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

Thermodynamic Correlation Analysis: Hydration and Perturbation Sensitivity of RNA Secondary Structures

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 thermodynamic parameters ΔH, ΔS, T_m, and ΔG of a total of 36 RNA strands, 22 tetralooped 22mers, and 14 heptalooped 25mers (same stem sequence) were analyzed with respect to enthalpy−entropy compensation (EEC). The EEC plots $\{\Delta H, \Delta S\}$ were compared with collected literature data from protein and nucleic acid unfolding studies (3224 and 241 datapoints, respectively) which all proved to be remarkably linear. The similarity of the compensation slopes and intercepts for all compounds indicate that, irrespective of the chemical nature and stability of the folding solutes, the exothermicity ΔH and entropic penalty $T{\cdot}\Delta S$ 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 (ΔG) and ratio (T_m) of both parameters. EEC plots $\{\Delta H, \Delta S\}$ are presented in an extended context, as 3D plots $\{\Delta H, \Delta S, T_m\}$ allowing for a correct analytical description of the enthalpy−entropy relationship and for more practical interpretations of large amounts of thermodynamic data when replotted as $\{\Delta H, T_m\}$ or $\{\Delta G_T, T_m\}$. The introduction of a variety of mismatches into nucleic acids, or limited irreguliarities into any supramolecular complex, and the analysis of the involved thermodynamics as shown in this study—i.e., scanning the "enthalpy−entropy space" of whole macromolecular subgroups—should permit to extract and quantify more "hidden information", such as hydration extent and sensitivity of macromolecular frameworks toward desolvation and structural perturbation, from thermodynamic analyses of large sample sizes.

Introduction

The preceding article presents the thermodynamic parameters of the folding of two RNA hairpins bearing the same set of helical stem sequences but closed by two different loops, a tetra- and a heptaloop, under a variety of conditions (different internal mismatches, pH, salt, and cosolvents). Here, an extensive correlation analysis of the datapoints $\{\Delta H, \Delta S\}$, $\{\Delta H, T_m\}$, $\{\Delta S, T_m\}$, and $\{\Delta G_T, T_m\}$ is put into context with a large amount of protein and nucleic acid unfolding thermodynamics and suggests a useful way of interpretation. The most indicative measure for the stability of folding is the T_m value, T_m(unimolecular) = $T_{\Delta G°=0} = \Delta H/\Delta S$. The extent or strength of solvation of the folded structures finds its signature in the accompanying exothermicity ΔH. The sensitivity of a family or subgroup of folded structures toward hydration layer-changing elements (such as mismatches) and possibly events (such as ligand binding) shows in the relation between the two, ΔH versus T_m, and their sensitivity toward destabilizing elements/events in general is best decribed by ΔG_T versus T_m.

Enthalpy−Entropy Compensation

It is a generally observed phenomenon that the exothermicity of folding ΔH provides a sensitive means of measuring the effect of small changes introduced into higher-order structures, such as internal mismatches into RNA hairpins, and that the entropic penalty of folding $T{\cdot}\Delta S$ concomitantly compensates to subtly different degrees for the often large exothermicities. This explains why the free energies of folding $\Delta G_{25°C}$ are comparatively small and why the order of stabilities is different from that of exothermicities. The enthalpy−entropy compensation (EEC) effect is well-known from a large body of data including calorimetric measurements of protein unfolding, optical detection of thermal denaturation of diverse nucleic acid systems, ligand binding studies and supramolecular associations in general, solvent transfer, enzyme kinetics, heterogeneous catalysis, heats of sublimation, etc.[1] Recently, a large number of calorimetrically determined protein denaturation thermodynamics were analyzed with respect to EEC (Figure 1).[1m]

The linear correlation between the denaturation enthalpies and entropies of proteins,

$$\Delta H = g + x_T{\cdot}T{\cdot}\Delta S \quad \text{or} \quad \Delta S = (\Delta H - g)/x_T{\cdot}T \qquad (1)$$

is most remarkable and can be roughly described by an intercept $g = -3.6 \pm 0.4$ kcal/mol and a slope $x_T{\cdot}T = 325 \pm 6$ K (90% confidence range). In other words, the enthalpic and entropic contributions to protein denaturation cancel each other out on the average at some 46 to 58 °C. A similar compilation of thermal denaturation parameters for nucleic acids (Figure 1, gray dots) including tRNA acceptor hairpins,[2] many other DNA and RNA sequences,[3] and a number of nonfuranosyl nucleic acids,[3h]

* E-mail: peter.strazewski@unibas.ch.
† Present address: Laboratoire de Synthèse de Biomolécules, Bâtiment E. Chevreul (5ème étage), Universite Claude Bernard-Lyon 1, 43 boulevard du 11 novembre 1918, F-69622 Villeurbanne (Cedex), France.

10.1021/ja016131x CCC: $22.00

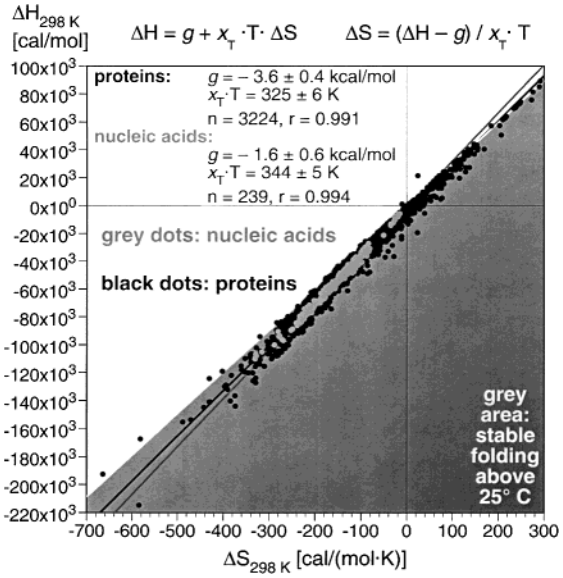

Figure 1. Folding thermodynamics of proteins and nucleic acids. Calorimetrically derived protein unfolding data ($n = 3224$ datapoints $\{\Delta H, \Delta S\}$) have been converted to $T = 298$ K assuming a temperature-independent change in heat capacity ΔC_p (eqs 8, 9) and were made available by Liu et al.[1m] These data are depicted as folding parameters in [cal/mol]. The nucleic acid thermodynamics mostly originate from optically derived temperature-dependent equilibrium shift profiles ("melting curves"), usually recorded at 260 nm, leading to folding enthalpies and entropies with no heat capacity corrections ($n = 241$ datapoints, 3 calorimetric measurements). Since the relative errors in ΔH and ΔS are similar and interdependent in both datasets,[1b,8] which translates into comparatively small errors in $\Delta G°$ (Tables 1 and 2 in ref 2b), each set of datapoints was analyzed by means of two linear regressions, once by $\Delta H = f(\Delta S)$ and once by $\Delta S = f(\Delta H)$. The given mean values and 90% confidence ranges result from a conservative interpretation of the double regression (arithmetic mean ± largest upper or lower 90% error range). The area of increasingly exergonic free energy of folding, $\Delta G_{25°C} \leq 0$, is shaded in gray.

quite nicely matches the protein data (Figure 1, black dots). The linear correlation may be a little different from the protein data and follows $g = -1.6 \pm 0.6$ kcal/mol and $x_T{\cdot}T = 344 \pm 5$ K (90% confidence range). Thus, the enthalpic and entropic contributions to nucleic acid denaturation appear to cancel each other out on the average at higher temperatures of somewhere between 66 and 76 °C.

The physical origin of the EEC, be it in kinetics or thermodynamics, is a question discussed since it was discovered.[1] In the case of a solution of molecules that can fold into higher-order structures or form host−guest complexes, the relatively weak interactions between solutes and between solute and solvent apparently result in a linear relationship between ΔH and ΔS, whereas EEC of gas-phase association or sublimation thermodynamics may deviate from linearity.[1g,h,n] Two (out of more) particular questions arise. One asks for the physical origin of compensation effects between enthalpy and entropy changes, i.e., of the fact that in all systems where denaturation, for instance, was measured exothermicity and entropic penalty of folding appear *positively* coupled to each other: the smaller the exothermicity the smaller the entropic penalty and vice versa.[4] It has been addressed many times and the explanations all follow the line that, in structurally adjustable systems (be it solute or solvent), favorable enthalpic interactions always cost degrees of freedom (of translation, rotation, vibration). The additivity of incremental effects results in a linear relationship. In very rigid systems, entropic freezing is limited so that eventually at exceptionally high exo- or endothermicities a deviation from linearity in EEC plots is expected.

