Accelerated Publications

Evidence that Folding of an RNA Tetraloop Hairpin Is Less Cooperative than Its DNA Counterpart†

Ellen M. Moody, Jessica C. Feerrar, and Philip C. Bevilacqua*

Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802

Received April 2, 2004; Revised Manuscript Received May 15, 2004

ABSTRACT: Hairpin secondary structural elements play important roles in the folding and function of RNA and DNA molecules. Previous work from our lab on small DNA hairpin loop motifs, d(cGNAg) and d(cGNABg) (where B is C, G, or T), showed that folding is highly cooperative and obeys indirect coupling, consistent with a concerted transition. Herein, we investigate folding of the related, exceptionally stable RNA hairpin motif, r(cGNRAg) (where R is A or G). Previous NMR characterization identified a complex network of seven hydrogen bonds in this loop. We inserted three carbon (C3) spacers throughout the loop and found coupling between G1 of the loop and the CG closing base pair, similar to that found in DNA. These data support a GNRA motif being expandable at any position but before the G. Thermodynamic measurements of nucleotide-analogue-substituted oligonucleotides revealed pairwise-coupling free energies ranging from weak to strong. When coupling free energies were remeasured in the background of changes at a third site, they remained essentially unchanged even though all of the sites were coupled to each other. This type of coupling, referred to as "direct", is peculiar to the RNA loop. The data suggest that, for small stable loops, folding of RNA obeys a model with nearest-neighbor interactions, while folding of DNA follows a more concerted process in which the stabilizing interactions are linked through a conformational change. The lesser cooperativity in RNA loops may provide a more robust loop that can withstand mutations without a severe loss in stability. These differences may enhance the ability of RNA to evolve.

Hairpins or stem-loops are common secondary structural elements in RNA and DNA and have important biological functions including regulating gene expression, initiating RNA folding, and forming tertiary structures (*1*−*4*). Thermally stable RNA and DNA tri- and tetraloop hairpin sequences have been identified by sequence comparison (*5*, *6*) and temperature-gradient gel electrophoresis (TGGE)[1] selection experiments (*7*−*9*). A number of structures of such loops have been solved (*10*−*13*) and reveal an array of hydrogen-bonding and stacking interactions. The stability of the RNA and DNA hairpins has been found to depend on both the base composition of the loop and its closing base pair (cbp) (*6*, *14*−*16*).

TGGE experiments led to the identification of 4 classes of stable DNA tetraloops: d(cGNABg), d(cGNNAg), d(cCNNGg), and d(gCNNGc) (*9*). The stable d(cGNAg) and d(cGNABg) loops have been extensively studied by NMR and nucleotide analogue substitutions and found to be stabilized by three interactions: two loop−loop hydrogen bonds in a sheared GA pair and a loop-closing base-pair interaction (*9*, *17*−*20*). In addition, the d(cGNAg) and d(cGNABg) loops were found to be expandable at any position, except at the 5′ end of the loop (*21*). Indeed, in both DNA and RNA, certain hairpin loop sequences have much greater thermodynamic stability with a CG cbp than expected from Watson−Crick base pairing alone ($\Delta\Delta G^\circ_{37}$ = 2−3 kcal/mol) (*6*−*9*, *22*). A stabilizing stacking interaction is therefore presumed to exist between position 1 of these loops and the CG cbp.

Recently, the cooperativity for folding of d(cGNAg) and d(cGNABg) loops was investigated by single, double, and higher order functional group mutagenesis cycles (*19*, *20*). Cooperativity is an important concept in molecular recognition and enzymology that links function and structure (*23*−*25*). Deletion of single functional groups resulted in unusually large thermodynamic penalties, ranging from 1.35 to 1.65 kcal/mol. Double mutant cycles revealed significant nonadditivity, with coupling free-energy terms (δ) ranging from −0.9 to −1.3 kcal/mol. Moreover, repeating of double mutant cycles in the absence of the third interaction revealed sharply diminished or even zero coupling free-energy terms (*19*, *20*). This behavior, referred to as indirect coupling (*23*), suggests that all three interactions need to be present simultaneously

† This work was supported by NSF CAREER Grant MCB-9984129 and a Camille Dreyfus Teacher−Scholar Award and Sloan Fellowship to P.C.B. and a Paul Berg Award to E.M.M.

* To whom correspondence should be addressed. Phone: (814) 863-3812. Fax: (814) 863-8403. E-mail: pcb@chem.psu.edu.

[1] Abbreviations: B, C, G, or T; C3, 3 carbon spacer; cbp, closing base pair; EDTA, ethylenediaminetetraacetic acid; I, inosine; N, A, C, G, or U(T); $P_{10}E_{0.1}$, 10 mM sodium phosphate and 0.1 mM Na_2EDTA (pH 7.0); R, A or G; TGGE, temperature-gradient gel electrophoresis; T_M, melting temperature; UV, ultraviolet spectroscopy.

10.1021/bi049350e CCC: $27.50
Published on Web 05/29/2004

to have significant stability. These studies supported the notion that simple, minimally stable DNA hairpin loops fold in a concerted fashion, in which the interactions are linked through a conformational change. We offered an analogy for indirect coupling of a three-legged stool, wherein removal of any one leg results in the collapse of the stool (*19*).

Because hairpins are more common in RNA than DNA and because RNA motifs tend to have more hydrogen bonds, we thought it would be interesting to investigate the origin of coupling in RNA hairpin loops. The model RNA hairpin chosen for this study is a member of the phylogenetically conserved (*5*) and unusually stable (*6*) r(GNRA) family, for which a number of structures have been determined (*11*, *26*−*30*). As seen in d(cGNA(B)g) loops, this loop contains a sheared GA pair and a loop-closing base-pair interaction, as well as five additional loop−loop hydrogen bonds (Figure 1). Together, these eight interactions and the surrounding 2′-hydroxyls provide a much different context in which to assess cooperativity and the origin of coupling.

