J|A|C|S
ARTICLES
Published on Web 12/06/2003

Folding of a Stable DNA Motif Involves a Highly Cooperative Network of Interactions

Ellen M. Moody and Philip C. Bevilacqua*

Contribution from the Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802

Received October 6, 2003; E-mail: pcb@chem.psu.edu

Abstract: Hairpins are structural elements that play important roles in the folding and function of RNA and DNA. The extent of cooperativity in folding is an important aspect of the RNA folding problem. We reasoned that an investigation into the origin of cooperativity might be best carried out on a stable nucleic acid system with a limited number of interactions, such as a stable DNA hairpin loop. The stable d(cGNAg) hairpin loop motif (closing base pair in lower case; loop in upper case; N = A, C, G, or T) is stabilized through only three interactions: two loop−loop hydrogen bonds in a sheared GA base pair and a loop−closing base pair interaction. Herein, we investigate this network of interactions and test whether the loop−loop and loop−closing base pair interactions communicate. Thermodynamic measurements of nucleotide analogue substituted oligonucleotides were used to probe the additivity of the interactions. On the basis of double mutant cycles, all interactions were found to be nonadditive and interdependent, suggesting that loop−loop and loop−closing base pair interactions form in a highly cooperative manner. When double mutant cycles were repeated in the absence of the other interaction, nonadditivity was significantly reduced suggesting that coupling is indirect and requires all three interactions in order to be optimal. A cooperative network of interactions helps explain the structural and energetic bases of stability in certain DNA hairpins and paves the way for similar studies in more complex nucleic acid systems.

Hairpins are common secondary structural elements in RNA and play important roles in initiating RNA folding, forming tertiary structures, and interacting with proteins.[1,2] Double stranded DNA can transiently form cruciforms in which the individual strands form hairpins.[3] These structures have been implicated in telomere replication, deletion mutations, and V(D)J recombination.[4−6] In addition, DNA hairpins can play a part in regulating replication and transcription.[1,7] The importance of DNA hairpins in these biological processes is part of the motivation for studying their structures and thermodynamics.

Research on the structure and stability of hairpins has concentrated mainly on RNA, where they are most common. Stable RNA triloop and tetraloop hairpin sequences have been reported,[8−11] and numerous structures have been solved by NMR and X-ray methods.[12−16] There have also been a number of studies on the structure and stability of DNA hairpin loops. As in RNA, the stability of DNA hairpins depends on the base composition of the loop as well as the closing base pair.[17−20] The stability of the stem can be accurately predicted using nearest-neighbor rules,[21] and studies on model hairpins and combinatorial libraries are providing further rules for DNA loop stability.[8,20]

It has also been observed in both DNA and RNA that for certain hairpin loop sequences a CG closing base pair provides much greater thermodynamic stability ($\Delta\Delta G^\circ_{37}$ 2−3 kcal/mol) than predicted by Watson−Crick base pairing alone.[8,10,11,20,22] Recent work in our lab demonstrated a large thermodynamic penalty for three carbon spacer (C3) insertion between the 5′ end of a stable hairpin loop and a CG closing base pair, with much smaller penalties for insertion throughout the rest of the loop or in hairpins with other closing base pairs.[23] These data support a stabilizing interaction between the G at position one of the loop and the CG closing base pair. Although the exact

(1) Varani, G. *Annu. Rev. Biophys. Biophys. Chem.* **1995**, *24*, 379−404.
(2) Tinoco, I., Jr.; Bustamante, C. *J. Mol. Biol.* **1999**, *293*, 271−281.
(3) Lilley, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 6468−6472.
(4) Choi, K. H.; Choi, B. S. *Biochim. Biophys. Acta* **1994**, *1217*, 341−344.
(5) Bzymek, M.; Lovett, S. T. *Genetics* **2001**, *158*, 527−540.
(6) Gellert, M. *Annu. Rev. Biochem.* **2002**, *71*, 101−132.
(7) Glucksmann-Kuis, M. A.; Dai, X.; Markiewicz, P.; Rothman-Denes, L. B. *Cell* **1996**, *84*, 147−154.
(8) Antao, V. P.; Lai, S. Y.; Tinoco, I., Jr. *Nucleic Acids Res.* **1991**, *19*, 5901−5905.
(9) Antao, V. P.; Tinoco, I., Jr. *Nucleic Acids Res.* **1992**, *20*, 819−824.
(10) Shu, Z.; Bevilacqua, P. C. *Biochemistry* **1999**, *38*, 15369−15379.
(11) Proctor, D. J.; Schaak, J. E.; Bevilacqua, J. M.; Falzone, C. J.; Bevilacqua, P. C. *Biochemistry* **2002**, *41*, 12062−12075.
(12) Cheong, C.; Varani, G.; Tinoco, I., Jr. *Nature* **1990**, *346*, 680−682.
(13) Heus, H. A.; Pardi, A. *Science* **1991**, *253*, 191−194.
(14) Allain, F. H. T.; Varani, G. *Nucleic Acids Res.* **1995**, *23*, 341−350.
(15) Chou, S. H.; Chin, K. H.; Wang, A. H. *Nucleic Acids Res.* **2003**, *31*, 2461−2474.
(16) Du, Z.; Yu, J.; Andino, R.; James, T. L. *Biochemistry* **2003**, *42*, 4373−4383.
(17) Senior, M. M.; Jones, R. A.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6242−6246.
(18) Paner, T. M.; Amaratunga, M.; Doktycz, M. J.; Benight, A. S. *Biopolymers* **1990**, *29*, 1715−1734.
(19) Vallone, P. M.; Paner, T. M.; Hilario, J.; Lane, M. J.; Faldasz, B. D.; Benight, A. S. *Biopolymers* **1999**, *50*, 425−442.
(20) Nakano, M.; Moody, E. M.; Liang, J.; Bevilacqua, P. C. *Biochemistry* **2002**, *41*, 14281−14292.
(21) SantaLucia, J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1460−1465.
(22) Sandusky, P.; Wooten, E. W.; Kurochkin, A. V.; Kavanaugh, T.; Mandecki, W.; Zuiderweg, E. R. *Nucleic Acids Res.* **1995**, *23*, 4717−4725.
(23) Moody, E. M.; Bevilacqua, P. C. *J. Am. Chem. Soc.* **2003**, *125*, 2032−2033.

nature of the interaction has yet to be determined, the imino proton and carbonyl of the first G of the loop appear to contribute to the effect.[23] Importantly, this study showed that even a simple secondary structure such as a DNA hairpin can be composed of multiple types of interactions, loop−loop and loop−closing base pair, opening the possibility for cooperativity in their formation.

We are interested in the RNA folding problem, which is a complex problem typically involving hierarchical assembly of preformed secondary structures into a functional tertiary structure.[2,24] An outstanding problem in RNA folding and catalysis is the contribution of cooperativity and nonadditivity to the overall energetic picture of folding. Kraut and co-workers have likened extensive nonadditivity in folding and catalysis to a house with poor structural integrity, wherein removal of a single beam (= an interaction) leads to loss of the entire structure.[25] Using similar logic, we reasoned that a stable DNA hairpin loop, rather than a stable RNA loop, might provide an ideal proving ground for such a concept since it has fewer stabilizing interactions. In particular, the d(cGNAg) motif is like the r(cGNRAg) motif, where "R" is purine, in that both contain a sheared GA base pair and a 2−3 kcal/mol bonus contribution of the closing base pair to ΔG°_{37} (ref 23 and E. M. Moody, J. C. Feerrar, W. G. O'Connell, and P. C. Bevilacqua, unpublished data); however, the r(cGNRAg) motif is further stabilized by four additional hydrogen bonds: (1) a G amino proton to the RpA phosphate, (2) the G imino proton to the RpA phosphate, (3) a bifurcated hydrogen bond from the G 2′-OH to the N7 and N6/O6 of the purine, and (4) the amino proton of A to the G 2′-OH.[13,26] The fewer number of interactions in the DNA loop might make each one essential to overall folding, thus allowing insight into the fundamental issue of cooperativity in nucleic acid energetics and possible demonstration of maximal nonadditivity.