The other question asks for the physical origin of a *common* linear EEC, i.e., a linear dependence of unfolding enthalpy and entropy, ΔH_{un} and ΔS_{un}, that is common to "all" solutes, irrespective of their structure, size, charge distribution, chemical functionalities, etc. On the basis of the pioneering work of Grunwald and Steel[1i] who developed a formalism of "compositional thermodynamics", the analysis of Liu et al.[1m] dissects the various contributions to the overall denaturation thermodynamics into changes of solvated solute, N(solv), U(solv) (N = folded solute, U = denatured, unfolded solute), solute-bound solvent, solv(N), solv(U), and "solvent-bound solvent", e.g., bulk water interactions, solv(solv). Solvent rearrangement processes involving solv(N), solv(U) ⇌ solv(solv) equilibria are, in the case of double-stranded B-DNA, thought to be effected by one or more "hydration shell layers" with 24−45 hydration molecules per nucleotide[5] and exhibit, irrespective of the type of solute or solvent, fully compensating thermodynamics:[1i,o,6]

$$H_{solv(N)} = T{\cdot}S_{solv(N)} \quad (2a)$$

$$H_{solv(U)} = T{\cdot}S_{solv(U)} \quad (2b)$$

$$H_{solv(solv)} = T{\cdot}S_{solv(solv)} \quad (2c)$$

It can be derived from first thermodynamic principles that, while the overall free energy of (un)folding results from the

(1) A vast amount of studies addressing EEC and the isoequilibrium and isokinetic relationship can be found in the literature. Here we would like to point out only a few—biased and by no means comprehensive—where many of the cited references allow for an in-depth study of the phenomenon: (a) Lumry, R.; Rajender, S. *Biopolymers* **1970**, *9*, 1125−227. (b) Exner, O. *Progress in Physical Organic Chemistry*, vol. 10; Streithwieser, A., Jr., Taft R. W., Eds.; New York, 1973; pp 411−482. (c) Jencks, W. P. *Adv. Enzymol. Relat. Mol. Biol.* **1975**, *43*, 219−410. (d) Williams, D. H. *Aldrichimica Acta* **1991**, *24*, 71−80. (e) Searle, M. S.; Williams, D. H. *J. Am. Chem. Soc.* **1992**, *114*, 10690−97. (f) Searle, M. S.; Williams, D. H. *Nucleic Acids Res.* **1993**, *21*, 2051−56. (g) Searle, M. S.; Westwell, M. S.; Williams, D. H. *J. Chem. Soc., Perkin Trans. 2* **1995**, 141−151. (h) Dunitz, J. D. *Chem. Biol.* **1995**, *2*, 709−12. (i) Grunwald, E.; Steel, C. *J. Am. Chem. Soc.* **1995**, *117*, 5687−92. (j) Serra, M. A.; Nissen, P. *FASEB J.* **1999**, *13*, A1384. (k) Cooper, A. *Curr. Op. Chem. Biol.* **1999**, *3*, 557−63. (l) Rekharsky, M.; Inoue, Y. *J. Am. Chem. Soc.* **2000**, *122*, 4418−35. (m) Liu, L.; Yang, C.; Guo, Q.-X. *Biophys. Chem.* **2000**, *84*, 239−251. (n) Williams, D. H.; O'Brien, D. P.; Bardsley, B. *J. Am. Chem. Soc.* **2001**, *123*, 737−38. (o) Liu, L.; Guo, Q.-X. *Chem. Rev.* **2001**, *101*, 673−696. (p) Liu, L.; Yang, C.; Guo, Q.-X. *Bull. Chem. Soc. Jpn.* **2001**, *74*, 2311. EEC should not be treated as synonymous to the isoequilibrium or isokinetic relationship but is formally related to it, see ref 1o.

(2) (a) Meroueh, M.; Chow, C. S. *Nucleic Acids Res.* **1999**, *27*, 1118−25. (b) Biała, E.; Strazewski, P. *J. Am. Chem. Soc.* **2002**, *124*, 3540−3545, preceding article in this issue.

(3) (a) 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. (b) Antao, V. P.; Tinoco, I., Jr. *Nucleic Acids Res.* **1992**, *20*, 819−24. (c) Filimonov, V. V.; Breslauer, K. J. *Thermodynamic Data for Biochemistry and Biotechnology*; Hinz, H.-J., Ed.; Springer, Berlin, 1986; pp 377, and pp 402, respectively. (d) Morse, S. E.; Draper, D. E. *Nucleic Acids Res.* **1995**, *23*, 302−6. (e) Bevilacqua, J. M.; Bevilacqua, P. C. *Biochemistry* **1998**, *37*, 15877−84. (f) 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. (g) Kierzek, R.; Burkard, M. E.; Turner, D. H. *Biochemistry* **1999**, *38*, 14214−23. (h) Micura, R.; Kudick, R.; Pitsch, S.; Eschenmoser, A. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 680−82.

(4) *Negative* coupling of exothermicity and entropic penalty in DNA-binding ligands: (a) Breslauer, K. J.; Remeta, D. P.; Chou, W.-Y.; Ferrante, R.; Curry, J.; Zaunczkowski, D.; Snyder, J. G.; Marky, L. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8922−26, and in solvation and ligand binding of small organic molecules: (b) Gallicchio, E.; Kubo, M. M.; Levy, R. M. *J. Am. Chem. Soc.* **1998**, *120*, 4526−27.

(5) (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.

(6) Lee, B. *Biophys. Chem.* **1994**, *51*, 271−78 and cited references therein.

difference of free energies between solvated folded solute and solvated denatured solute, $\Delta G_{un} = G_{N(solv)} - G_{U(solv)}$, with a *zero free energy contribution from solvent rearrangement processes*, $G_{solv(N)} = G_{solv(U)} = G_{solv(solv)}$, the largely compensating overall folding enthalpies and entropies, ΔH_{un} and ΔS_{un}, are apparently *strongly dominated by* the large enthalpic and entropic *fully compensating contributions from solvent rearrangement processes*. A general phenomenon seems to be that the enthalpic and entropic *contributions from conformational changes of macromolecular solutes are much smaller*:

$$|H_{N(solv)} - H_{U(solv)}| \ll |H_{solv(N)} - H_{solv(U)}| \quad (3)$$

$$|S_{N(solv)} - S_{U(solv)}| \ll |S_{solv(N)} - S_{solv(U)}| \quad (4)$$

The fact that the observed exothermicities and entropy changes of all thermal or other denaturation experiments largely originate from rearrangement processes of solvent molecules—usually present in much higher molar concentrations than macromolecular solutes—clearly necessitates for the same solvent a common EEC, i.e., a representative intercept g and slope $x_T \cdot T$ as shown in Figure 1:

$$\Delta H_{un} - g = x_T \cdot T \cdot \Delta S_{un} \quad (5)$$

Any group of solutes capable of stable folding must exhibit compensation slopes higher than the temperature they function: $x_T > 1.0$. At room temperature ($T = 298$ K), for instance, the favorable folding enthalpies in average proteins ($x_{298K} = 1.09$) and nucleic acids ($x_{298K} = 1.15$) generally outweigh their entropic folding penalties by roughly 9 to 15%, irrespective of the actual folding exothermicity, and hence, irrespective of the amount of solvent molecules involved in solute-induced rearrangements and irrespective of the size and nature of the solute.

However, overall EEC's can only be approximate. The smaller, not fully compensating enthalpic and entropic contributions from changes in solute conformation are in principle different from the ones of solvent rearrangement and do show up in the variations of g and $x_T \cdot T$. When analyzing different types of compounds separately such as nucleic acids and proteins as shown in Figure 1 or, even more so, when we analyze subgroups of one family of compounds such as groups of similar nucleic acids, as shown in Figure 2, the subtle differences that originate from differences in loops of hairpins or in the backbones of the nucleic acids manifest themselves in different EEC's. Subgroups of nucleic acids are naturally smaller in number but they often show a more defined and clearer linear relationship with r factors of up to 0.99996 (captions to Figures 2 and 3). The statistically most significant differences appear to be the slopes $x_T \cdot T$, i.e., the temperatures at which the overall unfolding enthalpic contributions of a subgroup precisely offset the entropic ones or, alternatively, the enthalpy−entropy ratio x_T at a given temperature, both irrespective of the actual exothermicity. We observe compensation slopes $x_T \cdot T$ of up to 413 K ($x_{298K} = 1.39$). At very low folding exothermicities of the solute, $\Delta H \rightarrow 0$, $\Delta S \rightarrow 0$, we calculate compensation intercepts of subgroups of nucleic acids between $g = +8$ to -8 kcal/mol (Figure 2).