MATERIALS AND METHODS

Preparation of RNA. RNA was prepared by solid-phase synthesis and deprotected as per the suggestions of the manufacturer (Dharmacon). Select oligonucleotides were purified on a C18 HPLC column using an ammonium acetate and acetonitrile gradient. The major peak was collected and desalted by dialysis or elution on a disposable C-18 column (Waters), and the molecular weight was confirmed by electrospray ionization mass spectrometry. Oligonucleotides were stored in $P_{10}E_{0.1}$ [=10 mM sodium phosphate and 0.1 mM Na_2EDTA (pH 7.0)]. All RNA had the general sequence 5′-r(ggaX$L_1L_2L_3L_4$X′ucc), where X and X′ are complementary nucleotides forming the cbp and "L" indicates a loop nucleotide. All oligonucleotides have the same three beginning (5′-gga) and ending (ucc-3′) nucleotides; therefore, only the loop and cbp are provided in the text.

UV Melting Experiments. UV absorbance melting profiles were obtained in $P_{10}E_{0.1}$ at 260 and 280 nm and analyzed using Kaleidagraph version 3.5 (Synergy Software), as described elsewhere (*9*, *31*). Melts were found to be independent of strand concentration, consistent with the hairpin conformation. A low ionic strength buffer was chosen for three reasons: to favor the hairpin conformation over the duplex (*31*), to prevent the T_M from being so high that an upper baseline was hard to define, and to facilitate the comparison to earlier studies on DNA done in this buffer (*9*, *19*−*21*).

Analysis of Double Mutant Cycles. The additivity of ΔG°_{37} values was analyzed similarly as described (*19*, *23*, *24*, *32*). The free-energy change associated with mutation A is denoted ΔG_A, while the change associated with mutation A in the presence of mutation B is denoted ${}^B\Delta G_A$. The magnitude of the nonadditive effect between mutations A and B is the coupling free energy, δ_{AB}, which was calculated according to the equations

$$\delta_{AB} = \Delta G^\circ_{37}(M_{00}) + \Delta G^\circ_{37}(M_{11}) - [\Delta G^\circ_{37}(M_{10}) + \Delta G^\circ_{37}(M_{01})] \quad (1a)$$

$$\delta_{AB} = \Delta G_{AB} - [\Delta G_A + \Delta G_B] \quad (1b)$$

where M_{00}, M_{01}, M_{10}, and M_{11} signify the unmodified, two single mutant, and double mutant sequences, respectively. A negative value for δ_{AB} reflects deletion of the first (set of) interaction(s) weakening the second (set of) interaction(s) and signifies positive coupling between the functional groups (*23*). A positive value of δ_{AB}, on the other hand, reflects deletion of the first (set of) interaction(s) strengthening the second (set of) interaction(s) and signifies negative coupling. These ideas are expressed well in eq 1b, wherein nonzero values of δ_{AB} reflect effects of double mutants not being the sum of the effects of single mutants. A δ_{AB} value of 0 supports no coupling. A double mutant is considered "completely nonadditive" if δ_{AB} equals the smaller of $-\Delta G_A$ or $-\Delta G_B$, which causes either ${}^B\Delta G_A$ or ${}^A\Delta G_B$ to approach zero (*20*).

Certain double mutant cycles were repeated in the background of a change elsewhere in the loop to test whether the coupling was direct or indirect. The equations for these cases are

$${}^C\delta_{AB} = \Delta G^\circ_{37}(M_{001}) + \Delta G^\circ_{37}(M_{111}) - [\Delta G^\circ_{37}(M_{101}) + \Delta G^\circ_{37}(M_{011})] \quad (2a)$$

$${}^C\delta_{AB} = {}^C\Delta G_{AB} - [{}^C\Delta G_A + {}^C\Delta G_B] \quad (2b)$$

where a superscript C denotes a mutational configuration at site C. If δ_{AB} equals ${}^C\delta_{AB}$, then coupling is direct, otherwise it is indirect (*23*). Errors were propagated as described (*19*).

RESULTS

Probing Loop-Closing Base-Pair Interactions in r(GAAA). We recently tested d(cGCAg) and d(cGCACg) hairpins for the presence of loop-closing base-pair interactions by inserting C3 spacers throughout the loop to interrupt potential interactions (*21*). Spacers were tolerated at all positions except between the cbp and the first position of the loop, where a large destabilization penalty ($\Delta\Delta G^\circ_{37} = 1.61$ kcal/mol) was incurred. This effect was absent with a GC cbp, which is 2−3 kcal/mol less stable than a CG cbp. These results supported a stable stacking interaction between position 1 of the loop and a CG cbp.

For r(cGAAAg), we likewise inserted C3 spacers throughout the loop and determined the thermodynamic consequences (Figure 2, Table 1). The pattern of energetic effects was qualitatively similar to that found for DNA. A very large destabilization ($\Delta G^\circ_{37} = 1.89$ kcal/mol in RNA versus 1.61 kcal/mol in DNA) was incurred for C3 spacer insertion between the CG cbp and position 1 of the loop, with smaller effects (from +0.18 to −0.44 kcal/mol) for substitution elsewhere in the loop.

Next, we looked at the importance of having a CG cbp for a r(GNRA) loop; surprisingly, to our knowledge, this effect has not yet been reported. A destabilizing effect ($\Delta G^\circ_{37} = 1.27$ kcal/mol) was found for the CG to GC swap (Table 1), consistent with a free-energy bonus for a CG cbp. Nevertheless, the magnitude of this effect, although significant, was not as large as that found for a r(UNCG) loop (2.3 kcal/mol) (*6*) or for d(cGNAg) (3.06 kcal/mol) and d(cGNABg) (1.88 kcal/mol) loops (*9*).