Herein, we investigate loop−loop and loop−closing base pair interactions in a stable d(cGNAg) hairpin loop. Communication between the loop−loop and loop−closing base pair interactions is investigated by probing the additivity of free energies in double and triple mutant cycles. Results support a highly cooperative system in which all loop−loop and loop−closing base pair interactions are essential for stable folding.

Materials and Methods

Preparation of DNA. Solid phase synthesized DNA was deblocked and desalted by the manufacturer (IDT). Oligonucleotides were dialyzed against $P_{10}E_{0.1}$ [= 10 mM sodium phosphate and 0.1 mM Na_2EDTA, pH 7.0] for 3−4 h using a Gibco BRL microdialysis apparatus with a flow rate of ∼1 L/h and a Spectra-Por membrane with a molecular weight cutoff of 1000. DNA oligonucleotides containing 7-deazaguanosine, 7-deazaadenosine, and purine were synthesized by either the Nucleic Acids Facility at Pennsylvania State University or the HHMI-Keck Facility at Yale University using reagents from Glen Research. These oligonucleotides were dialyzed or sep-packed before melting and were stored in $P_{10}E_{0.1}$ at −20 °C. Electrospray ionization mass spectrometry was used to confirm the molecular weights of oligonucleotides with modifications. For representative sequences, the DNA was end-labeled by phosphorylation with [γ-^{32}P] ATP and polynucleotide kinase (New England Biolabs), and the purity was confirmed by observation of a single band by denaturing PAGE (in 8.0 M urea at 50 °C); to prevent secondary structure on the gel, the DNA was glyoxylated before electrophoresis.[20,27] All DNA had the general sequence 5′-d(ggaX$L_1L_2L_3$X′tcc), where X and X′ are complementary nucleotides forming the closing base pair and "L" indicates a loop nucleotide. Most oligonucleotides have the same three beginning (5′gga) and ending (tcc3′) nucleotides; in these cases, only the loop and closing base pair 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, using a Gilford Response II spectrophotometer and a heating rate of 0.5 or 1 °C/min. Melts were performed in 1, 5, or 10 mm path length cuvettes, at a minimum of three different strand concentrations ranging from 1 to 70 μM. Melts were found to be independent of strand concentration, consistent with the hairpin conformation.[20] At the start of each experiment, the hairpin was denatured by heating to 90 °C in melting buffer. Data for forward and reverse melts were similar, consistent with reversibility of the melting transition. Concentrations were calculated using absorbance values at 90 °C and extinction coefficients from a nearest-neighbor analysis.[28,29] Data were the average of three or more melts. Thermodynamic parameters were obtained by nonlinear least-squares fitting to a two-state model with sloping baselines using a set of parametric equations defined in Kaleidagraph v3.5 (Synergy software).[30]

Analysis of Double Mutant Cycles. The additivity of ΔG°_{37} values for double mutant cycles was analyzed similarly to previously described.[31−33] This approach can be illustrated using a thermodynamic box in which the ΔG°_{37} values of the unmodified sequence (M_{00}), two single mutants (M_{10} and M_{01}), and the double mutant (M_{11}) comprise the four corners of a box (see Figure 4A as an example). (The subscripts on "M" are placeholders for sites A and B, respectively, with "0" being the unmodified configuration and "1" being the mutant configuration.) Moving vertically from the bottom left-hand corner of the box gives the free energy change associated with mutation A, ΔG_A; moving horizontally from the bottom left-hand corner gives the free energy change associated with mutation B, ΔG_B; and moving diagonally from the bottom left-hand corner of the box gives the effect of both mutations, ΔG_{AB}. (Note that these three values are simply $\Delta\Delta G^\circ_{37}$s referenced to the unmodified sequence.) The free energy change associated with mutation A in the presence of B, $^B\Delta G_A$, and that associated with mutation B in the presence of A, $^A\Delta G_B$, are given along the other two edges of the box. A coupling free energy term, δ_{AB}, is defined as the magnitude of the nonadditive effect between mutations A and B, such that $^B\Delta G_A = \Delta G_A + \delta_{AB}$, and was calculated according to the following equations,

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

Nonadditivity is derived from the idea that if the two changes were additive, i.e., they did not communicate, ΔG_{AB} would be the sum of ΔG_A and ΔG_B. Equation 1a can be rearranged to give eq 1b which emphasizes this point,

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

(24) Brion, P.; Westhof, E. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 113−137.
(25) Kraut, D. A.; Carroll, K. S.; Herschlag, D. *Annu. Rev. Biochem.* **2003**, *72*, 517−571.
(26) Jucker, F. M.; Heus, H. A.; Yip, P. F.; Moors, E. H.; Pardi, A. *J. Mol. Biol.* **1996**, *264*, 968−980.
(27) Latham, J. A.; Zaug, A. J.; Cech, T. R. *Methods Enzymol.* **1990**, *181*, 558−569.
(28) Borer, P. N. In *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*, 3rd ed.; Fasman, G. D., Ed.; CRC Press: Cleveland, OH, 1975; Vol. I, p 597.
(29) Richards, E. G. In *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*, 3rd ed.; Fasman, G. D., Ed.; CRC Press: Cleveland, OH, 1975; Vol. I, p 597.
(30) Nakano, S.; Cerrone, A. L.; Bevilacqua, P. C. *Biochemistry* **2003**, *42*, 2982−2994.
(31) Horovitz, A.; Fersht, A. R. *J. Mol. Biol.* **1990**, *214*, 613−617.
(32) Di Cera, E. *Chem. Rev.* **1998**, *98*, 1563−1592.
(33) Klostermeier, D.; Millar, D. P. *Biochemistry* **2002**, *41*, 14095−14102.

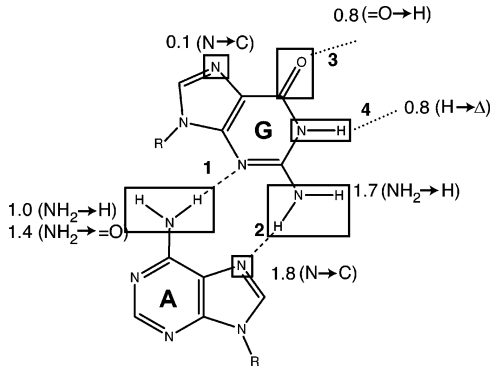

Figure 1. Thermodynamic consequences of single functional group changes in d(cGCAg) hairpins. Values are $\Delta\Delta G^\circ_{37}$ values for substitutions with purine derivatives (see Table 1); the functional group substitutions are also provided. Certain substitutions with pyrimidines were also performed (Table 1). Hydrogen bonds 1 and 2 (dashed lines) are shown, as are two stacking interactions, 3 and 4 (dotted lines), from the G to the CG closing base pair of the stem, which lies below based on the structure of d(cGCAg).[35] Dashed lines are not used for interactions 3 and 4 because it is not known if they involve hydrogen bonds.[23]

A negative δ_{AB} value occurs when the effect of a mutation is diminished in the presence of another, and signifies positive coupling between the functional groups.[32] A positive value of δ_{AB}, on the other hand, occurs when the effect of a mutation is enhanced in the presence of another, and signifies negative coupling. A δ_{AB} value of 0 supports the absence of coupling. We term a double mutant as "completely nonadditive" if $\delta_{AB} = -\Delta G_A = -\Delta G_B = -\Delta G_{AB}$.