Two subgroups are of a particular interest, since they are related to relatively small and systematic structural changes in the nucleic acids and to compositional changes of the solvent. The 3·70 variants of the 22mer and the 25mer tRNA acceptor hairpins[2b] differ from each other and from the closely related 16 and 24mer tRNA acceptor hairpins analyzed by Meroueh and Chow[2a] in the loop-closing base pair, the loop sequence, and loop length. The thermodynamics refer all to thermal denaturation in a 0.1 M aqueous NaCl solution. Despite the mid-transition temperatures T_m covering a quite large range of between 55.7 and 100 °C, which reflects the effect of rather different base pairing stabilities at position 3·70, the compensation intercepts g are all in the positive range and the compensa-

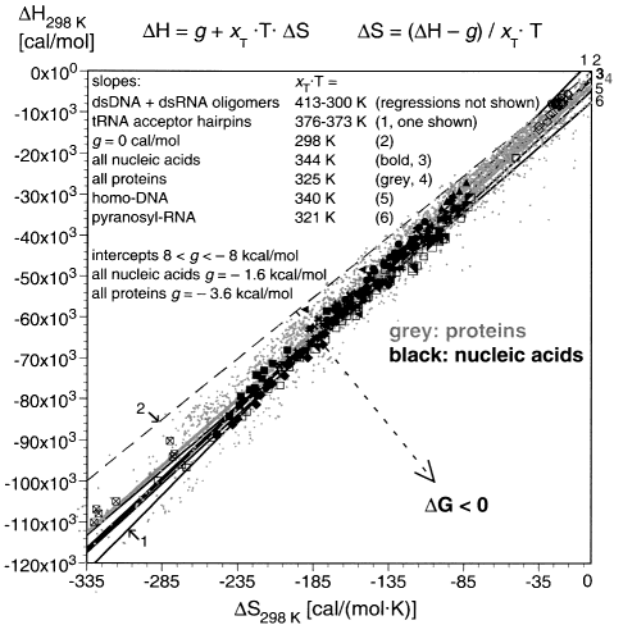

Figure 2. Folding thermodynamics of various nucleic acids (gray background data: proteins from ref 1m). The line of zero free energy of folding at room temperature, $\Delta G_{25°C} = 0$, is marked as dashed line 2 ($g = 0$, $x_T \cdot T = 298$ K, $x_{298K} = 1.0$); perpendicular to this line (dashed arrow) $\Delta G_{25°C}$ becomes increasingly exergonic. Folding enthalpies/entropies $\Delta H/\Delta S$ and, if $r \geq 0.96$, linear regression compensation parameters (g, $x_T \cdot T$, n = sample size, r = regression coefficient; regressions and confidence range: see caption to Figure 1) of (◇) 5′-terminal dinucleotide stacking in dsRNA in 1 M NaCl[3a] ($g = 0.7 \pm 1.8$ kcal/mol, $x_T \cdot T = 413 \pm 70$ K, $n = 10$, $r = 0.986$); (Δ) 5′-terminally mismatched dsRNA 7mers and 8mers in 0.2 M NaCl[3c] ($g = -6.2 \pm 1.9$ kcal/mol, $x_T \cdot T = 288 \pm 18$ K, $n = 4$, $r = 0.994$); (○) average base pair stack in polyDNA and polyRNA pH 6.9−7.5[3c] ($n = 6$, $r = 0.959$); (⊙) average base pair stack in bulk DNA from diverse organisms[3c] ($n = 10$, $r = 0.83$); (▲) RNA and DNA 12mer hairpins (4 bp + tetraloop) in 1 M NaCl[3b] ($g = 4.0 \pm 0.3$ kcal/mol, $x_T \cdot T = 374 \pm 3$ K, $n = 16$, $r = 0.9996$); (#) internally mismatched dsRNA 7mers in 1 M NaCl[3g] (from curve fit parameters: $g = -2.3 \pm 1.5$ kcal/mol, $x_T \cdot T = 331 \pm 9$ K, $n = 7$, $r = 0.997$; from $1/T_m$ plots (not depicted): $g = -0.8 \pm 1.5$ kcal/mol, $x_T \cdot T = 367 \pm 10$ K, $n = 7$, $r = 0.996$); (●) 16mer and 24mer tRNA acceptor hairpins in 0.1 M NaCl[2a] ($g = 4.2 \pm 3.8$ kcal/mol, $x_T \cdot T = 374 \pm 17$ K, $n = 22$, $r = 0.968$); (◆) 22mer tRNA acceptor hairpins in 0.1 M NaCl[2b] ($g = 3.0 \pm 2.7$ kcal/mol, $x_T \cdot T = 373 \pm 14$ K, $n = 25$, $r = 0.995$); (■) 25mer tRNA acceptor hairpins in 0.1 M NaCl[2b] ($g = 7.3 \pm 12.0$ kcal/mol, $x_T \cdot T = 377 \pm 70$ K, $n = 14$, $r = 0.963$); (solid triangle pointing to the left) self-complementary dsDNA 6mers and 8mers in 1 M NaCl (calorimetric measurement)[3c] ($n = 3$, $r = 0.46$); (solid triangle pointing to the right) self-complementary dsDNA 4mers and 5mers in 1 M NaCl[3c] ($g = 3.5 \pm 2.2$ kcal/mol, $x_T \cdot T = 395 \pm 10$ K, $n = 14$, $r = 0.993$); (▼) non self-complementary dsRNA 7mers to 9mers in 1 M NaCl[3c] ($g = -1.5 \pm 1.0$ kcal/mol, $x_T \cdot T = 329 \pm 5$ K, $n = 3$, $r = 0.9998$); (◩) ds homo-DNA 6mers and 8mers in 0.15 M NaCl[3h] ($g = -4.4 \pm 1.5$ kcal/mol, $x_T \cdot T = 340 \pm 10$ K, $n = 18$, $r = 0.989$); (◹) ds pyranosyl-RNA 6mers and 8mers in 0.15 M NaCl[3h] ($g = -6.8 \pm 1.6$ kcal/mol, $x_T \cdot T = 321 \pm 12$ K, $n = 18$, $r = 0.984$); (□) self-complementary G·A-mismatched dsRNA 9mers in 1 and 0.1 M NaCl, pH 7.0 and 5.0[3d] ($g = 6.5 \pm 0.6$ kcal/mol, $x_T \cdot T = 331 \pm 3$ K, $n = 31$, $r = 0.998$); (box with X) 13 bp hairpin RNA 42mers in 0.1 M NaCl, pH 7.0 [3e] ($g = 2 \pm 7$ kcal/mol, $x_T \cdot T = 317 \pm 24$ K, $n = 7$, $r = 0.979$). Compare also with ref 1j.

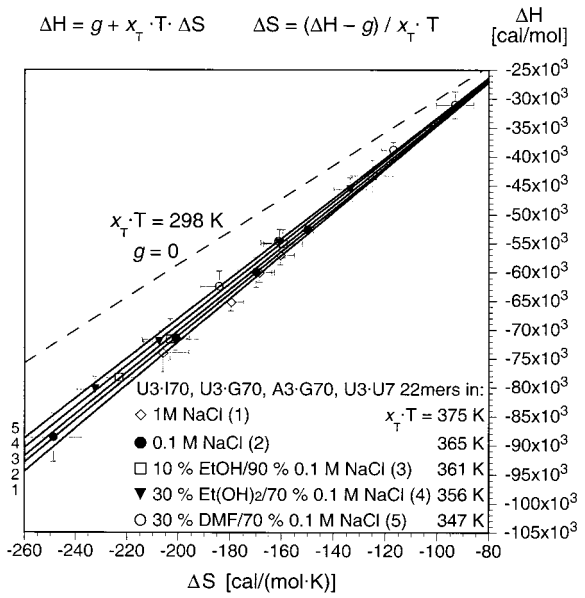

Figure 3. Solvent dependence of folding enthalpies and entropies (from Table 3 in ref 2b) and of compensation parameters of four singly mismatched (U3·I70, U3·G70, A3·G70, U3·U70) 22mer tRNA acceptor hairpins (◇) in 1 M NaCl ($g = -2.8 \pm 13.0$ kcal/mol, $x_T{\cdot}T = 375 \pm 72$ K, $n = 4$, $r = 0.998$), (●) in 0.1 M NaCl ($g = -2.1 \pm 2.0$ kcal/mol, $x_T{\cdot}T = 365.2 \pm 10.0$ K, $n = 4$, $r = 0.99996$), (□) in 10% (*v*/*v*) ethanol/0.1 M aqueous NaCl ($g = 2.1 \pm 4.7$ kcal/mol, $x_T{\cdot}T = 361 \pm 26$ K, $n = 4$, $r = 0.9997$), (▼) in 30% (*v*/*v*) ethylene glycol/0.1 M aqueous NaCl ($g = 2.0 \pm 4.0$ kcal/mol, $x_T{\cdot}T = 356 \pm 20$ K, $n = 4$, $r = 0.9998$), and (○) in 30% (*v*/*v*) dimethylformamide/0.1 M aqueous NaCl ($g = 1.4 \pm 2.0$ kcal/mol, $x_T{\cdot}T = 347 \pm 14$ K, $n = 4$, $r = 0.9999$). UV profiles at 260 nm, except in 30% DMF at 285 nm. Linear regressions as described in the caption to Figure 1. Error bars (gray) represent experimental standard deviations (50% confidence). The dashed line of zero free energy of folding at room temperature, $\Delta G_{25°C} = 0$, was added for comparison ($g = 0$, $x_T{\cdot}T = 298$ K, $x_{298K} = 1.0$). The area of crossing regression lines (not shown) lies between $\Delta H = -13$ and -15 kcal/mol and $\Delta S = -41.8$ and -47.5 cal/(mol·K), respectively, corresponding to free energies of folding $\Delta G_{25°C}$ of between -0.6 and -0.8 kcal/mol.

tion slopes $x_T{\cdot}T$ are remarkably similar: 373 K for the 22mers, 376 K for the 25mers, and 374 K for Meroueh and Chow's hairpins (caption to Figure 2).