The smaller energetic penalty for a CG to GC switch for a r(GNRA) loop suggests that a GC cbp might also be able to interact stably with this loop. To test this idea, we

Table 1: Free-Energy Parameters for C3 Spacer Insertion in the r(GAAA) Loop

sequence[a]	$\Delta H^{\circ b}$ (kcal mol^{-1})	$\Delta S^{\circ b}$ (cal mol^{-1} K^{-1})	$\Delta G^{\circ}_{37}{}^{b,c}$ (kcal mol^{-1})	$T_M{}^b$ (°C)	$\Delta G_A{}^{b,c}$ (kcal mol^{-1})
cGAAAg	−41.7 ± 0.8	−122.6 ± 2.4	−3.65 ± 0.09	66.8 ± 0.5	
*C3*GAAA	−35.9 ± 1.4	−110.0 ± 4.4	−1.76 ± 0.14	53.0 ± 1.0	1.89 ± 0.17
G*C3*AAA	−46.5 ± 4.0	−136.8 ± 11.9	−4.09 ± 0.29	67.0 ± 1.1	−0.44 ± 0.30
GAA*C3*A	−44.1 ± 3.1	−131.1 ± 9.4	−3.47 ± 0.17	63.5 ± 0.7	0.18 ± 0.19
GAAA*C3*	−44.0 ± 2.0	−129.4 ± 5.8	−3.90 ± 0.16	67.1 ± 0.6	−0.25 ± 0.18
gGAAAc	−39.6 ± 3.3	−119.9 ± 10.4	−2.38 ± 0.07	56.9 ± 1.4	
*C3*GAAA	−33.6 ± 8.6	−105.1 ± 26.9	−1.03 ± 0.29	46.8 ± 1.5	1.35 ± 0.30

[a] Sequences in bold type are the reference for the sequences below. Changes from the reference sequence are italicized and listed from 5′ to 3′ positions in the loop. [b] Errors are the standard deviations from three or more measurements and were propagated by standard methods. [c] An extra significant figure is provided to avoid round-off error in subsequent calculations.

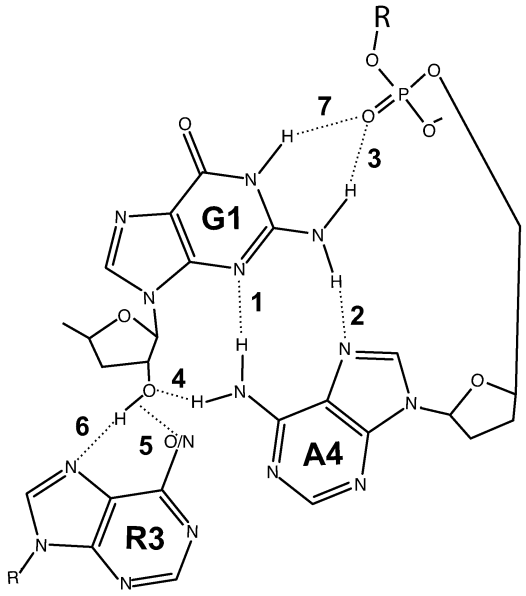

FIGURE 1: Hydrogen-bonding network proposed in GNRA tetraloops by Jucker and Pardi (*28*). Bases are numbered by loop position from 5′ to 3′. Hydrogen bonds are numbered as follows: 1, GN3−AH6; 2, GH2−AN7; 3, GH2−AOP; 4, GO2′−AH6; 5, G 2′OH to position 6 of the R nucleotide; 6, G 2′OH to the N7 of the R nucleotide; 7, GH1−AOP.

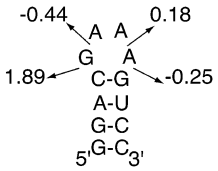

FIGURE 2: Energetic effects of C3 spacers. Values are $\Delta\Delta G^{\circ}_{37}$ (kcal/mol) relative to those of unsubstituted loops. Positions of substitution are numbered from 5′ to 3′. Negative $\Delta\Delta G^{\circ}_{37}$ values indicate stabilization, and positive values indicate destabilization.

measured the effect of C3 spacer insertion before the 5′ end of a r(gGAAAc) loop. In contrast to results for related DNA loops (*21*), a significant thermodynamic penalty still resulted, with a ΔG°_{37} of 1.35 kcal/mol as compared to 1.89 kcal/mol for a CG cbp (Table 1). This finding further supports a GC cbp having an appreciable and therefore interruptible stabilizing interaction with a r(GAAA) loop.

Significant stability of r(gGAAAc) is also consistent with the common occurrence of GC cbp's for r(GNRA) loops phylogenetically. For example, in a survey of prokaryotic 16S rRNAs, GNRA loops were found to contain both GC and CG cbp's with high frequency, whereas UNCG loops were almost always associated with a CG cbp (*5*, *33*).

Loop−Loop Pairwise-Coupling Terms in r(GAAA). The r(GNRA) loop contains a complex network of seven hydrogen bonds, all of which emanate from G1: the two hydrogen bonds of the sheared GA pair, the imino and amino protons of G1 to the RpA phosphate, and a set of three bifurcated hydrogen bonds from the 2′-OH of G1 to the N7 and N6/O6 of R and the H6 of A (Figure 1) (*11*, *26*−*30*). We carried out a series of functional group changes, both alone and in pairwise combinations, and measured the effects on stability.

The first set of mutants was designed to probe the sheared GA pair. Substituting G1 with I eliminates hydrogen bonds 2 and 3 and destabilized the r(c*I*AAAg) mutant, with a ΔG°_{37} of 0.75 ± 0.11 kcal/mol (Table 2). Substituting A4 with I replaces the 6-amino group with a carbonyl and eliminates hydrogen bonds 1 and 4. As expected, r(cGAA*I*g) was also significantly destablized, with a ΔG°_{37} of 1.02 ± 0.19 kcal/mol. Simultaneously, mutating G1 and A4 to inosines, r(c*I*AA*I*g), was destabilizing by 1.62 ± 0.11 kcal/mol, which is close to the sum of the single mutant effects. This gives rise to a small $\delta_{14/23}$ value of −0.2 ± 0.2 kcal/mol.

The effect of the G1 to I change in RNA (0.75 kcal/mol) is much smaller than in DNA (1.45−1.70 kcal/mol) (*9*). Moreover, a δ value close to 0 for the sheared GA pair in RNA contrasts sharply with the results in DNA, where interactions 1 and 2 were found to be highly coupled with a δ_{12} of −0.9 ± 0.2 kcal/mol for d(c*I*C*I*g) and −1.1 ± 0.2 kcal/mol for d(c*I*C*I*Cg).