In a similar fashion, double mutant cycles were repeated in the presence of a third mutation to test whether coupling was direct or indirect.[31,32] Two equations can be defined for this case, with the following example in the presence of a mutational configuration at site C.

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

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

As Di Cera points out, if the coupling between two sites, A and B, is direct, then δ_{AB} should equal $^{C}\delta_{AB}$; otherwise the coupling requires one or more additional sites to be established and is referred to as indirect.[31,32]

Propagation of Errors. Errors in ΔG°_{37}, ΔH°, and ΔS° values are the standard deviations of three or more melts. Errors in ΔG°_{37} are less than in ΔH° and ΔS° because of enthalpy−entropy compensation.[34] Errors in $\Delta\Delta G^\circ_{37}$ values are the square root of the sum of the squares of the errors in the ΔG°_{37} values. Errors in δ_{AB} values are calculated from eq 1a rather than eq 1b to avoid overcounting.

Results

Thermodynamic Effects of Nucleotide Analogue Substitutions. The thermodynamic consequence of functional group substitution in a sheared GA base pair was probed by examining the stability of a series of nucleotide analogues in the d(cGNAg) triloop, whose NMR structure is known (Figure 1).[35] The nucleotide analogues used and representative UV melts are provided in Figures 2 and 3.

When substituted for G, 2-aminopurine (2AP) eliminates the NH1 imino proton and O6-carbonyl group; 7-deazaguanosine (7dG) changes the 7-nitrogen to carbon; and inosine (I) eliminates the 2-amino group (Figure 2A). When substituted for an adenosine, I has a carbonyl group at position 6 and an imino proton at position 1; 7-deazaadenosine changes the 7-nitrogen to carbon; and purine eliminates the 6-amino group (Figure 2B). In many cases, it was possible to delete a functional group altogether instead of changing it to another group in order to lessen the possibility of creating new, compensating interactions, which is an important consideration for the nonadditivity studies (see below).

Substitution of 7dG at position 1 of the loop, d(c*7dG*CAg), did not give a significant thermodynamic penalty ($\Delta\Delta G^\circ_{37}$ = 0.06 kcal/mol) (Table 1), which is consistent with the sheared GA pairing. For comparison, a G to 7dG substitution was also made in a stem with a related loop, d(g*7dg*acGCACgtcc), and gave a similar destabilization with a $\Delta\Delta G^\circ_{37}$ of 0.34 kcal/mol. The effect of 7dG in the stem is similar to effects reported for RNA duplexes, where $\Delta\Delta G^\circ_{37}$ values of 0.14 to 0.36 kcal/mol were observed.[36] These comparisons suggest that the small free energy penalty for 7dG substitution in the loop is not significant. Substitution of the loop G with I, d(c*I*CAg), which eliminates the 2-amino group of G and thereby hydrogen bond 2, gave a large energetic penalty with a $\Delta\Delta G^\circ_{37}$ of 1.65 kcal/mol (Figure 3, Table 1). Interaction 2 was also probed by eliminating the hydrogen bond acceptor, the N7 of A, by substituting 7-deazaadenosine (7dA) for the A at position 3 of the loop, d(cGC*7dA*g). This nucleotide analogue also had a large destabilizing effect with a $\Delta\Delta G^\circ_{37}$ of 1.82 kcal/mol. For comparison, substitution of 7dA for an A in a stem with a related loop, d(gg*7da*cGCACgtcc), gave a much smaller destabilization with a $\Delta\Delta G^\circ_{37}$ of 0.50 kcal/mol. Interaction 1 in the loop was probed by substitution of the A at position 3 with I or purine (Pur) and was also destabilizing with a $\Delta\Delta G^\circ_{37}$ of 1.35 and 0.96 kcal/mol, respectively. Thus both interactions 1 and 2 are critical to the stability of the loop.

The G at position 1 of the loop was further probed by substitution with 2AP, which was found to be destabilizing with a $\Delta\Delta G^\circ_{37}$ value of 0.79 kcal/mol. This substitution eliminates the imino proton and carbonyl of the loop G suggesting that these two functionalities contribute to stacking interactions with the CG closing base pair.[20] Interestingly, the imino proton and carbonyl of the G of a sheared GA base pair are known to contribute to folding stability in an RNA hairpin as well.[37] In the RNA hairpin study, a water-mediated hydrogen bond from the imino proton of the loop G to a nearby phosphate was used to rationalize the 2AP substitution effect;[37] however, there is no such phosphate nearby in the DNA loop structure.[35]

Investigating the Additivity of Loop−Loop Interactions. The nucleotide analogue modifications described above were consistent with the sheared GA conformation (Figure 1). To test for cooperativity of the interactions within the loop, double and triple mutants were constructed and analyzed. Double mutant cycles are represented by thermodynamic boxes composed of the reference sequence, the two single mutants, and the double mutant (Figure 4A).

A term representing the nonadditive effect, δ_{AB}, was calculated according to eq 1a or 1b. A completely additive set of mutations has a δ_{AB} of 0, and a completely nonadditive set of

(34) SantaLucia, J., Jr.; Turner, D. H. *Biopolymers* **1997**, *44*, 309−319.
(35) Hirao, I.; Kawai, G.; Yoshizawa, S.; Nishimura, Y.; Ishido, Y.; Watanabe, K.; Miura, K. *Nucleic Acids Res.* **1994**, *22*, 576−582.
(36) Burkard, M. E.; Turner, D. H. *Biochemistry* **2000**, *39*, 11748−11762.
(37) SantaLucia, J., Jr.; Kierzek, R.; Turner, D. H. *Science* **1992**, *256*, 217−219.

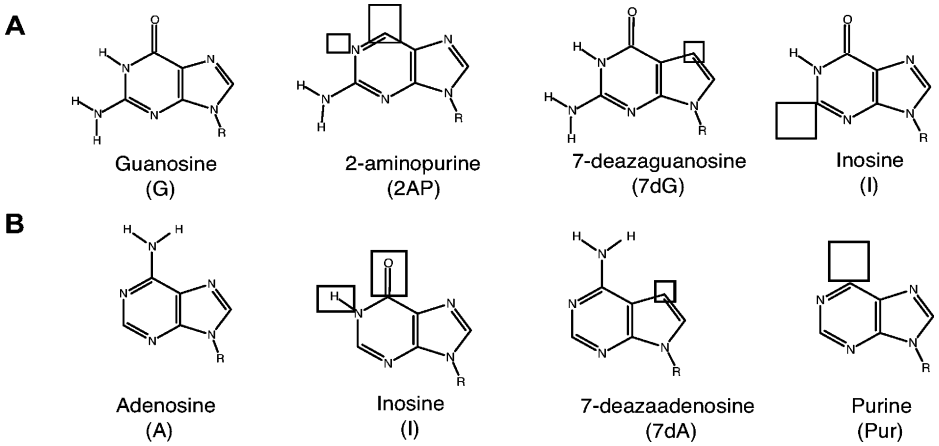

Figure 2. Nucleotide analogues used in this study. (A) Guanosine and analogues used to replace it, and (B) adenosine and analogues used to replace it. Boxes represent areas of change between the natural base and the analogue.