The other subgroup involves four 3·70 variants of the 22mer tRNA acceptor hairpins,[2b] I3·U70, U3·G70, A3·G70, and U3·U70, exhibiting in a 0.1 M NaCl solution a well-dispersed free energy range and a particularly obvious linear EEC ($r = 0.99996$). The same variants fold under different compositional conditions—higher ionic strength or nonaqueous cosolvents—with characteristic changes in their EEC parameters g and $x_T{\cdot}T$. The compensation slopes for this set of hairpins follow the hierarchy of solvent systems a, b, c, d, and e: $x_T{\cdot}T$ [°C] = 102° (a: 1 M NaCl, pH 7.5), 92° (b: 0.1 M NaCl, pH 7.5), 88° (c: 10% ethanol/90% b), 83° (d: 30% ethylene glycol/70% b), 74 °C (e: 30% DMF/70% b), or $x_{298K} = 1.26$ (a), 1.22 (b), 1.21 (c), 1.19 (d), 1.16 (e). The compensation intercepts g are close to nil, although two, the ones in solvent systems a and b without organic additives, might be slightly in the exergonic range, -2.8 ± 13.0 kcal/mol (a), -2.1 ± 2.0 kcal/mol (b), while the others seem to lie on the slightly endergonic side (see caption to Figure 3).

Figure 3 depicts these trends and shows the corresponding regression lines. Irrespective of sequence effects or differences between single atomic groups, the four hairpin variants exhibit in the various solvent mixtures compensating entropic penalties to exothermicity-dependent degrees. The less exothermically the hairpins form, owing to the presence of hydration-destabilizing cosolvents, the closer the compensation slopes are to 25 °C (the smaller the enthalpy−entropy ratios are), i.e., the more severly compensating are the entropic folding penalties. If one were to further destabilize the hairpin down to an exothermicity of a double-stranded di- or trinucleotide, i.e., at ΔH values of -13 to -15 kcal/mol and $\Delta G_{25°C}$ of -0.6 to -0.8 kcal/mol, the solvent dependencies would appear to coalesce (crossover region not shown).

Discussion and Reanalysis

During the past decade a large amount of thermodynamic data on nucleic acids accumulated in the literature, which so far was not (or not "officially")[1j] analyzed with respect to the relationship between enthalpic and entropic contributions to folding. The body of experimental thermodynamic data on nucleic acid folding still is probably about an order of magnitude smaller than on protein denaturation. In addition, the range of unfolding enthalpies ΔH_{un} of oligonucleotides was concentrated before Morse and Draper's,[3d] Bevilacqua and Bevilacqua's[3e] and this investigation[2b] mostly in the region between 30 and 65 kcal/mol, rarely above, with some additional information for averaged base pair stacks in oligo- and polynucleotides at 5 to 15 kcal/mol (Figure 2). A highly linear correlation between ΔH and ΔS of nucleic acid unfolding has been noticed and mentioned explicitly,[3f] but was ascribed to an analytical artifact owing to the narrow range of experimental T_m values (approximately 20 to 65 °C). The thermodynamic analysis of certain double-mismatched G·A-containing RNA duplices[3d] and particularly stable 7 bp- [2b] and 13 bp-long[3e] hairpin stems extends the observed unfolding enthalpies to 110 kcal/mol (T_m 56−100 °C, ◆, ■, □, box with X in Figure 2). Hence, an analysis of the enthalpy−entropy relationship in nucleic acids, often ignored possibly due to experimental uncertainty, has become statistically obtrusive. It may help open new ways of interpreting thermodynamic parameters and put them into context as characterizing "folding units".

One very strong argument for the presence of a dominance of hydration thermodynamics over folding thermodynamics hydration-exempt as a general rule is given by the statistical analysis of a large body of experimental data on a variety of macromolecules as shown in Figures 1 and 2. What do we gain by studying compensation slopes and intercepts of a whole subgroup of closely related macromolecules? The slopes $x_T{\cdot}T$ and intercepts g are, respectively, compensation temperatures and free enthalpies (free of entropic penalties). The dominating component originates from fully compensating hydration keeping x_T close to 1.0. The experimental $x_T{\cdot}T$ and g values reflect *a restricted combination of folding enthalpies* ΔH *and entropies* $T{\cdot}\Delta S$ *for each subgroup, a confined range of enthalpy−entropy ratios (*T_m *values) realisable in a given macromolecular framework*. The dominant effect on the observed thermodynamics is that *the structure of the macromolecule limits the range of adoptable states for hydration molecules*, more so in the folded than in the denatured state, *such that a favorable net enthalpy results*. The more *stably* a structure folds in comparison to the general compensation line of the subgroup it belongs to (the further out in the direction following the *perpendicular*, dashed arrow in Figure 2), the less dominant are exothermic

but fully compensated hydration events that accompany the folding. The limiting slope $x_T \cdot T = 25$ °C for room-temperature conditions corresponds to a hypothetical full dominance of hydration interactions with no stabilizing folding interactions at all. On the other hand, the more *exothermically* a structure folds, producing datapoints *parallel* to the general compensation line of the subgroup it belongs to, the more dominant are hydration events that stabilize the folding.

Nucleic acids appear to be slightly more sensitive to destabilizing structural elements than proteins, as judged by their respective compensation slopes (Figure 1). The global difference in $x_T \cdot T$ between the two classes of compounds could, however, be an artifact of detection methods and/or, less likely, of sample number. The protein data are calorimetrically derived and thus model-independent. The nucleic acid data are optically derived thermodynamic parameters from van't Hoff plots or from curve fittings of temperature-dependent equilibrium shift profiles. These usually assume a two-state model, and temperature-independent folding enthalpies and entropies, thus, ignore any temperature-dependent heat capacity changes ΔC_p. "Mixed-sequence" polynucleic acids as well as a double-stranded DNA dodecamer were recently shown to exhibit positive ΔC_p values of, respectively, 40 to 85[7a] and 57 to 93[7b] cal/(mol·K) per base pair, owing to uncooperative single-strand ordering/denaturing.[7b,c] A linear correction to standard conditions of optically derived folding enthalpies and entropies with a nonzero ΔC_p value would necessarily diminish any enthalpy−entropy ratio (compensation slope) by some amount, since $\Delta H_{298K} = \Delta H_T + \Delta C_p \cdot (298\ K - T)$, $\Delta S_{298K} = \Delta S_T + \Delta C_p \cdot \ln(298\ K/T)$, and $|298\ K - T| > |\ln(298\ K/T)|$ (for $T > 1.0$ K). A deviation at 298 K using a high value of $\Delta C_p = 85$ cal/(mol·K) and $\Delta T = 75$ K, for instance, would diminish the enthalpy term by 6.4 kcal/mol and the entropy term by 5.6 kcal/mol per base pair corresponding to a 14% reduction of x_{298K}.

The more closely related the folding subgroups are the more similar are their heat capacity changes allowing for a more stringent analysis of their compensation behavior. For instance, comparing bi- with unimolecular associations, i.e., nucleic acid double-strands with hairpins, we observe—with the caveat of a relatively small number of analyzed subgroups to compare with—larger compensation slopes for hairpins than for a number of double-stranded systems (caption to Figure 2). This entails compensation intercepts that appear on the endergonic side (positive g), while many double-stranded nucleic acids show exergonic (negative) g values. If the relation between ΔH and ΔS were truly linear, as the global protein data may suggest (Figure 1), hairpins would consequently appear to be more sensitive to destabilizing structural elements, such as mismatches for instance, than double-stranded nucleic acids. At some strong, possibly external destabilization, through binding perhaps, a looped stem would become unstable (positive g) under conditions where the stem without the loop would still retain some residual stability (negative g). One could argue that a loop only folds with some stabilization from the stem; a "loop" without a stabilizing stem is usually denatured. However, a more extended analysis of the EEC shows that the relation between ΔH and ΔS can only appear to be linear within the amenable range of stabilities, it cannot be strictly linear for principal reasons (see Conclusions).

