration shows in the solvent dependence of the folding thermodynamics of four selected tRNA acceptor hairpin 22mers (Table 3). Through addition of an increasingly stronger denaturant, the exothermicity of folding of a subset of hairpins diminishes, as judged by the difference between the extremes (mostly U3·I70 versus U3·U70), by more or less the same 30 to 35 kcal/mol, an amount worth the net formation of six to seven hydrogen bonds per hairpin. The formation of a highly hydrated[18] U·I wobble pair releases enthalpy (and costs entropy), in the absence or presence of destabilizing protic cosolvent, equal to the net formation of two or three hydrogen bonds more than of an isosteric U·G wobble base pair bearing an additional unpaired 2-amino group in the shallow groove. Nevertheless, the U3·G70 variant exhibits slightly higher transition temperatures than the U3·I70 variant under most of the tested conditions.

More comparisons between variants (in 0.1 M NaCl) differing in one or two atomic groups within the same pairing mode, G3·C70 versus A3·U70 or C3·G70 versus U3·A70 (Watson−Crick) and G3·U70 versus I3·U70 or U3·G70 versus U3·I70 (wobble), show that the presence of a "protruding group" in the shallow groove of A-RNA, the 2-amino group of guanine, stabilizes the folded structure at elevated temperatures (higher T_m) whether or not it "pairs" with its vis-à-vis (G·C > A·U, G·U > I·U, etc., Table 1). At increasingly lower temperatures this difference gradually diminishes and eventually inverts (crossing melting profiles)[6] in favor of the "weaker" base pairs, A·U, I·U, that are more hydrated in the shallow groove and thus form more exothermically (Table 2). We can ascribe this relative stabilization of guanine-containing base pairs at higher and destabilization at lower temperatures to the perturbation of the first hydration layer by the 2-amino group, which shows in the relatively low exothermicities of the corresponding variants. At high temperatures the hydration is less pronounced; hence, the entropic penalty (of hydration) cannot effectively compensate despite the relatively weak exothermicity, which raises T_m. At lower temperatures, where the entropic penalty generally weighs less, an extensive and tight hydration of the so-called weak base pairs becomes a more favorable, stabilizing feature (more negative $\Delta G°_T$).

Other studies confirm that A·U base pairs in A-RNA stems may fold more exothermically than G·C pairs,[5a,10c] again pointing at the decisive influence of the hydration layer on folding exothermicity. Another indication are proton-stabilized A·C-containing nucleic acids that fold more exothermically when protonated by an average 13 ± 6 kcal/mol per A·C mismatch in RNA (in 0.1 M NaCl at pH 7.5 versus 5.5, this work) and by an average 9 ± 4 kcal/mol per A·C mismatch in DNA (in 1 M NaCl at pH 7.0 versus 5.0).[9s] Protons are a paradigm of tight hydration, although additional exothermicity from Coulomb attraction of the opposing strands by A^+·C under low salt conditions cannot be excluded. More generally, dramatic or unexpectedly large changes in $\Delta H°$ and $\Delta S°$ between only slightly different singly mismatched duplices, hairpins, or aqueous solutions of slightly differing complexes of, most likely, any kind originate from extended differential effects in their hydration shell rather than in their covalent structure.[19]

Conclusion

The analysis of the *shapes* of optically derived melting curves, in particular, the steepness of thermal transitions, translates into an analysis of exothermicities and entropic penalties. Although they are accompanied by relatively large errors (in $\Delta H°$ and $\Delta S°$), the interdependence of these errors together with a fairly large amount of datapoints on structurally similar compounds (sequence variants, mutants) and the usage of the same analytical precision for the data analysis of each member of a subgroup (curve fitting, baseline corrections, etc.) result in statistically relevant information on the stability, hydration (this article), and perturbation sensitivity (next article) of nucleic acid higher-order structures.

This study shows that (i) among all potentially more basic internal single-mismatches in A-RNA only the A·C-variants can be stabilized by protonation at pH 5.5, (ii) this stabilization is accompanied by a rather high exothermicity, suggesting that tight hydration is characteristic of highly exothermic pairing systems (protons are highly hydrated); (iii*)* under "close to physiological" conditions (100 mM NaCl, 10 mM Na-phosphate, pH 7.5) the "weaker" A·U and I·U pairs form more exothermically than the "stronger" ones, G·C and G·U, which again reflects the tighter hydration of the former. This difference in the means of gaining stability—relatively high exothermicity and high entropic penalty of highly hydrated base pairs (steeper transitions in the melting profiles) versus relatively low entropic penalty and low exothermicity of typically high-melting base pairs (shallower transitions in the melting profiles)—reduces, in general, the difference between weak and strong pairing regions in A-RNA at ambient temperatures: at lower temperatures C·G/G·C pairs are similar to U·A/A·U pairs while at higher temperatures they are more stable, which may have an important bearing on the evolution of thermophilic organisms exhibiting distinct G + C contents.

Acknowledgment. We thank Prof. William McClain and Dr. Kay Gabriel, Department of Bacteriology, University Wisconsin—Madison, for their collaboration and help. The financial support by the Swiss National Science Foundation and the Novartis Foundation is gratefully acknowledged.

Supporting Information Available: Details of RNA synthesis (synthetic protocol, deprotection, HPLC purification, desalting, storage) and analyses (HPLC, CE, MALDI-TOF MS, phosphodiesterase I/alkaline phosphatase degradation), UV measurements (pH 5.5 and 7.5, concentration dependence), data collection and curve fitting procedures (illustrated in color plots), fitting parameters (including pH, concentration and solvent dependence), error statistics, and detailed comparisons with MFOLD-calculated stabilities (PDF file). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0161305

(17) (a) Grunwald, E.; Steel, C. *J. Am. Chem. Soc.* **1995**, *117*, 5687−92. (b) Lee, B. *Biophys. Chem.* **1994**, *51*, 271−78, and cited references therein.
(18) Cruse, W. B. T.; Aymani, J.; Kennard, O.; Brown, T.; Jack, A. G. C.; Leonard, G. A. *Nucleic Acids Res.* **1989**, *17*, 55−72.
(19) Williams, D. H.; O'Brien, D. P.; Bardsley, B. *J. Am. Chem. Soc.* **2001**, *123*, 737−38.