Table 1. Thermodynamic Parameters for Folding of Single, Double, and Triple Mutants in the d(cGCAg) Hairpin[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)	$\Delta G_A{}^{b-d}$ (kcal mol^{-1})
			Stem Modifications		
GCAC[e]	−33.9 ± 0.6	−99.6 ± 2.2	−2.97 ± 0.09	67.0 ± 1.6	
7dg	−31.1 ± 1.5	−91.8 ± 4.5	−2.63 ± 0.13	65.6 ± 0.7	0.34 ± 0.17
7da	−29.9 ± 0.9	−88.5 ± 2.7	−2.47 ± 0.07	64.9 ± 0.4	0.50 ± 0.14
			Loop Modifications		
GCA[e]	−31.7 ± 1.4	−90.7 ± 4.1	−3.60 ± 0.14	76.8 ± 0.5	
*I*CA	−28.8 ± 1.3	−86.6 ± 4.0	−1.95 ± 0.05	59.6 ± 0.9	1.65 ± 0.15
*2AP*CA	−32.6 ± 1.6	−95.9 ± 4.7	−2.81 ± 0.15	66.3 ± 0.8	0.79 ± 0.20
*7dG*CA	−27.0 ± 3.3	−75.6 ± 9.6	−3.55 ± 0.32	84.4 ± 3.3	0.06 ± 0.35
GC*T*	−26.6 ± 1.1	−80.4 ± 3.2	−1.56 ± 0.08	56.4 ± 1.0	2.05 ± 0.16
GC*7dA*	−26.5 ± 2.6	−79.7 ± 7.9	−1.78 ± 0.13	59.4 ± 0.9	1.82 ± 0.19
GC*I*	−28.8 ± 0.9	−85.5 ± 2.9	−2.25 ± 0.11	63.3 ± 1.3	1.35 ± 0.17
GC*Pur*	−30.0 ± 1.8	−88.1 ± 5.4	−2.64 ± 0.09	67.1 ± 1.3	0.96 ± 0.16
*IC**I*	−26.2 ± 0.6	−79.4 ± 1.8	−1.53 ± 0.05	56.3 ± 0.9	2.08 ± 0.15
*IC**Pur*	−26.7 ± 2.9	−80.8 ± 8.9	−1.60 ± 0.14	56.9 ± 1.1	2.00 ± 0.20
IC7dA	−26.7 ± 1.4	−79.9 ± 4.2	−1.91 ± 0.15	60.8 ± 1.1	1.70 ± 0.20
*C3 2AP*CA	−26.8 ± 2.5	−81.0 ± 7.8	−1.65 ± 0.08	57.6 ± 1.7	1.95 ± 0.16
*2AP C3*CA	−31.5 ± 2.2	−92.7 ± 7.7	−2.77 ± 0.19	69.1 ± 1.3	0.83 ± 0.23
*2APCC3*A	−32.4 ± 2.1	−95.4 ± 6.6	−2.84 ± 0.06	66.9 ± 1.8	0.76 ± 0.15
2APCAC3	−28.5 ± 2.8	−83.8 ± 8.5	−2.47 ± 0.17	66.5 ± 1.1	1.13 ± 0.22
*C3*GC*I*	−29.2 ± 0.8	−88.7 ± 2.6	−1.67 ± 0.05	55.8 ± 0.9	1.93 ± 0.16
GC*3*C*I*	−30.1 ± 1.2	−90.8 ± 3.7	−1.89 ± 0.06	57.9 ± 0.8	1.71 ± 0.15
GC*C3 I*	−29.4 ± 2.7	−90.5 ± 8.7	−1.29 ± 0.07	51.4 ± 1.4	2.31 ± 0.15
GC*I C3*	−31.9 ± 1.5	−95.7 ± 4.6	−2.24 ± 0.10	60.4 ± 1.2	1.36 ± 0.17
*C3*GC*7dA*	−27.4 ± 1.6	−83.2 ± 5.2	−1.60 ± 0.02	62.5 ± 0.5	2.00 ± 0.14
G*C3*C*7dA*	−27.6 ± 2.1	−82.3 ± 6.7	−2.09 ± 0.07	62.5 ± 1.7	1.51 ± 0.16
GC*C3 7dA*	−26.9 ± 2.9	−81.8 ± 9.2	−1.54 ± 0.13	56.0 ± 1.8	2.06 ± 0.19
GC*7dA C3*	−27.8 ± 1.0	−82.7 ± 3.0	−2.11 ± 0.12	62.5 ± 0.5	1.49 ± 0.18
*C3 I*CA	−29.6 ± 3.8	−90.1 ± 11.9	−1.70 ± 0.15	56.0 ± 1.7	1.90 ± 0.21
C3 ICI	−27.2 ± 2.2	−82.7 ± 6.9	−1.59 ± 0.06	56.3 ± 1.5	2.01 ± 0.15

[a] All sequences are DNA and are for loops that have a CG closing base pair. These hairpin loops conform to the d(cGNAg) motif.[20,47] [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. [d] ΔG_A is the free energy change associated with mutation A and is calculated as $\Delta\Delta G^{\circ}_{37}$ for mutation A referenced to the unmodified sequence. [e] Sequence in bold type is the reference for the sequences below. Changes from reference sequence are italicized. Sequences are listed in order of position and then most penalizing change. "7dg" and "7da" refer to sequences in which a substitution was made at positions 2 and 3 of the stem, respectively. Double and triple mutants follow single mutants. Thermodynamic parameters for the unmodified d(cGCAg) and d(cGCACg) entries are from a previous study.[20]

mutations has a $\delta_{AB} = -\Delta G_A = -\Delta G_B = -\Delta G_{AB}$ (see Materials and Methods for details).

The first double mutant examined was d(c*IC7dA*g), which tests the presence of hydrogen bond 2 in the sheared GA (Figure 1). This double substitution had a destabilizing effect of $\Delta\Delta G^{\circ}_{37}$ = 1.70 kcal/mol (Table 2), similar to each of the single modifications which had $\Delta\Delta G^{\circ}_{37}$ values of 1.65 (I) and 1.82 (7dA) kcal/mol. These values lead to complete nonadditivity with a δ of −1.8 ± 0.2 kcal/mol (Figure 4B). The complete nonadditivity of d(c*IC7dA*g) is expected since these changes are redundant in that they affect the same hydrogen bond (Figure 1). In essence, once a hydrogen bond is broken by deleting one of the participants, it cannot be "broken again" by deleting the other participant. The large magnitude of $\Delta\Delta G^{\circ}_{37}$ values may reflect either the intrinsic value of the hydrogen bond in interaction 2 or, as supported by the data below, the loss of additional, coupled interactions.

The double mutant d(c*ICI*g) eliminates both hydrogen bonds 1 and 2 of the sheared GA (Figure 1). This double modification gave one of the largest destabilizations in this study, with a $\Delta\Delta G^{\circ}_{37}$ of 2.08 kcal/mol (Table 1). (Substitution of T for the A of the triloop, which also breaks both hydrogen bonds, gave

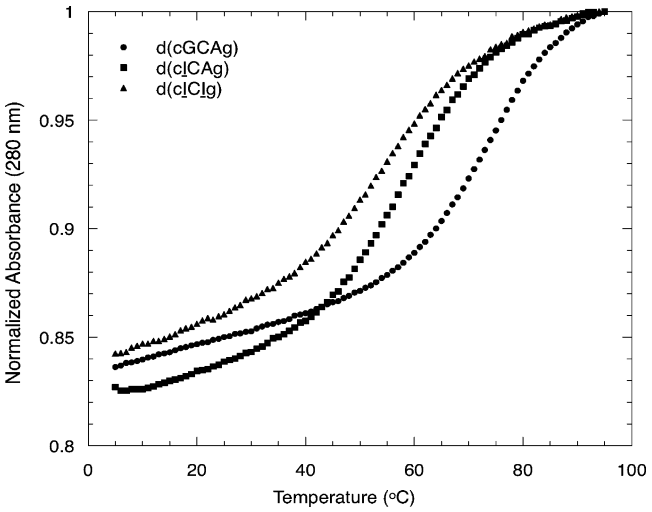

Figure 3. Representative UV melting curves. d(cGCAg) (●) was the reference sequence used in these studies. d(c*I*CAg) (■) was destabilized, with a $\Delta\Delta G^\circ_{37}$ of 1.65 kcal/mol and ΔT_M of −17.2 °C. d(c*I*C*I*g) (▲) was only slightly more destabilized with a $\Delta\Delta G^\circ_{37}$ of 2.07 kcal/mol and ΔT_M of −20.5 °C. Absorbance values were normalized by dividing each trace by its maximum absorbance value.