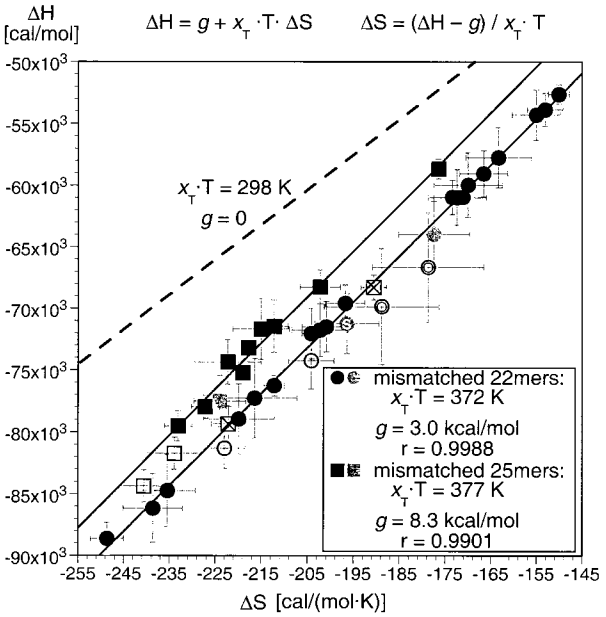

Figure 4. 2D EEC plot $\{\Delta H, \Delta S\}$ of 22mer (circles) and 25mer (squares) tRNA acceptor hairpins in 0.1 M NaCl. Dark symbols: 3·70 and 4·69 single-mismatches. Gray symbols: G3·U70 and G3·C70-G4·U69 (s for "shift") variants. Empty symbols: all-Watson−Crick A3·U70 and U3·A70 variants. Diagonally crossed squares and double-circles: all-Watson−Crick C3·G70 and G3·C70 variants. Error bars (gray) represent experimental standard deviations (50% confidence) as listed in Table S1 of the Supporting Information to ref 2b. The regression lines correspond to the trends of the mismatched variants only. Regression parameters and confidence values obtained as described in the caption of Figure 1. Regression parameters of the mismatched 22mers excluding G3·U70 and G3·C70-G4·U69 (●): $x_T \cdot T = 372$ K, $g = 3.1$ kcal/mol, $r = 0.9994$. Dashed line: $\Delta G_{25°C} = 0$ ($g = 0$, $x_T \cdot T = 298$ K, $x_{298K} = 1.0$).

In Figure 4 the difference in compensation behavior between the $tRNA^{Ala}$ acceptor 22mer and the 25mer hairpins is separately illustrated. All singly mismatched variants are treated as two subgroups according to their common helix-destabilizing property, a mispair in position 3·70 or 4·69 (black and gray symbols). The all-Watson−Crick A3·U70 and U3·A70 variants (empty circles and squares) appear slightly in the more stable area below the linear regression lines derived from the single-mismatches only. The G3·C70 and C3·G70 variants (double-circles and diagonally crossed squares) show an even greater deviation from the compensation lines imposed by a single structural perturbation of otherwise identical hairpin stems. The variants G3·U70 and G3·C70-G4·U69 (grey symbols, s for "shift") fold quite a bit less exothermically and are somewhat less stable than A3·U70 and U3·A70. Their compensation behavior, however, seems to be more in line with A3·U70 and U3·A70 than with the other mismatched variants.

The errors in determining ΔH and ΔS are, albeit relatively large, strongly correlated. The error bars depicted in Figure 4 represent standard deviations obtained from the experimental determination of several independent profiles for each strand (see Supporting Information to ref 2b). In the plot they describe ellipses with their main axes in the x and y directions. The real error ranges are within ellipses with their axes parallel and

(7) (a) Chalikian, T. V.; Völker, J.; Plum, G. E.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 7853−8. (b) Holbrook, J. A.; Capp, M. W.; Saecker, R. M.; Record, M. T., Jr. *Biochemistry* **1999**, *38*, 8409−8422. (c) Vesnaver, G.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 3569−73.

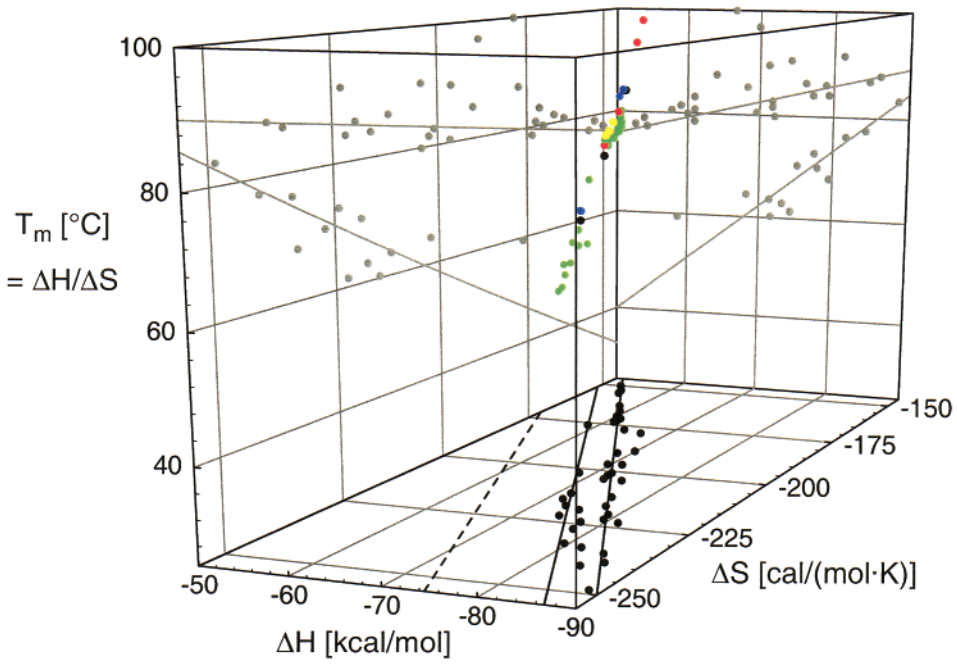

Figure 5. 3D EEC plot $\{\Delta H, \Delta S, T_m\}$ of 22mer and 25mer tRNA acceptor hairpins in 0.1 M NaCl (colored dots irrespective of strand length). Green: mismatches, except black for G3·U70 and G3·C70-G4·U69; blue, A3·U70 and U3·A70; red, C3·G70 and G3·C70; yellow, A·C-containing hairpins at pH 5.5. Colored dots constitute the surface $T_m = \Delta H/\Delta S$ (almost coplanar perspective) intersecting at $T_m = 25$ °C (dashed line). Black dots and lines in the isothermic $\{\Delta H, \Delta S\}$ plane at $T_m = 25$ °C: see Figure 4. Gray dots in the other planes: isoenthalpic and isoentropic projections ("shadows") of the colored dots. Gray lines and curves in the isoentropic and isoenthalpic planes: linear regressions and curve fittings according to eqs 6 and 7, respectively; fitting parameters for $\Delta H(T_m)$ in Figure 6 and $\Delta S(T_m)$ in the caption to Figure 6.