Loop-Closing Base-Pair Pairwise-Coupling Terms in r(GAAA). Next, we examined coupling between the loop and the cbp using a C3 spacer before G1 to interrupt loop-closing base-pair interactions and inosines to interrupt loop−loop interactions. The single mutants r(c*C3*GAAAg) and r(cGAA*I*g) were destabilizing by 1.89 ± 0.17 and 1.02 ± 0.19 kcal/mol, respectively (Table 2). The double mutant, r(c*C3*GAA*I*g), on the other hand, was destabilized by only 1.59 ± 0.20 kcal/mol. This gives rise to a $\delta_{14/cbp}$ value of −1.3 ± 0.3 kcal/mol. This large, negative δ value indicates that loop−loop and loop-closing base-pair interactions are highly coupled. This is similar to what was found for DNA loops, where d(c*C3*GC*I*g) gave a δ value of −0.9 ± 0.2 kcal/mol.

The single mutant r(c*I*AAAg) was destabilizing by 0.75 ± 0.11 kcal/mol, while the double mutant, r(c*C3I*AAAg), was destabilizing by 2.02 ± 0.19 kcal/mol, which gives a $\delta_{23/cbp}$ of −0.6 ± 0.3 kcal/mol. This significant, negative δ value indicates that these interactions are nonadditive and

Table 2: Thermodynamic Parameters for Folding of Single, Double, and Triple Mutants in the r(cGAAAg) Hairpin

sequence[a]	$\Delta H^{\circ b}$ (kcal mol^{-1})	$\Delta S^{\circ b}$ (cal mol^{-1} K^{-1})	$\Delta G^{\circ}_{37}{}^{b,c}$ (kcal mol^{-1})	T_M (°C)	$\Delta G_A{}^{b,c}$ (kcal mol^{-1})
GAAA	−41.7 ± 0.8	−122.6 ± 2.4	−3.65 ± 0.09	66.8 ± 0.5	
*C3*GAAA	−35.9 ± 1.4	−110.0 ± 4.4	−1.76 ± 0.14	53.0 ± 1.0	1.89 ± 0.17
GAA*I*	−37.7 ± 1.2	−112.9 ± 3.6	−2.63 ± 0.17	60.3 ± 1.2	1.02 ± 0.19
*d*GAAA	−36.8 ± 2.6	−110.0 ± 8.1	−2.67 ± 0.18	61.3 ± 1.5	0.98 ± 0.20
*I*AAA	−39.1 ± 0.6	−116.8 ± 2.0	−2.90 ± 0.06	61.5 ± 0.4	0.75 ± 0.11
*C3I*AAA	−34.9 ± 2.7	−107.1 ± 8.2	−1.63 ± 0.17	53.0 ± 0.7	2.02 ± 0.19
*I*AA*I*	−34.3 ± 1.4	−104.2 ± 4.5	−2.03 ± 0.06	56.5 ± 0.5	1.62 ± 0.11
*C3*GAA*I*	−39.8 ± 3.8	−121.6 ± 11.7	−2.06 ± 0.18	54.0 ± 1.2	1.59 ± 0.20
*d*IAAA	−36.3 ± 1.2	−110.0 ± 3.9	−2.20 ± 0.08	57.0 ± 1.2	1.45 ± 0.12
*d*GAA*I*	−37.3 ± 2.3	−111.6 ± 6.7	−2.66 ± 0.19	60.8 ± 0.5	0.99 ± 0.21
*C3I*AA*I*	−37.5 ± 1.5	−114.1 ± 4.6	−2.05 ± 0.08	55.0 ± 0.8	1.60 ± 0.12
*dI*AA*I*	−37.3 ± 2.4	−112.5 ± 7.3	−2.40 ± 0.l3	58.4 ± 0.9	1.25 ± 0.16

[a] Sequences in bold type is the reference for the sequences below. Changes from reference sequence are italicized. Sequences are listed in order of most penalizing change and grouped by single, double, and triple mutations. [b] Errors are the standard deviations from three or more measurements and were propagated by standard methods. [c] An extra significant figure is provided to avoid round-off error.

are interacting, although not to the same extent, as 14 and the cbp. For comparison, in DNA d(c*C3I*CAg) gave a $\delta_{2/cbp}$ of −1.3 ± 0.2 kcal/mol, which is somewhat larger than that for $\delta_{23/cbp}$.

Examining the Origin of Coupling: The r(cC3IAAIg) Triple Mutant. Next, the three pairwise δ values between C3 and inosine at the first and fourth positions were calculated in the presence of the other site being modified (Figure 3A, Table 3). A term representing the nonadditive effect under these conditions, ${}^{C}\delta_{AB}$, was calculated according to eqs 2a and 2b. First, r(c*I*AAAg), in which interactions 2 and 3 are absent, was used as a reference sequence. The ${}^{23}\delta_{14/cbp}$ value for this case is −1.3 ± 0.2 kcal/mol, which is equal to the $\delta_{14/cbp}$ value of −1.3 ± 0.3 kcal/mol. Next, (c*C3*GAAAg), in which the cbp interaction is absent, was used as a reference sequence. The ${}^{cbp}\delta_{14/23}$ value is −0.1 ± 0.3 kcal/mol, which is close to −0.2 ± 0.2 kcal/mol for $\delta_{14/23}$. Last, r(cGAA*I*g), in which interactions 1 and 4 are absent, was used as a reference sequence. The ${}^{14}\delta_{23/cbp}$ value is −0.6 ± 0.3 kcal/mol, which is the same as the $\delta_{23/cbp}$ value of −0.6 ± 0.3 kcal/mol. Overall, these comparisons indicate that δ_{AB} values are, within error, equal to ${}^{C}\delta_{AB}$ values. This supports direct coupling of the interactions.

Examining the Origin of Coupling: The r(cdIAAIg) Triple Mutant. Next, we measured several other ${}^{C}\delta_{AB}$ values to see if they also couple directly (Figure 3B, Table 3). These mutants involve a change of the 2′-OH of G1 to 2′-H. First, we consider single and double mutants involving changes at the 2′-OH of G1. The single mutants r(c*d*GAAAg) and r(cGAA*I*g) were destabilized by 0.98 ± 0.20 and 1.02 ± 0.19 kcal/mol, respectively, while the double mutant, r(c*d*-GAA*I*g), was destabilized by only 0.99 ± 0.21 kcal/mol. This gives rise to a large, negative $\delta_{14/456}$ value of −1.0 ± 0.3 kcal/mol, consistent with complete nonadditivity. These changes are redundant in that they both interrupt interaction 4, providing a possible molecular basis for complete nonadditivity.