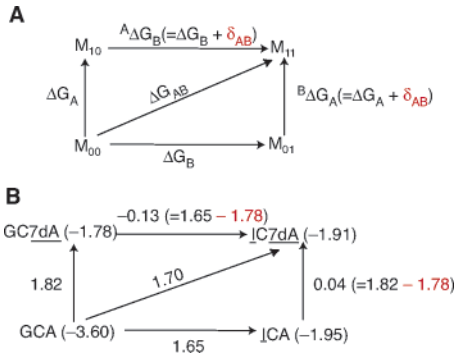

Figure 4. Thermodynamic boxes for a redundant double mutant cycle. (A) Cycle showing thermodynamic relationship between mutants A and B. M_{00} is the unmodified sequence, M_{10} is the single mutant A that affects the hydrogen bond acceptor for interaction 2, M_{01} is the single mutant B that affects the hydrogen bond donor for interaction 2, and M_{11} is the double mutant. As an example, ΔG_A is the free energy change associated with mutation A and is calculated as $\Delta\Delta G^\circ_{37}$ for mutation A referenced to the unmodified sequence. $^{B}\Delta G_A$ is the free energy change for mutation A in the background of mutation B. δ_{AB} (in red) represents the nonadditive free energy of combining the two mutations. δ_{AB} was calculated according to eq 1b. Note also that $^{B}\Delta G_A = \Delta G_A + \delta_{AB}$. (B) Free energy values at 37 °C for the redundant double mutant d(c*IC7d*Ag). Experimentally measured ΔG°_{37} values are at the corners of the box (Table 1). δ_{AB} (red) for this cycle is maximal at −1.8 ± 0.2 kcal/mol (Table 2).

a similarly large destabilization of 2.05 kcal/mol (Table 1).) The single modifications were also significantly destabilizing with $\Delta\Delta G^\circ_{37}$ values of 1.65 (I1) and 1.35 (I2) kcal/mol, respectively, giving a δ_{12} of −0.9 ± 0.2 kcal/mol (Table 2). (The subscript "12" on δ refers to coupling between interactions 1 and 2.) The simplest interpretation of this partial but still significant nonadditivity is that loss of one hydrogen bond significantly weakens the other. More complicated scenarios are also possible. For instance, the second hydrogen bond might not be weakened by removing the first, but upon removing the second hydrogen bond, a new interaction might be created elsewhere, perhaps by the new functional group. However, the next example, in which the hydrogen bonds are removed without introducing new functional groups, suggests that this is not the case.

The double mutant d(c*IC*Pur*g*) also eliminates both hydrogen bonds 1 and 2 in the sheared GA (Figure 1); however, this double mutant eliminates the 6 amino group without replacing it with a new functional group. The double mutant d(c*IC*Pur*g*) had a large destabilization, with a $\Delta\Delta G^\circ_{37}$ of 2.00 kcal/mol. The single modifications had $\Delta\Delta G^\circ_{37}$ values of 1.65 (I) and 0.96 (Pur) kcal/mol, giving a δ_{12} of −0.6 ± 0.2 kcal/mol (Table 2). The d(c*IC*Pur*g*) double mutant had within experimental error essentially the same thermodynamic penalty as d(c*IC*I*g*), consistent with the loss of one hydrogen bond weakening the other. The slightly smaller δ_{12} value for d(c*IC*Pur*g*) may be attributed to the smaller $\Delta\Delta G^\circ_{37}$ value for the purine single substitution, which may have a smaller desolvation penalty than inosine or somewhat different stacking interactions.

In summary, these double mutations support deletion of hydrogen bond 1 of the sheared GA pair weakening hydrogen bond 2, and vice versa. Loop−loop interactions can therefore be considered to be somewhat nonadditive.

Coupling of Loop-Loop and Loop-Closing Base Pair Interactions. Recently, we probed the presence of loop−closing base pair interactions in d(cGCAg) and d(cGCAGg) hairpin loops by using C3 spacers to interrupt potential interactions with the closing base pair.[23] Spacers were tolerated throughout the loop, except between the closing base pair and position 1 of the loop. Notably, this effect was absent with a GC closing base pair, which is less stable than a CG by 2−3 kcal/mol, indicating that the loop−closing base pair interaction is specific to a CG closing base pair. Studies using 2AP and DAP substitutions at position 1 of the loop suggested that the imino proton and carbonyl of the loop G help mediate a portion of the interaction with the CG base pair (interactions 3 and 4 in Figure 1).[23] The loop−closing base pair interaction was therefore expected to be interrupted either by a C3 spacer insertion before position 1 of the loop or by a 2AP substitution at the first position of the loop. To test this idea, an additional series of double mutants was investigated.

The double mutants d(c*2AP C3*CAg), d(c*2AP*C*C3*Ag), and d(c*2AP*CA*C3*g) gave only modest destabilizations and small positive δ values of 0.2 ± 0.3, 0.2 ± 0.3, and 0.6 ± 0.3, consistent with additive effects (Table 2). These effects were as expected since the C3 spacer at these positions does not affect the closing base pair interaction.[23] In contrast, the d(c*C3 2AP*CAg) mutant gave a negative δ value of −0.4 ± 0.3 kcal/mol. The negative sign on this δ is consistent with the C3 at the 5′ end of the loop and the 2AP affecting the same interaction, i.e., being somewhat redundant. However, 2AP does not appear to eliminate loop−closing base pair interactions to the same extent as a C3 spacer. For example, the $\Delta\Delta G^\circ_{37}$ value for 2AP substitution of 0.79 kcal/mol is about half the value of 1.51 kcal/mol for C3 insertion before loop position 1. In addition, 2AP and DAP substitutions destabilize CG−closing base pair sequences only ∼0.3 kcal/mol more than GC−closing base pair sequences.[23] The small thermodynamic penalty upon 2AP substitution for the first G of the loop may seem surprising, since one might expect removal of the guanine carbonyl group to shield the protons on the amino group and weaken interaction 2. Perhaps, other interactions compensate for any weakening of interaction 2. These results suggest that a C3 spacer at the 5′ end of the loop provides the best means of breaking loop−closing base pair interactions and was therefore used in subsequent double mutant cycles.

Table 2. Free Energy Parameters and δ Values for Double Mutant Cycles in the d(cGCAg) Hairpin[a]