perpendicular to the regression lines.[8] Thermodynamic analyses as described in the preceding article basically result in ΔH, ΔS, and T_m values ($T_m = T_{\Delta G°=0}$). The ratio $\Delta H/\Delta S$ corresponds to T_m of unimolecular interactions. In bimolecular interactions $R\cdot\ln c_{tot}$ or $R\cdot\ln(c_{tot}/4)$ is to be added to ΔS, dependent on whether, respectively, identical or equimolar different partners form a complex, with c_{tot} being the molar total concentration of all binding partners. Usually, the data are supplemented with calculated ΔG_T values. The relative experimental errors in determining T_m are much smaller than those of ΔH, ΔS, and ΔG owing to the sigmoidal curvature of a melting profile which, in turn, is the basis for the correlation of errors in ΔH and ΔS. Therefore, it seems reasonable to extend an EEC plot like the one in Figure 4 with a third dimension, the ratio of both parameters as shown in Figure 5. The $\{\Delta H, \Delta S\}$ basis of this 3D EEC plot at $T_m = 25$ °C depicts the datapoints and compensation lines from Figure 4 as black dots and lines (no difference in the various datapoint symbols for simplicity). Hence, a conventional EEC plot can be viewed as an isothermic 2D projection of the datapoints $\{\Delta H, \Delta S, T_m\}$, all part of the surface $T_m = \Delta H/\Delta S$, which have been visualized here as colored dots (color codes in the Figure caption). Over unrealistically large values $\{\Delta H, \Delta S, T_m\}$ the surface describes the positive half of a hyperbolic parabaloid showing a zero inclination at very large values $\{\Delta H, \Delta S\}$, an increasingly rising inclination as $\{\Delta H, \Delta S\}$ approach the origin, and an infinite inclination at the origin (parallel to the T_m-axis, examples of 3D plots of the function in the Supporting Information). Within measurable temperatures and realistic thermodynamics, however, the surface is essentially a quite strongly inclined and slightly twisted plane. The perspective in Figure 5 was chosen to illustrate that the data all lie in this plane. The line of intersection of the datasurface $\{\Delta H, \Delta S, T_m\}$ at 25 °C is shown as a dashed line corresponding to $g = 0$, $x_T\cdot T = 298$ K or $x_{298K} = 1.0$ (compare to Figure 4). Different ways of characterizing the hairpins, or any noncovalent complex for that matter, with respect to EEC behavior are visualized by the other two projections in Figure 5, the equivalent isoenthalpic and isoentropic 2D projections (grey dots to the left and back plane, respectively). The different tendencies of both hairpins become immediately visible as different regression lines or curves (gray). The fact that the 25mers are generally less stable, thus, lower in T_m and closer to the dashed line (of intersection) than the 22mers, together with the shape and highly pronounced inclination of the $\{\Delta H, \Delta S, T_m\}$ datasurface necessitates strong differences between the isoentropic regression lines or the isoenthalpic regression curves that represent two different hairpin frameworks.The analysis of the exothermicities ΔH versus the corresponding mid-transition temperatures T_m is shown more closely in Figure 6. The typical clustering and differences among the variants become much more significant than in Figure 4. Both subgroups have been characterized by two separate linear regressions each involving the mismatched variants only, according to,

$$\Delta H(T_m) = \kappa\cdot(\tau - T_m) \quad \text{and} \quad \Delta S(T_m) = \kappa\cdot\ln(\tau/T_m) \qquad (6, 7)$$

where $\kappa\cdot\tau$ is the linear regression intercept in [cal/mol] at $T_m = 0$, τ the intercept in [Kelvin] at $\Delta H = 0$ (eq 6) or $\Delta S = 0$ (eq 7) and κ is the negative regression slope in [cal/(mol·K)]. Positive κ values reflect "enthalpy-driven" (-dominated) and negative κ values "entropy-driven" (-dominated) foldings. The analogy between the empirically derived eq 6 and the well-

(8) Kita, F.; Adam, W.; Jordan, P.; Nau, W. M.; Wirz, J. *J. Am. Chem. Soc.* **1999**, *121*, 9265−75.

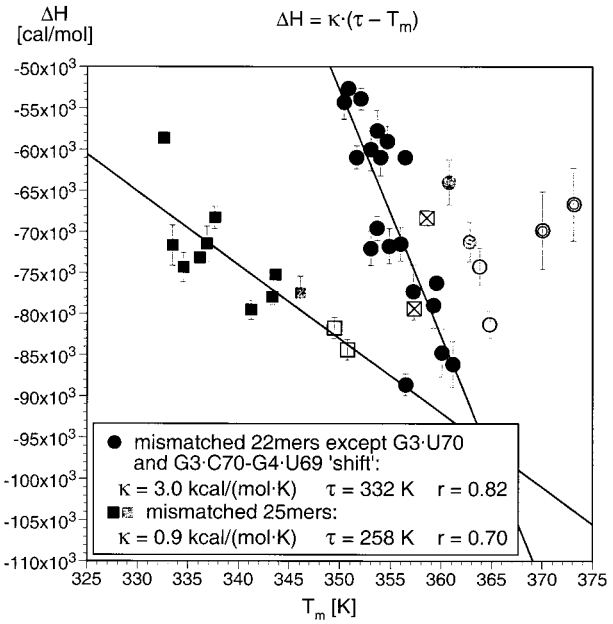

Figure 6. ΔH versus T_m (= $\Delta H/\Delta S$) of 22mer (circles) and 25mer (squares) tRNA acceptor hairpins in 0.1 M NaCl. Error bars (gray) represent experimental standard deviations (50% confidence), the S.D. for T_m is ≥ ±0.1° and ≤ ±1.0°. Symbols as in Figure 4. The regression lines and parameters correspond to the trends of the mismatched variants only, G3·U70 and G3·C70-G4·U69 22mers excluded. The regression parameters in the figure insert refer to eq 6. Temperature τ corresponds to the intercept at $\Delta H = 0$. Parameter κ is the negative regression slope. The regression parameters for eq 7 (plot not shown) are mismatched 22mers (●) $\kappa = 2.8$ kcal/(mol·K), $\tau = 332$ K, $r = 0.81$; and mismatched 25mers (solid box, gray box) $\kappa = 0.7$ kcal/(mol·K), $\tau = 248$ K, $r = 0.60$.

known relation describing a linear temperature dependence of transition enthalpies and entropies, respectively,

$$\Delta H_T = \Delta H_{ref} + \Delta C_p \cdot (T - T_{ref}) \quad (8)$$

$$\Delta S_T = \Delta S_{ref} + \Delta C_p \cdot \ln(T/T_{ref}) \quad (9)$$

insinuates the meaning of κ as a measure of heat capacity of a higher-order structure (*kappa* for capacity) and τ as the reference temperature for $\Delta H_{ref} = \Delta S_{ref} = 0$. Therefore, if the relationship between ΔH and T_m is linear (T_m-independent τ and κ for a given framework or subgroup)—as apparently may seem to be the case for the 22mers and 25mers studied here ($r = 0.82$ and 0.70, Figure 6) and also for other nucleic acids derived from optical[2a,3b] (analyses not shown) and calorimetric measurements (compilation and first {ΔH, T_m} plot published in ref 7a)[9]—then the relationship between ΔS and T_m must be logarithmic (see eq 7; {ΔS, T_m} plot—similar to {ΔH, T_m} plot—not shown; corresponding fitting parameters in the caption to Figure 6, $r = 0.81$ and 0.60). Because the EEC for the studied hairpins appears highly linear within the measured range of stabilities, the relationship between ΔG and T_m, according to

$$\Delta G_T = h - T_m \cdot s_T \quad \text{or} \quad \Delta G_T = (\tau - T_m) \cdot s_T \quad (10, 11)$$

is even more linear than the one between ΔH and T_m (r values 0.89 to 0.97, Figure 7). In eq 10 the intercept h in [cal/mol] represents the folding enthalpy of a higher-order structure at a hypothetical $T_m = 0$ K and therefore is the least reliable parameter. In eq 11, parameter $\tau = h/s_T$ in [Kelvin] is the intercept and reference temperature at $\Delta G_T = 0$. Parameter s_T is the negative slope in [cal/(mol·K)]. τ, h, and s_T are a measure for the resistance of a molecular framework toward, respectively, enthalpic (τ, h) and entropic, when $s_T > 0$, *de*stabilization, i.e., a measure for structural rigidity. The difference in regression slopes s_T shows that the 22mers are approximately twice as resistant to structural perturbations as the 25mers (inserts Figure 7).

To be able to perform the same correlation analysis with the nucleic acid denaturation data available from the literature (as done for ΔH versus ΔS in the legend of Figure 2), that is, mostly of bimolecular equilibria, a modified version of eqs 10 or 11 and a concentration-normalized slope parameter $s_{1M,T}$ was used (eqs 12, 13),

$$\Delta G_T = h - (\Delta H/\Delta S) \cdot s_{1M,T} \quad (12)$$

$$\Delta G_T = \{\tau - (\Delta H/\Delta S)\} \cdot s_{1M,T} \quad (13)$$

in which $s_{1M,T}$ is proportional to s_T according to eq 14 for self-complementary and eq 15 for non-self-complementary bimolecular systems ($s_{1M,T} = s_T$ in unimolecular systems):

$$s_T = \{(R \cdot \ln c_{tot} + \Delta S)/\Delta S\} \cdot s_{1M,T} \quad (14)$$

$$s_T = \{(R \cdot \ln(c_{tot}/4) + \Delta S)/\Delta S\} \cdot s_{1M,T} \quad (15)$$

The corresponding linear regression parameters are listed in the legend of Figure 7 and often show astonishingly high regression coefficients of above 90%.