The single mutant r(c*I*AAAg) was destabilizing by 0.75 ± 0.11 kcal/mol, while the double mutant, r(c*dI*AAAg), was destabilizing by 1.45 ± 0.12 kcal/mol. This gives rise to a small $\delta_{23/456}$ of −0.3 ± 0.2 kcal/mol, indicating that these interactions are largely additive.

Next, we consider triple mutant cycles involving changes at the 2′-OH of G1. First, r(c*d*GAAAg), in which interactions

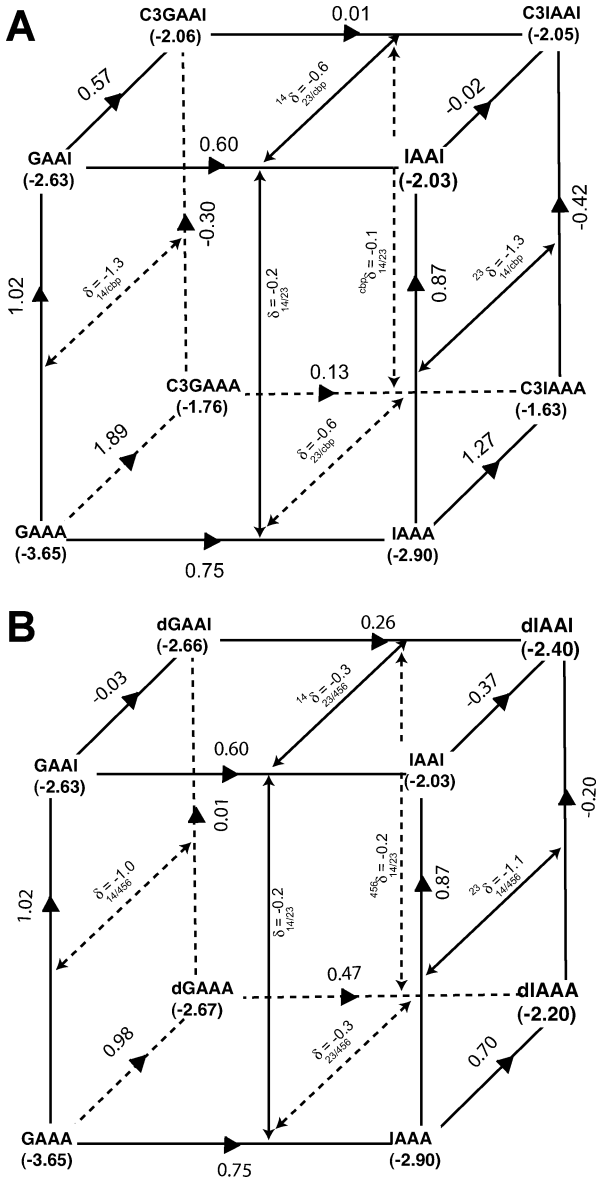

FIGURE 3: Thermodynamic cubes for triple mutant cycles. (A) Cube for r(c*C3I*AA*I*g). (B) Cube for r(c*dI*AA*I*g). Experimentally measured ΔG°_{37} values are at the vertices of the cube, and the free-energy change associated with a mutation is given along an edge. δ values are provided on each of the six faces. Pairs of δ values on two opposite faces of the cube are approximately the same, consistent with direct coupling. δ values are provided in Table 3.

Table 3: Free-Energy Parameters and δ Values for Double Mutant Cycles in the r(cGAAAg) Hairpin

	interaction probed[a]	ΔG_A[b] (kcal/mol)	ΔG_B[c] (kcal/mol)	ΔG_{AB}[d] (kcal/mol)	ΔG_{AB}[e] (if additive) (kcal/mol)	δ_{AB}	δ_{AB}[f] (kcal/mol)
*C3I*AAA	23/cbp	0.75 ± 0.11	1.89 ± 0.17	2.02 ± 0.19	2.64 ± 0.18	$\delta_{23/cbp}$	−0.6 ± 0.3
*I*AA*I*	14/23	1.02 ± 0.19	0.75 ± 0.11	1.62 ± 0.11	1.77 ± 0.20	$\delta_{14/23}$	−0.2 ± 0.2
*C3*GAA*I*	14/cbp	1.02 ± 0.19	1.89 ± 0.17	1.59 ± 0.20	2.91 ± 0.24	$\delta_{14/cbp}$	−1.3 ± 0.3
*d*IAAA	23/456	0.75 ± 0.11	0.98 ± 0.20	1.45 ± 0.12	1.73 ± 0.21	$\delta_{23/456}$	−0.3 ± 0.2
*d*GAA*I*	14/456	1.02 ± 0.19	0.98 ± 0.20	0.99 ± 0.21	2.00 ± 0.26	$\delta_{14/456}$	−1.0 ± 0.3
	interaction probed[a,g]	$^{C}\Delta G_A$[g] (kcal/mol)	$^{C}\Delta G_B$[g] (kcal/mol)	$^{C}\Delta G_{AB}$[g] (kcal/mol)	$^{C}\Delta G_{AB}$[g] (if additive) (kcal/mol)	$^{C}\delta_{AB}$[g]	$^{C}\delta_{AB}$[g] (kcal/mol)
*C3I*AA*I*	^{14}23/cbp	0.60 ± 0.18	0.57 ± 0.25	0.58 ± 0.19	1.17 ± 0.25	$^{14}\delta_{23/cbp}$	−0.6 ± 0.3
	^{23}14/cbp	0.87 ± 0.08	1.27 ± 0.18	0.85 ± 0.1	2.14 ± 0.19	$^{23}\delta_{14/cbp}$	−1.3 ± 0.2
	^{cbp}14/23	−0.30 ± 0.23	0.13 ± 0.22	−0.29 ± 0.16	−0.17 ± 0.28	$^{cbp}\delta_{14/23}$	−0.1 ± 0.3
*d*IAA*I*	^{456}14/23	0.01 ± 0.26	0.47 ± 0.20	0.27 ± 0.22	0.48 ± 0.27	$^{456}\delta_{14/23}$	−0.2 ± 0.3
	^{23}14/456	0.87 ± 0.08	0.70 ± 0.10	0.50 ± 0.14	1.57 ± 0.12	$^{23}\delta_{14/456}$	−1.1 ± 0.2
	^{14}23/456	0.60 ± 0.18	−0.03 ± 0.25	0.23 ± 0.21	0.57 ± 0.26	$^{14}\delta_{23/456}$	−0.3 ± 0.3