	interaction probed[b]	ΔG_A[c] (kcal/mol)	ΔG_B[d] (kcal/mol)	ΔG_{AB}[e] (kcal/mol)	ΔG_{AB} (if additive)[f] (kcal/mol)	δ[g]	δ (kcal/mol)
IC7dA		1.65 ± 0.15	1.82 ± 0.19	1.70 ± 0.20	3.47 ± 0.24	red.	−1.8 ± 0.2
ICI	2, 1	1.65 ± 0.15	1.35 ± 0.17	2.08 ± 0.15	3.00 ± 0.23	δ_{12}	−0.9 ± 0.2
ICPur	2, 1	1.65 ± 0.15	0.96 ± 0.16	2.00 ± 0.20	2.61 ± 0.22	δ_{12}	−0.6 ± 0.2
*C3 2AP*CA	3, 3	1.51 ± 0.16	0.79 ± 0.20	1.95 ± 0.16	2.30 ± 0.26		−0.4 ± 0.3
*2AP C3*CA		−0.11 ± 0.23	0.79 ± 0.20	0.83 ± 0.23	0.68 ± 0.30	N. I.	0.2 ± 0.3
*2APCC3*A		−0.18 ± 0.15	0.79 ± 0.20	0.76 ± 0.15	0.61 ± 0.25	N. I.	0.2 ± 0.3
2APCAC3		−0.21 ± 0.17	0.79 ± 0.20	1.13 ± 0.22	0.58 ± 0.26	N. I.	0.6 ± 0.3
*C3*GC*I*	3, 1	1.51 ± 0.16	1.35 ± 0.17	1.93 ± 0.16	2.86 ± 0.23	δ_{13}	−0.9 ± 0.2
G*C3*C*I*		−0.11 ± 0.23	1.35 ± 0.17	1.71 ± 0.15	1.24 ± 0.29	N. I.	0.5 ± 0.3
GC*C3 I*		−0.18 ± 0.15	1.35 ± 0.17	2.31 ± 0.15	1.17 ± 0.23	N. I.	1.1 ± 0.2
GC*I C3*		−0.21 ± 0.17	1.35 ± 0.17	1.36 ± 0.17	1.14 ± 0.24	N. I.	0.2 ± 0.2
*C3*GC*7dA*	3, 2	1.51 ± 0.16	1.82 ± 0.19	2.00 ± 0.14	3.33 ± 0.25	δ_{23}	−1.3 ± 0.2
G*C3C7dA*		−0.11 ± 0.23	1.82 ± 0.19	1.51 ± 0.16	1.71 ± 0.30	N. I.	−0.2 ± 0.3
GC*C3 7dA*		−0.18 ± 0.15	1.82 ± 0.19	2.06 ± 0.19	1.64 ± 0.24	N. I.	0.4 ± 0.2
G*C7dA C3*		−0.21 ± 0.17	1.82 ± 0.19	1.49 ± 0.18	1.61 ± 0.25	N. I.	−0.1 ± 0.2
*C3 IC*A	3, 2	1.51 ± 0.16	1.65 ± 0.15	1.90 ± 0.21	3.16 ± 0.22	δ_{23}	−1.3 ± 0.2
	interaction probed[b]	$^{c}\Delta G_A$[h] (kcal/mol)	$^{c}\Delta G_B$[h] (kcal/mol)	$^{c}\Delta G_{AB}$[h] (kcal/mol)	$^{c}\Delta G_{AB}$ (if additive)[h] (kcal/mol)	δ[h]	δ (kcal/mol)
C3 ICI	3, 2 in 1 bkg	0.58 ± 0.12	0.72 ± 0.12	0.66 ± 0.13	1.3 ± 0.17	$^{1}\delta_{23}$	−0.6 ± 0.1
C3 ICI	3, 1 in 2 bkg	0.25 ± 0.16	0.42 ± 0.07	0.36 ± 0.08	0.67 ± 0.17	$^{2}\delta_{13}$	−0.3 ± 0.2
C3 ICI	2, 1 in 3 bkg	0.39 ± 0.17	0.42 ± 0.10	0.50 ± 0.10	0.81 ± 0.20	$^{3}\delta_{12}$	−0.3 ± 0.2

[a] All sequences are DNA and are for loops that have a CG closing base pair. [b] "Interaction probed" refers to the interactions shown in Figure 1 for the sheared GA conformation. Interactions 3 and 4 are represented by the number 3 for simplification because both interactions change simultaneously upon substitution. [c] ΔG_A values are the free energy changes associated with the single modification that breaks the first of the "interactions probed" listed and are calculated as $\Delta\Delta G°_{37}$ for mutation A referenced to the unmodified sequence. Values for this column and columns 4 and 5 are from Table 1, except the C3 spacer values, which were determined in a previous study.[23] [d] ΔG_B values are the free energy changes associated with the single modification that breaks the second of the "interactions probed" listed. [e] ΔG_{AB} values are the free energy changes associated with both modifications in a single oligonucleotide. [f] Values are the sum of ΔG_A and ΔG_B. [g] δ values were calculated from the difference between column 5, which is the actual free energy associated with the double mutant, and column 6, which is the expected free energy change if the single mutants were additive. Errors were propagated from eq 1a. "Red." refers to redundant modifications that affect the same interaction. "N.I." is for a case where a second interaction was not interrupted. In all other cases, the two interactions, and any background interaction, are given for δ. [h] Superscript "c" refers to quantities measured in the background of a third change.

Since the d(cGNAg) hairpin loop is characterized by loop−loop and loop−closing base pair interactions, we wanted to test whether altering the loop−closing base pair interaction might affect the strength of the loop−loop hydrogen bonds. To test this idea, additional double mutant cycles were investigated. A set of double mutants involving A to I substitutions at position 3 of the loop was investigated in the background of C3 spacers at different loop positions to determine the extent of additivity of interactions 1 and 3/4. For the d(cG*C3*C*I*g) and d(cGC*I C3*g) hairpin loops, the δ values were small at 0.5 ± 0.3 and 0.2 ± 0.2 kcal/mol, indicating that these modifications are nearly additive. The double mutant d(c*C3*GC*I*g), on the other hand, gave a significant δ value of −0.9 ± 0.2 kcal/mol (=δ_{13}), suggesting that interruption of loop−closing base pair interaction via the C3 spacers disrupts or significantly weakens loop−loop interaction 1. (The subscript "13" on δ refers to coupling between interactions 1 and 3/4.) Curiously, the d(cGC*C3 I*g) double mutant had a positive δ value of +1.1 ± 0.2 kcal/mol. This relatively large, positive δ value was unexpected, but since it is positive, it does not indicate the same type of coupling seen in the rest of this study. A positive δ value indicates a synergestic effect, in which a double mutant is less stable than the sum of the two single mutants. This may be caused by the C3 spacer positioning the carbonyl group of inosine in an unfavorable interaction.

A 7dA in the triloop should disrupt hydrogen bond 2, and a C3 spacer at the 5′ end of the loop should disrupt interactions 3/4. Thus, this double mutant cycle was used to probe coupling of loop−loop and loop−closing base pair interactions through loop−loop hydrogen bond 2. The double modifications, d(cG*C3*C*7dA*g), d(cGC*C3 7dA*g), and d(cGC*7dA C3*g) gave δ value of −0.2 ± 0.3, +0.4 ± 0.2, and −0.1 ± 0.2, respectively (Table 2). These values are close to zero, indicating that the C3 spacers at these positions and the 7dA are roughly additive and therefore not communicating with each other. In contrast, the d(c*C3*GC*7dA*g) double mutant gave a large negative δ of −1.3 ± 0.2 kcal/mol (=δ_{23}). The large negative δ value strongly supports strong positive coupling between interactions 2 and 3/4 and the notion that weakening of the loop−closing base pair interactions weakens loop−loop hydrogen bonding and vice versa. Overall, the d(c*C3*GC*I*g) and d(c*C3*GC*7dA*g) double mutant cycles support the existence of a cooperative and highly coupled network of interactions for these stable hairpin loops.

Next, a series of double and triple mutant cycles were investigated to probe the origin of the positive coupling. These cycles were carried out using I at position 3 of the loop to test interaction 1, I at position 1 of the loop to test interaction 2, and C3 before position 1 of the loop to test interactions 3/4. First, the coupling between interactions 2 and 3/4 was retested using the double mutant, d(c*C3 I*CAg). A large positive coupling of $\delta_{23} = -1.3 \pm 0.2$ kcal/mol was found (Table 2), in agreement with the double mutant cycle using d(c*C3*GC*7dA*g). This result indicates that coupling between these interactions is not specific to one set of mutants.

Next, the three δ values between interactions 1, 2, and 3 were calculated in the presence of the other site being modified. A term representing the nonadditive effect under these conditions, $^{C}\delta_{AB}$, was calculated according to eqs 2a or 2b. The triple mutant cycle for d(c*C3 IC*I*g*) is given in Figure 5, and the δ values are near the bottom of Table 2.