Perturbation Sensitivity. In both {ΔH, T_m} and {ΔG, T_m} plots, the all-Watson−Crick variants (empty, double, and crossed symbols in Figure 6, red and blue symbols in Figure 7) clearly prompt as exceptionally stable, more or less significantly distinct from the regression lines of the mismatched variants. Those, in turn, show a now significant difference between the 22mers and the 25mers—visibly different regression slopes κ or s_T and intercepts τ—indicating that a UUCG loop renders an A-RNA stem less sensitive to structural perturbations of the hydration layer (κ) and the folded covalent structure (s_T) than the other loop: the steeper the slope, the larger the possible enthalpy, entropy, and free energy range within one Kelvin, i.e., the larger the heat capacity, the larger the resistance toward destabilization of a folded subgroup. With respect to stability *and* exothermicity, thus, "pairing strength" and "hydration quality", a G·U wobble pair (gray-filled symbols in Figure 6, purple and more stable cyan symbols in Figure 7) behaves like an A·U or U·A pair (empty and blue symbols, respectively), especially within the UUCG-looped stem. These G·U pairs stabilize the hairpin mainly through a relatively low entropic compensation (penalty) without being highly exothermic. Finally, the most stable G3·C70 and C3·G70 variants of both hairpins are distinct from all the other ones. These base pairs gain stability from a combination of an exceptionally low entropic penalty and moderate exothermicity more than any other base pair.[2b]

Finally, the results from the 22mer variants that showed an exceptionally linear EEC in 0.1 M NaCl, U3·I70, U3·G70, A3·G70, and U3·U70, and that were analyzed in 1 M NaCl and with various organic cosolvents, have been replotted as ΔH

(9) Note added in proof: {ΔH, T_m} and {ΔS, T_m} plots were used as in ref 7a) to calculate average ΔC_p values. Diamond, J. M.; Turner, D. H.; Mathews, D. H. *Biochemistry* **2001**, *40*, 6971−81. Mathews, D. H.; Turner, D. H. *Biochemistry* **2002**, *41*, 869−80.

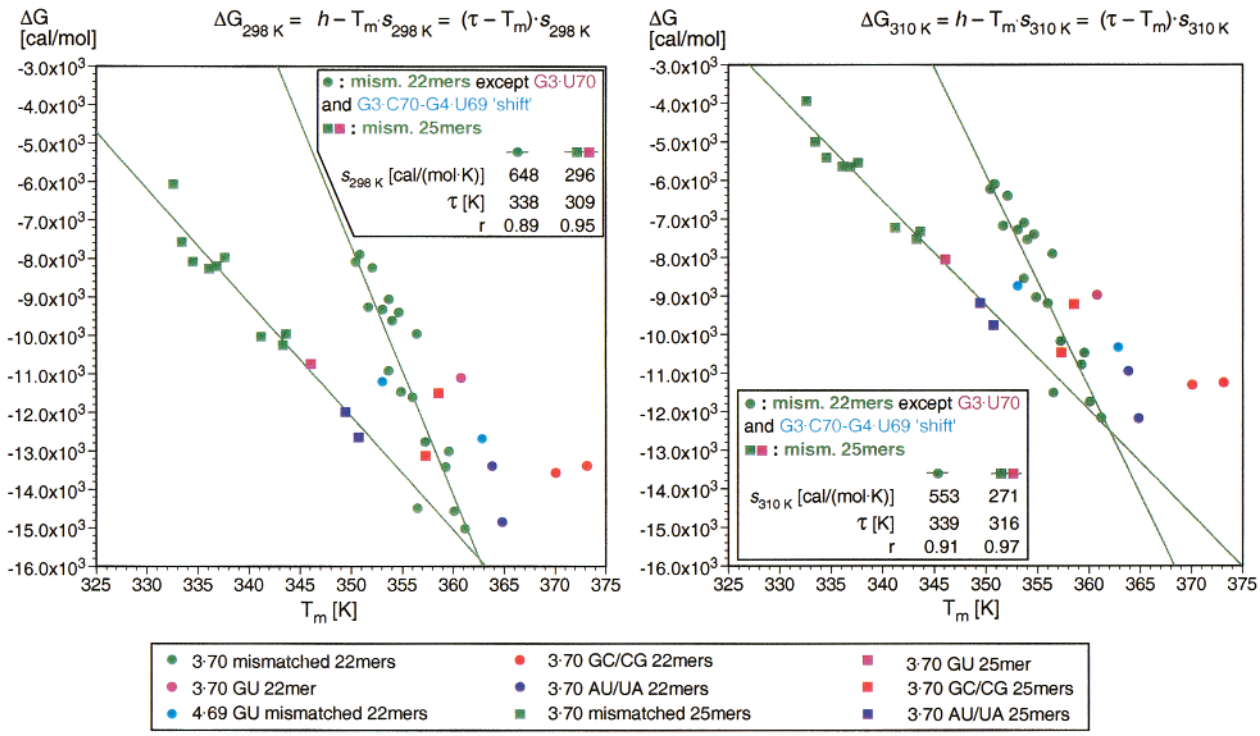

Figure 7. ΔG_T (= $\Delta H - T{\cdot}\Delta S$) versus T_m (= $\Delta H/\Delta S$) of 22mer (circles) and 25mer (squares) tRNA acceptor hairpins in 0.1 M NaCl at 25 °C (Figure 7a, left) and 37 °C (Figure 7b, right). *Average* 90% confidence levels (see Figure S11 of the Supporting Information to ref 2b): dev. (ΔG, 90%) = ±0.8 kcal/mol and dev. (T_m, 90%) = ±1.0°. Note that the G3·C70 and C3·G70 variants (red symbols) have exceptionally high transition temperatures as compared to their stability at lower temperatures. The regression lines according to eq 11 correspond to the trends of the mismatched variants only, G3·U70 and G3·C70-G4·U69 22mers excluded. Regression slopes s_T, intercepts (at $\Delta G_T = 0$) τ, and regression coefficients r for T = 25° and 37 °C are listed in the plots. Corresponding regression parameters for nucleic acid denaturation data $\Delta G_{310\ K}$ versus ($\Delta H/\Delta S$) from the literature using eq 13 are listed below (symbol code as in the caption of Figure 2): (◇) $s_{1M,310K}$ = 31.4 cal/(mol·K), τ = 324 K, r = 0.781; (△) $s_{1M,310K}$ = 49.9 cal/(mol·K), τ = 270 K, r = 0.909; (○) $s_{1M,310K}$ = 17.9 cal/(mol·K), τ = 309 K, r = 0.971; (⊙) $s_{1M,310K}$ = 19.7 cal/(mol·K), τ = 306 K, r = 0.991; (▲) $s_{1M,310K}$ = 197.4 cal/(mol·K), τ = 321 K, r = 0.974; (#) $s_{1M,310K}$ = 127.9 cal/(mol·K), τ = 300 K, r = 0.768 (from curve fit parameters); (●) s_{310K} = 148.3 cal/(mol·K), τ = 311 K, r = 0.953; (solid triangle pointing to the right) $s_{1M,310K}$ = 183.6 cal/(mol·K), τ = 330 K, r = 0.858; (◩) $s_{1M,\ 310K}$ = 41.3 cal/(mol·K), τ = 193 K, r = 0.428 (over possibly two subgroups); (◹) $s_{1M,310K}$ = 63.6 cal/(mol·K), τ = 246 K, r = 0.699 (over possibly two subgroups); (box with X) s_{310K} = 261.9 cal/(mol·K), τ = 307 K, r = 0.949.

versus T_m in Figure 8. As expected, the {ΔH, T_m} datapoints are by far not as linearly related as they are in the EEC plot {ΔH, ΔS} depicted in Figure 3. Yet, the converging compensation slopes $x_T{\cdot}T$ in Figure 3 are not as informative as they may seem, since this kind of convergence (or coalescence) is a mere consequence of the fact that EEC plots are 2D projections of a {ΔH, ΔS, T_m} surface. The surface shows a slightly twisted rising inclination as ΔH and ΔS approach nil (not shown, see Supporting Information). A 2D projection of parallel lines (Figure 8) on a twisted surface always produces converging lines (Figure 3).

In Figure 8 we immediately see that organic denaturants probably do not alter the enthalpy-driven ($\kappa > 0$) "folding mechanism", since the relation between stability and exothermicity of folding remains essentially the same (similar slopes κ, i.e., approximately parallel lines for □, ▼, and ○) despite the destabilizing action of the denaturants (different intercepts τ for □, ▼, and ○) and despite mismatch-dependent differences (A3·G70 vs U3·U70). We obtain a similar picture when analyzing {ΔG_T, T_m} according to eq 11 (plot not shown). In short, the resistance of these hairpins toward mismatch-induced destabilization as judged by s_{298K} (caption to Figure 8) is the highest in 0.1 M NaCl (1.01 kcal/(mol·K)), somewhat lower in the presence of protic (alcoholic) additives (0.77, 0.68 kcal/(mol·K)) and the lowest in the presence of aprotic additive DMF (0.56 kcal/(mol·K)). In 1 M NaCl, U3·G70 strongly deviates from the other variants that show a "rigidity" of 0.86 kcal/(mol·K) (= s_{298K} value without U3·G70).