[a] "Interaction probed" refers to Figure 1. Cbp represents a C3 spacer. [b] ΔG_A values are for breaking the first of the interactions probed listed. [c] ΔG_B values are for breaking the second of the interactions probed listed. [d] ΔG_{AB} values are for both modifications in a single oligonucleotide. [e] Values are the sum of ΔG_A and ΔG_B. [f] δ values were calculated as the differences between columns 5 and 6, and errors were propagated from eq 1a. [g] Superscript refers to quantities measured in the background of a third change.

4−6 are absent, was used as a reference sequence. The value for $^{456}\delta_{14/23}$ is −0.2 ± 0.3 kcal/mol, which is equal to the $\delta_{14/23}$ value of −0.2 ± 0.2. Next, r(cGAA*I*g), in which interactions 1 and 4 are absent, was used as a reference sequence. The $^{14}\delta_{23/456}$ value is −0.3 ± 0.3 kcal/mol, which is equal to the $\delta_{23/456}$ value of −0.3 ± 0.2 kcal/mol. Last, r(c*I*AAAg), in which interactions 2 and 3 are absent, was used as a reference sequence. The $^{23}\delta_{14/456}$ value is −1.1 ± 0.2 kcal/mol, which is the same as the $\delta_{14/456}$ value of −1.0 ± 0.3 kcal/mol. Overall, these three comparisons, like the three above, support direct coupling in that δ_{AB} values are, within error, equal to $^{C}\delta_{AB}$ values.

DISCUSSION

There have been extensive structural and energetic studies on unusually stable RNA and DNA hairpin loop sequences (*6*−*9*, *22*, *34*, *35*). The goal of the present study was to delineate similarities and differences in the cooperativity of RNA and DNA hairpin loop folding. We chose to study a r(GNRA) loop sequence because it shares extensive similarities with the DNA loop sequences, d(cGNAg) and d(cGNABg), including unusual thermodynamic stability, a sheared GA pair, and a stable loop-closing base-pair interaction (*11*, *17*, *18*, *26*−*30*). In addition, both r(GNRA) and d(cGNAg) loops have structures solved.

On one hand, we did find many qualitative similarities between the behaviors of the two loops. Both systems were significantly destabilized by modifying interactions 1 and 2 in the sheared GA pairing. In addition, introduction of a GC cbp significantly destabilized both systems, as did introduction of a C3 spacer before position 1 of the loop but not elsewhere in the loop. The latter result supports expandability of r(GNRA) loops at all positions except before the G. As illustrated and summarized by Abramovitz and Pyle, numerous experiments in the literature also support this model (*36*). However, a large number of differences were found between the two systems.

Differences can be divided into effects of single changes, double mutant cycles, and triple mutant cycles. Thermodynamic effects associated with single functional group changes were in general much larger in DNA than in RNA. This came as somewhat of a surprise because deletion of a single functional group in RNA eliminates more interactions than in DNA. For example, deletion of the amino group of G1 deletes interactions 2 and 3 in RNA but only interaction 2 in DNA, yet the penalty in RNA is only 0.75 kcal/mol versus a 1.45−1.70 kcal/mol penalty in DNA (*19*, *20*).

Single mutant effects have also been reported on a different r(cGNRAg) loop, r(cGCAAg), by SantaLucia and coworkers, although at an ionic strength of 100 mM NaCl (*34*). They found that a G to I substitution was destabilizing by a ΔG°_{37} of 0.54 kcal/mol, similar to our value of 0.75 ± 0.11 kcal/mol. A purine substitution at the last position of the loop, which eliminates hydrogen bonds 1 and 4, was destabilizing by 0.28 kcal/mol, whereas in our study an inosine substitution at the last position, which also eliminates hydrogen bonds 1 and 4, was destabilizing by 1.02 ± 0.19 kcal/mol. The differences in these two values may be due to the addition of the carbonyl group upon substituting A with I. Substitution of a deoxy G at position 1 of the loop had a ΔG°_{37} of 0 in their studies but was found herein to be destabilizing by 0.98 ± 0.20 kcal/mol. The $\Delta\Delta G^\circ_{37}$ differences may be due to the higher ionic strength in their system leading to a higher T_M (68.1 versus 61.3 °C) and therefore a longer extrapolation back to 37 °C or perhaps to an indirect influence of nucleotide 2 on the loop. Overall, the single mutant results are in qualitative agreement between the two studies.

In general, coupling free-energy terms for double mutant cycles were also significantly larger in DNA than RNA. For example, the free-energy coupling term between interactions 1 and 2 in the loop was near 0 in RNA but approximately −1.0 kcal/mol in d(cGNAg) and d(cGNABg) loops (*19*, *20*). Likewise, the coupling term between the cbp and interaction 23 was only −0.6 kcal/mol in RNA but −1.3 kcal/mol in DNA.

Last, triple mutant cycles showed markedly different behavior in RNA than in DNA. In particular, for the six sets of δ's in r(GAAA), δ_{AB} values were, within error, equal to $^{C}\delta_{AB}$ values (Table 3). For DNA, on the other hand, δ_{AB} values were considerably larger (by $\approx$ −0.6 kcal/mol) than $^{C}\delta_{AB}$ values. Thus, in RNA, the interactions are directly

coupled, whereas in DNA they are indirectly coupled. The direct coupling effect in RNA is especially noticeable because most of the pairwise interactions are coupled, yet the coupling constant for each pair does not depend on the configuration of the third site (*23*). For example, the cbp is coupled to interactions 14 and 23, $\delta_{14/cbp} = -1.3$ kcal/mol and $\delta_{23/cbp} = -0.6$ kcal/mol, yet $\delta_{14/cbp} \approx {}^{23}\delta_{14/cbp}$ and $\delta_{23/cbp} \approx {}^{14}\delta_{23/cbp}$.