It can be noted that $^{C}\delta_{AB}$ is defined on a face of the cube opposite that for δ_{AB} (Figure 5A). Analysis of the mutants led to $^{1}\delta_{23}$, $^{2}\delta_{13}$, $^{3}\delta_{12}$ values of −0.6 ± 0.1, −0.3 ± 0.2, and −0.3 ± 0.2, respectively. In all three cases, the nonadditivity was

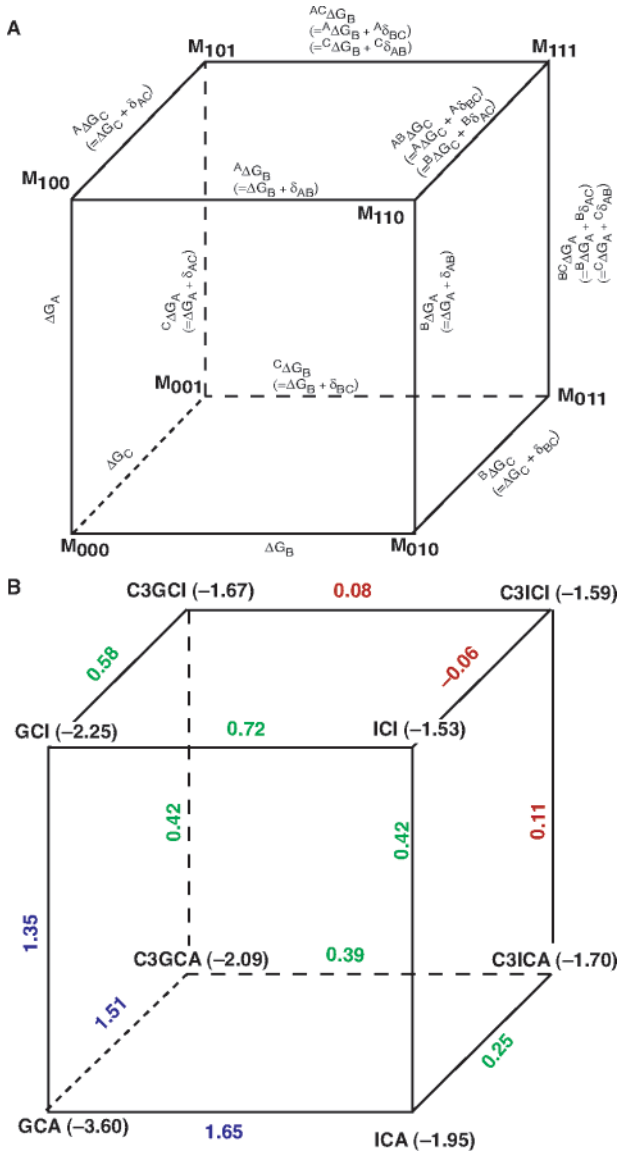

Figure 5. Thermodynamic cubes for triple mutant cycle. (A) Cycle showing thermodynamic relationship between mutants A, B, and C. M_{000} is the unmodified sequence, M_{100} is the mutation to hydrogen bond 1, M_{010} is the mutation to hydrogen bond 2, and M_{001} is the mutation that affects interaction 3/4. (See Figure 1 for details.) Formulas along the edges give free energy changes in terms of the nonadditivity free energies. δ values were calculated according to eqs 1b and 2b and are given in Table 2. Notation is further explained in the caption to Figure 4. (B) Example for the triple mutant d(c*C3 ICI*g). Experimentally measured ΔG°_{37} values are at the vertexes of the cube (Table 1), and the free energy change associated with a mutation is given along an edge. Changes in free energy in going from the unmodified to a single mutant are given in blue; changes in going from a single to a double mutant are given in green, and changes in going from a double to the triple mutant are given in red. δ values are provided in Table 2.

significantly diminished relative to what was found in the presence of the other interaction. This finding suggests that the coupling is not a result of direct interaction between any of the two sites and requires all three interactions in order to be optimal.[32]

Discussion

One of the goals of the RNA folding problem is to understand the thermodynamics and kinetics of secondary and tertiary structure formation. The energetics of structure formation can be complicated by cooperativity in folding, making it impossible to understand contributions of interactions to stability by simple single mutant analysis. A more complete understanding requires double and higher-order mutational analyses in order to identify the presence of coupling and to understand its origin.[31−33] DNA, although less recognized in forming complex structures than RNA, has important structural roles in nature, as pointed out in the Introduction. More importantly for this study, DNA provides a minimally stable system in which to examine cooperativity in nucleic acid folding.

Coupling Is Indirect and Increases with the Number of Interactions. The simplest source of nonadditivity from double mutant cycles is if both modifications affect the same interaction, such as hydrogen bond 2 (Figure 1). This redundant double mutant cycle led to complete nonadditivity and confirmed the identity of hydrogen bond 2 (Table 2); however, it did not reveal any information about coupling of the molecule. Of more fundamental interest is the nonadditivity of a double mutant cycle that arises between *different* interactions. In this study, we report a network of coupled interactions within the stable d(cGNAg) hairpin loop. By using double mutant cycles, hydrogen bonds 1 and 2 within the loop were shown to couple to each other as well as to interactions 3/4 between the loop and closing base pair.

The coupling free energies for d(c*C3 ICI*g) in the presence of a third mutation, ${}^{C}\delta_{AB}$, are −0.3 to −0.6 kcal/mol, which are substantially smaller than those in the unmodified background, δ_{AB} of −0.9 to −1.3 kcal/mol (Table 2). While the negative ${}^{C}\delta_{AB}$ values indicate that the simultaneous presence of any two interactions is worth more than the sum of any single interactions, the positive cooperativity seen in δ_{AB} values is about twice as large, indicating that the presence of all three interactions is worth substantially more than any two interactions. Thus, in these small DNA hairpins, cooperativity is positive and increases with the number of interactions. This has the potential to provide both specificity and stability to folding. Specificity can be achieved since the hairpin is not appreciably stable until all three interactions are simultaneously present. This may result in DNA loops being stable with only a small subset of all possible loop sequences, which is in agreement with the outcome of our selection experiments on DNA tetraloops.[20] Stability is also achieved through the strong positive cooperativity in that the stability of the unsubstituted loop is substantially greater than the free energy of loops in which only one interaction is removed.

The observation that coupling free energy depends on the configuration of a third site indicates that the coupling between any two sites is indirect.[31,32] Simple cases have been used to clearly illustrate the concept of direct coupling; for example, electrostatic interactions between protonation sites on glutamic acid couple directly.[32] Apparently, the coupling in stable DNA hairpins cannot be written down as a simple summation of pairwise couplings between the three interactions. The free energy penalty for deleting any of the three interactions first is similar (Figure 5B, blue), suggesting that this large, indirect coupling is almost completely dependent on all three interactions being present at once.

Indirect Coupling: Considerations of Enthalpy−Entropy Compensation and Stacking Interactions. As discussed in the previous section, the positive coupling requires all three

interactions to be optimal. It is worth considering the origin of such coupling. One possibility is that enthalpy−entropy compensation, in freezing out translation, rotational, and vibrational motions of the loop, prevents this entropic "price" from having to be paid each time a new interaction is made. Moving from M_{111} to M_{000} in Figure 5B provides a useful way of viewing how the energetics of the DNA hairpin are distributed as the interactions are "built-up". In particular, as one proceeds from the triply substituted sequence, M_{111}, to any of the doubly substituted sequences, almost no change in free energy is found (Figure 5B, red). This suggests that in the triply substituted DNA loops there are not enough other interactions to stabilize addition of the first of the three interactions. However, as one proceeds from any of the doubly substituted sequences to the singly substituted sequences, free energy changes of −0.25 to −0.7 kcal/mol are found (Figure 5B, green). Thus, the presence of one interaction, while not stable itself, apparently allows a second interaction (and the first) to form. A simple explanation for these differences in free energy increments is that the entropy loss required for formation of only one interaction is too great to allow it to form but that it is compensated for by the formation of the second interaction.