Accepting that exothermicity mainly reflects favorable hydration, the data in Figure 8 show that the hydration layer remains within the explored limits much the same in all protic media with similar salt but different denaturant content. The aprotic cosolvent DMF destabilizes *and* diminishes $\Delta\Delta H$ most (confirming the finding for double-stranded DNA in ref 10), read as a sign for a disturbed hydration layer. On the other hand, 1 M NaCl stabilizes the hairpins but by no means mismatch-independently (U3·I70 vs U3·G70). The rather different regression slopes κ for 0.1 and 1 M NaCl (◇, ●) including U3·G70, but similar κ when U3·G70 was excluded from the regression of the 1 M NaCl data, originates from an apparently exceptional stabilization of the U·G mismatch, although both U·I and U·G are known to adopt the same wobble geometry.[11] A high NaCl content produces through efficient scavenging of bulk water molecules into ion-solvating hydration shells more "anhydrous conditions". The NaCl-induced stabilization of both U3·G70 and U3·I70 variants is entropy-driven, but the more hydrated U·I mismatch with its undistrubed hydration network in the

(10) DePrisco Albergo, D.; Turner, D. H. *Biochemistry* **1981**, *20*, 1413−18.
(11) (a) 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. (b) Cruse, W. B. T.; Aymani, J.; Kennard, O.; Brown, T.; Jack, A. G. C.; Leonard, G. A. *Nucleic Acids Res.* **1989**, *17*, 55−72.

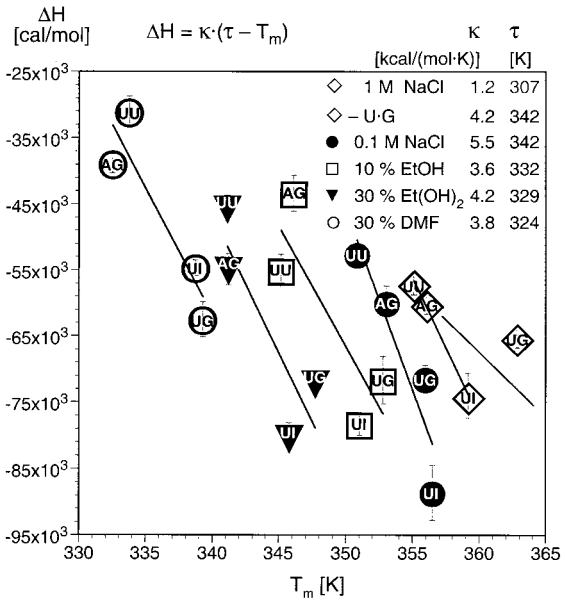

Figure 8. ΔH versus T_m of U3·I70, U3·G70, A3·G70, and U3·U70 22mer tRNA acceptor hairpins. Data from Table 3 in ref 2b. Error bars and linear regressions according to eq 6 as in the caption to Figure 6. Regression parameters κ in [kcal/(mol·K)] and τ in [Kelvin]. The regression lines do not implicate a strict linear relationship between ΔH and T_m in a subgroup of variants. They merely outline a feature of a cluster. Regression coefficients for $\Delta H(T_m)$: $r(\diamond) = 0.58$, $r(\diamond$ without U3·G70$) = 0.998$, $r(\bullet) = 0.92$, $r(\square) = 0.85$, $r(\blacktriangledown) = 0.87$, $r(\circ) = 0.92$. Regression parameters and coefficients ΔG_{298K} according to eq 11: 1 M NaCl ($\diamond$) $s_{298K} = 352$ cal/(mol·K), $\tau = 328$ K, $r = 0.77$; 1 M NaCl ($\diamond$ without U3·G70) $s_{298K} = 855$ cal/(mol·K), $\tau = 344$ K, $r = 0.998$; 0.1 M NaCl ($\bullet$) $s_{298K} = 1019$ cal/(mol·K), $\tau = 343$ K, $r = 0.93$; 10% EtOH ($\square$) $s_{298K} = 681$ cal/(mol·K), $\tau = 335$ K, $r = 0.90$; 30% $Et(OH)_2$ ($\blacktriangledown$) $s_{298K} = 771$ cal/(mol·K), $\tau = 332$ K, $r = 0.92$; 30% DMF ($\circ$) $s_{298K} = 560$ cal/(mol·K), $\tau = 327$ K, $r = 0.95$.

shallow groove responds to dehydration much more severly with entropic compensation (penalty) than does the U·G mismatch.

Conclusions

The folding of nucleic acids is exceptionally useful for studying enthalpy−entropy compensation (EEC). The experimental $\{\Delta H, \Delta S\}$ relationship appears to be highly linear and straightforward to analyze when compared to a number of other systems of molecular recognition, such as host−guest complexes and ligand binding.[11,12] Each of the binding partners in nucleic acids bears akin binding sites, the nucleobases, and predictable binding modes, base pairing. The differences between variants are rather small, and local changes can be introduced without large collateral, unforseen consequences that would otherwise complicate the structure−thermodynamics relationship. A small amount of structural changes is a necessary feature for significantly linear EECs to show up.

EEC plots (ΔH versus ΔS) were shown to be 2D projections of $\{\Delta H, \Delta S, T_m\}$ plots which may explain some of the generally observed characteristics of conventional EEC plots, such as converging compensation lines. An analysis of the function $T_m = \Delta H/\Delta S$ (see Supporting Information) shows that any EEC effect can only appear linear but indeed has to adopt a curvature possibly similar to the one first described empirically by Dudley Williams and co-workers.[1g] Therefore, this function provides for the first time a mathematical basis for the analytical description of a fundamental relationship between ΔH and ΔS proving that, in all cases, both values must pass the origin $\Delta H = \Delta S = 0$. The linear relationships as described by eqs 1 and 5 with $x_T \cdot T$ and g as fitting parameters turn out to be good tangential approximations of the true nonlinear relationship with a zero-intercept. One conclusion on a physical basis is that, when we speak about linear relationships between any ΔH and ΔS values, we probably mean "linear" on the T_m surface, not on its $\{\Delta H, \Delta S\}$ projection plane. The fact that incremental physical effects are additive makes the corresponding datapoints appear linear on a twisted surface which always introduces curvature on a projection plane.

The analysis of unfolding or denaturing thermodynamics of any macromolecular framework as ΔH versus $\Delta H/\Delta S$ or, equivalently, ΔS or ΔG versus $\Delta H/\Delta S$ ($= T_m$(unimolecular)) according to eqs 6, 7, and 11 compiles statistically significant and readily readable information about the folding thermodynamics of whole higher-order structure families, groups of variants of any kind, as opposed to single-populations. Broader systematic investigations in this direction could help understand and usefully interpret measured ΔH and ΔS values in terms of EEC and solvation dominance. They would give us a means of quantifying a dynamic response behavior characteristic for any hairpinned or other folded structure, or folding domain, foldamer, found, for instance, in nature (such as in rRNA) through the use of the fitting parameters $x_T \cdot T$, τ, κ, s_T, and $s_{1M,T}$, most significantly the latter two. We quantify a "general folding stability" through temperature $x_T \cdot T$ (slope in [K], more reliable, eq 1) and τ (intercept in [K], less reliable, eqs 6, 7, 11); a "resistance to structural perturbations" through s_T or $s_{1M,T}$ (slope in [cal/(mol·K)], eqs 10−13) derived, for instance, from a certain "sensitivity of the hydration layer" through capacity κ (slope in [cal/(mol·K)], eqs 6, 7). This kind of analysis should also become useful for (particularly aqueous) solutions of any kind of supramolecular complex. More datapoints may allow one in the future to obtain $x_T \cdot T$, κ, and s_T values and a common τ value from one simultaneous fitting of all parameters in 4D phase space $\{\Delta H, \Delta S, \Delta H/\Delta S, \Delta H - T \cdot \Delta S\}_T$.

Acknowledgment. Many thanks to Lei Liu, presently at the Department of Chemistry, Columbia University, New York, and Prof. Qing-Xiang Guo, Department of Chemistry, University of Science and Technology of China, Hefei, P. R. China, for making available their protein unfolding thermodynamic data collection. The financial support by the Swiss National Science Foundation and the Novartis Foundation is gratefully acknowledged.

Supporting Information Available: 3D plots of the functions $T_m = \Delta H/\Delta S$, $\Delta G = \Delta H - T \cdot \Delta S$ and data as in Figure 5 but including the origin $\Delta H = \Delta S = 0$ (PDF file). This material is available free of charge via the Internet at http://pubs.acs.org.

JA016131X

(12) Ackroyd, P. C.; Cleary, J.; Glick, G. D. *Biochemistry* **2001**, *40*, 29911−22.