What might be the molecular origins of the different cooperative behavior of these two systems? For the DNA loops, deletion of a given functional group was found to be progressively less penalizing as other interactions were removed (*19*). Given this scenario, one might have expected in RNA, where there are more stabilizing interactions than in DNA, that deletion of a given functional group would have been the most penalizing. Clearly, this is not the case. For DNA, the system, having only three interactions, was minimally stable to begin with. Thus, loss of any of the interactions simply could not be tolerated and led to significant weakening of the others. This molecular scenario is consistent with a concerted change in which the three interactions are linked through a conformational change. The observation that no loop−loop interactions could be detected in a GC cbp background supports a high degree of cooperativity from the DNA loops (*20*).

In RNA, the smaller effects of single, double, and triple mutants suggest that a given mutant is simply not as destabilized as in DNA, and therefore the overall the folding of the RNA loop is less cooperative. One molecular scenario is that RNA is able to compensate for the loss in functionality upon mutation by "opportunistic" hydrogen bonding and stacking elsewhere in the loop. Given the presence of multiple 2′OHs, there is a much greater number of hydrogen-bonding groups available for interaction in RNA than in DNA. This is also reflected in the hydration patterns of GNRA tetraloops. High-resolution crystal structures (at least 2.0 Å resolution or better) reveal seven structural water molecules, primarily lining the major groove of the loop (*30*), which could presumably partake in stabilizing hydrogen bonds. Also, the purine at position 3 of the RNA loop allows for more stacking interactions throughout the system than in DNA, where the purine is absent. The possibility for more interactions may explain why, for example, a GC cbp is relatively stabilized in a r(GNRA) loop. If small DNA hairpins with indirect coupling can be compared to a three-legged stool (*19*), RNA hairpins with direct coupling might be likened to a eight-legged stool. With the greater number of legs (interactions), loss of any one leg does not lead to collapse of the entire stool (structure). Similar arguments have been used to explain nonadditivity in enzymes (*25*).

Greater cooperativity often manifests itself in greater specificity (*23*, *25*). In the case of the hairpin loops, the DNA loops were generally less tolerant of mutations and therefore more specific than the RNA loops. The greater mutability or plasticity of the RNA loops has implications for RNA evolution. Shuster and Fontana have emphasized the importance of neutral genetic drift in evolution in which RNA can mutate with no significant change in its function, searching for the appropriate context from which to evolve a new function (*37*). The observation that loop−loop interactions continue to couple similarly despite changes elsewhere in the sequence is consistent with modularity in RNA motifs as simple as a tetraloop. The ability of RNA to tolerate changes in its loops with less catastrophic results than DNA might enhance the ability of RNA to evolve, supporting RNA as a potentially important prebiotic polymer.

CONCLUSIONS

In summary, small RNA hairpins were found to exhibit less cooperativity than their DNA counterparts. This conclusion is based on three properties of the DNA loops: larger energetic effects of single functional group substitutions, greater nonadditivity of double mutant cycles, and indirect coupling, as revealed through triple mutant cycles. All three of these observations support a concerted folding of DNA loops as opposed to a modular, stepwise folding of RNA loops. Importantly, this finding does not disagree with the importance of cooperativity in Mg^{2+} binding to RNA tertiary structures, which is driven by electrostatic forces rather than hydrogen bonding and stacking of the nucleobases (*38*). The molecular basis for the differences in cooperativity between RNA and DNA loops is likely due to the extra hydrogen bonding afforded by the 2′-hydroxyls and seven waters of hydration, as well as the extra stacking afforded by the purine at position 3 of the RNA loops. The lesser cooperativity in RNA loops may lead to a more robust loop that can withstand mutations without a severe loss in stability. These differences may enhance the ability of RNA to evolve.

REFERENCES

1. Varani, G. (1995) Exceptionally stable nucleic acid hairpins, *Annu. Rev. Biophys. Biophys. Chem. 24*, 379−404.
2. Tinoco, I., Jr., and Bustamante, C. (1999) How RNA folds, *J. Mol. Biol. 293*, 271−281.
3. Babitzke, P., Schaak, J. E., Yakhnin, A. V., and Bevilacqua, P. C. (2003) Role of RNA structure in transcription attenuation in *Bacillus subtilis*: The trpEDCFBA operon as a model system, *Methods Enzymol. 371*, 392−404.
4. Winkler, W. C., and Breaker, R. R. (2003) Genetic control by metabolite-binding riboswitches, *ChemBioChem 4*, 1024−1032.
5. Woese, C. R., Winker, S., and Gutell, R. R. (1990) Architcture of ribosomal RNA: Constraints on the sequence of "tetra-loops", *Proc. Natl. Acad. Sci. U.S.A. 87*, 8467−8471.
6. Antao, V. P., Lai, S. Y., and Tinoco, I., Jr. (1991) A thermodynamic study of unusually stable RNA and DNA hairpins, *Nucleic Acids Res. 19*, 5901−5905.
7. Shu, Z., and Bevilacqua, P. C. (1999) Isolation and characterization of thermodynamically stable and unstable RNA hairpins from a triloop combinatorial library, *Biochemistry 38*, 15369−15379.
8. Proctor, D. J., Schaak, J. E., Bevilacqua, J. M., Falzone, C. J., and Bevilacqua, P. C. (2002) Isolation and characterization of a family of stable RNA tetraloops with the motif YNMG that participate in tertiary interactions, *Biochemistry 41*, 12062−12075.
9. Nakano, M., Moody, E. M., Liang, J., and Bevilacqua, P. C. (2002) Selection for thermodynamically stable DNA tetraloops using temperature gradient gel electrophoresis reveals four motifs: d(cGNNAg), d(cGNABg), d(cCNNGg), and d(gCNNGc), *Biochemistry 41*, 14281−14292.
10. Cheong, C., Varani, G., and Tinoco, I., Jr. (1990) Solution structure of an unusually stable RNA hairpin, 5′GGAC(UUCG)GUCC, *Nature 346*, 680−682.
11. Heus, H. A., and Pardi, A. (1991) Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops, *Science 253*, 191−194.
12. Allain, F. H. T., and Varani, G. (1995) Divalent metal ion binding to a conserved wobble pair defining the upstream site of cleavage of group I self-splicing introns, *Nucleic Acids Res. 23*, 341−350.
13. Du, Z., Yu, J., Andino, R., and James, T. L. (2003) Extending the family of UNCG-like tetraloop motifs: NMR structure of a CACG tetraloop from coxsackievirus B3, *Biochemistry 42*, 4373−4383.
14. Senior, M. M., Jones, R. A., and Breslauer, K. J. (1988) Influence of loop residues on the relative stabilities of DNA hairpin structures, *Proc. Natl. Acad. Sci. U.S.A. 85*, 6242−6246.

15. Paner, T. M., Amaratunga, M., Doktycz, M. J., and Benight, A. S. (1990) Analysis of melting transitions of the DNA hairpins formed from the oligomer sequences d[GGATAC(X)4GTATCC] (X = A, T, G, C), *Biopolymers 29*, 1715−1734.
16. Vallone, P. M., Paner, T. M., Hilario, J., Lane, M. J., Faldasz, B. D., and Benight, A. S. (1999) Melting studies of short DNA hairpins: Influence of loop sequence and adjoining base pair identity on hairpin thermodynamic stability, *Biopolymers 50*, 425−442.
17. Hirao, I., Kawai, G., Yoshizawa, S., Nishimura, Y., Ishido, Y., Watanabe, K., and Miura, K. (1994) Most compact hairpin-turn structure exerted by a short DNA fragment, d(GCGAAGC) in solution: An extraordinarily stable structure resistant to nucleases and heat, *Nucleic Acids Res. 22*, 576−582.
18. Yoshizawa, S., Kawai, G., Watanabe, K., Miura, K., and Hirao, I. (1997) GNA trinucleotide loop sequences producing extraordinarily stable DNA minihairpins, *Biochemistry 36*, 4761−4767.
19. Moody, E. M., and Bevilacqua, P. C. (2003) Folding of a stable DNA motif involves a highly cooperative network of interactions, *J. Am. Chem. Soc. 125*, 16285−16293.
20. Moody, E. M., and Bevilacqua, P. C. (2004) Structural and energetic consequences of expanding a highly cooperative stable DNA hairpin loop, *J. Am. Chem. Soc.*, in press.
21. Moody, E. M., and Bevilacqua, P. C. (2003) Thermodynamic coupling of the loop and stem in unusually stable DNA hairpins closed by CG base pairs, *J. Am. Chem. Soc. 125*, 2032−2033.
22. Sandusky, P., Wooten, E. W., Kurochkin, A. V., Kavanaugh, T., Mandecki, W., and Zuiderweg, E. R. (1995) Occurrence, solution structure, and stability of DNA hairpins stabilized by a GA/CG helix unit, *Nucleic Acids Res. 23*, 4717−4725.
23. Di Cera, E. (1998) Site-specific thermodynamics: understanding cooperativity in molecular recognition, *Chem. Rev. 98*, 1563−1592.
24. Klostermeier, D., and Millar, D. P. (2002) Energetics of hydrogen bond networks in RNA: Hydrogen bonds surrounding G+1 and U42 are the major determinants for the tertiary structure stability of the hairpin ribozyme, *Biochemistry 41*, 14095−14102.
25. Kraut, D. A., Carroll, K. S., and Herschlag, D. (2003) Challenges in enzyme mechanism and energetics, *Annu. Rev. Biochem. 72*, 517−571.
26. Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994) Three-dimensional structure of a hammerhead ribozyme, *Nature 372*, 68−74.
27. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) RNA tertiary structure mediation by adenosine platforms, *Science 273*, 1696−1699.
28. Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H., and Pardi, A. (1996) A network of heterogeneous hydrogen bonds in GNRA tetraloops, *J. Mol. Biol. 264*, 968−980.
29. Correll, C. C., Wool, I. G., and Munishkin, A. (1999) The two faces of the *Escherichia coli* 23 S rRNA sarcin/ricin domain: The structure at 1.11 Å resolution, *J. Mol. Biol. 292*, 275−287.
30. Correll, C. C., and Swinger, K. (2003) Common and distinctive features of GNRA tetraloops based on a GUAA tetraloop structure at 1.4 Å resolution, *RNA 9*, 355−363.
31. Proctor, D. J., Kierzek, E., Kierzek, R., and Bevilacqua, P. C. (2003) Restricting the conformational heterogeneity of RNA by specific incorporation of 8-bromoguanosine, *J. Am. Chem. Soc. 125*, 2390−2391.
32. Horovitz, A., and Fersht, A. R. (1990) Strategy for analysing the cooperativity of intramolecular interactions in peptides and proteins, *J. Mol. Biol. 214*, 613−617.
33. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pande, N., Shang, Z., Yu, N., and Gutell, R. R. (2002) The Comparative RNA Web (CRW) Site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs, *Bioinformatics 3*, 2.
34. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1992) Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin, *Science 256*, 217−219.
35. Hirao, I., Nishimura, Y., Tagawa, Y., Watanabe, K., and Miura, K. (1992) Extraordinarily stable mini-hairpins: Electrophoretical and thermal properties of the various sequence variants of d(GCGAAAGC) and their effect on DNA sequencing, *Nucleic Acids Res. 20*, 3891−3896.
36. Abramovitz, D. L., and Pyle, A. M. (1997) Remarkable morphological variability of a common RNA folding motif: The GNRA tetraloop-receptor interaction, *J. Mol. Biol. 266*, 493−506.
37. Fontana, W., and Schuster, P. (1998) Continuity in evolution: On the nature of transitions, *Science 280*, 1451−1455.
38. Misra, V. K., and Draper, D. E. (1998) On the role of magnesium ions in RNA stability, *Biopolymers 48*, 113−135.

BI049350E