As one proceeds from any of the singly substituted sequences to the unsubstituted sequence, exceptionally large free energy changes of 1.35 to 1.65 kcal/mol are found (Figure 5B, blue). Our initial expectation was that this large effect might be due entirely to increasing bonuses from enthalpy−entropy compensation of hydrogen bond formation. However, we do not favor this model, in part based upon considerations of other studies. In the case of the exceptionally stable RNA hairpin loop sequence, cGCAAg, deletion of single functional groups was worth only 0.3 to 0.5 kcal/mol in ΔG°_{37}. (The latter comes from dividing $\Delta\Delta G^\circ_{37}$ for each deletion by the number of hydrogen bonds and excludes 2AP data since the source of this free energy is uncertain.[37]) If enthalpy−entropy compensation of hydrogen bonding was the sole basis for the 1.35 to 1.65 kcal/mol effects seen in DNA, then deleting a functional group from a "rigid" loop such as in RNA where more interactions are present should have led to equally large or even larger penalties. In addition, coupling free energies for double mutant cycles of G+1 and U42 in the hairpin ribozyme were found to be worth only −0.24 to −0.62 kcal/mol, despite these nucleotides being involved in more interactions (five each) than the G in the DNA loops.[33] If enthalpy−entropy compensation in hydrogen bond formation dominated, then deleting interactions in this more rigid system should have led to even greater coupling than that found in DNA. (Of course, increasingly complex scenarios are also possible wherein addition of hydrogen bonds leads to weakening of otherwise optimized interactions and therefore negative cooperativity; however, the magnitudes of such negatively cooperative interactions would have to be on the order of +1 kcal/mol to reverse the effects found in the DNA hairpins.) Apparently, the energetics of hydrogen bonds are limited to modest values (see next section).

If the large coupling seen in DNA is not consistent with enthalpy−entropy compensation of hydrogen bond formation, then what is its origin? Since the coupling is indirect and involves all three interactions being present simultaneously, it is possible that some type of stable stacking interaction is responsible, which might also involve enthalpy−entropy compensation as is typical for stacking. The presence of a special stacking bonus is supported by the large free energy penalties found for insertion of a C3 spacer following a CG closing base pair.[23] In addition, the large magnitude of the coupling free energy, −0.9 to −1.3 kcal/mol, is consistent with values for stacking of dangling ends in DNA and RNA.[38−40] This suggests that the stacking interaction may involve optimal positioning of induced or permanent dipoles in the hairpin. Alternatively, cooperativity may arise in part because modification of one functional group may affect an unmodified functional group via conjugation in the nucleobase, which might diminish hydrogen bonding, stacking, and propeller twisting.[41] Further studies will be necessary to test these possibilities.

What is a Hydrogen Bond Worth? An interesting debate has taken place in the literature over the energetic contribution of a hydrogen bond. Deletion of functional groups from GC base pairs in RNA duplexes by replacement with IC base pairs has led to estimates of −1.6 to −0.7 kcal/mol for a hydrogen bond.[42] Similarly large values (−2.2 to −1.4 kcal/mol) have been reported for removing hydrogen bonding groups from self-splicing RNAs.[43]

On the other hand, substantially smaller values have been reported elsewhere for the energetic contribution of a hydrogen bond. For example, hydrogen bonds in a cGCAAg tetraloop in RNA were analyzed by functional group substitution and found to be worth only −0.3 to −0.5 kcal/mol in ΔG°_{37}.[37] Likewise, individual hydrogen bonds involved in the tertiary structure of the *Tetrahymena* group I intron P4−P6 domain were found to be worth only −0.4 to −0.5 kcal/mol each in a ribose zipper.[44]

The differences among these various findings may be explained most readily by a correction for cooperativity in the hydrogen bond formation. Deleting a single hydrogen bond participant in the unsubstituted DNA loops studied here not only deletes that particular hydrogen bond but also weakens other interactions. A more accurate quantitation of the energetics of a hydrogen bond may be one-half the increment for d(c*C3* I*C*I*g*) to d(c*C3* GCAg), which adds two hydrogen bonds in the absence of loop−closing base pair coupling. This analysis results in a modest value for a hydrogen bond of 0.25 kcal/mol (Figure 5B). This estimate for a hydrogen bond in the absence of cooperativity is in good agreement with those from the studies on stable RNA hairpins and the tertiary structure mentioned above;[37,44] in the latter case, the authors also showed that deletion of multiple hydrogen bonds was additive (see below). This energetic value for hydrogen bonding in the absence of cooperativity also agrees with values from Rebek, Kool, and co-workers.[45,46]

Previous studies on RNA systems have demonstrated both additive and nonadditive interactions. Silverman and Cech investigated hydrogen bonds involved in tertiary interactions

(38) Freier, S. M.; Alkema, D.; Sinclair, A.; Neilson, T.; Turner, D. *Biochemistry* **1985**, *24*, 4533−4539.
(39) Burkard, M. E.; Kierzek, R.; Turner, D. H. *J. Mol. Biol.* **1999**, *290*, 967−982.
(40) Bommarito, S.; Peyret, N.; Santa Lucia, J., Jr. *Nucleic Acids Res.* **2000**, *28*, 1929−1934.
(41) SantaLucia, J. J.; Kierzek, R.; Turner, D. H. *J. Am. Chem. Soc.* **1991**, *113*, 4313−4322.
(42) Turner, D. H.; Sugimoto, N.; Kierzek, R.; Dreiker, S. D. *J. Am. Chem. Soc.* **1987**, *109*, 3783−3785.
(43) Bass, B. L.; Cech, T. R. *Nature* **1984**, *308*, 820−826.
(44) Silverman, S. K.; Cech, T. R. *Biochemistry* **1999**, *38*, 8691−8702.
(45) Kato, Y.; Conn, M. M.; Rebek, J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1208−1212.
(46) Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 1−22.
(47) Yoshizawa, S.; Kawai, G.; Watanabe, K.; Miura, K.; Hirao, I. *Biochemistry* **1997**, *36*, 4761−4767.

in the 160 nucleotide P4−P6 domain of the *Tetrahymena* group I intron and found that the modest contributions of tertiary hydrogen bonds were approximately additive, reflecting little or no cooperativity.[44] In this case, the formation of a tertiary structure from large secondary structures was studied, and the hydrogen bonds were spread out over several nucleotides. The study by Klostermeier and Millar on the hairpin ribozyme[33] did reveal positive coupling; however, it was substantially more modest than that found here for DNA hairpin loops. In the case of the hairpin ribozyme, the interactions mediated by the G+1 and U42 nucleotides were also spread out, being over five interactions per nucleotide.

It appears that the greater cooperativity in the stable DNA hairpin loops arises since one functional group cannot be deleted without severely weakening the other interactions. This may be because stable DNA loops have fewer interactions than the RNA motifs described above. Similar to the analogy of Kraut et al. on a poorly built house,[25] deletion of a single interaction in a DNA loop may be like removal of one leg of a three-legged stool. While each interaction "supports" one-third of the load when the stool is intact, deletion of any one leg results in the entire load being lost. This appears to not be the case for deletion of hydrogen bonds in more rigid backgrounds. The present study demonstrates that for minimally stable systems, the nonadditivity formalism commonly used to assess cooperativity in nucleic acid folding can lead to very large nonadditive effects. This complements the studies on RNA tertiary structures with a greater number of interactions. It will be interesting to see whether exceptionally stable RNA hairpin loops, which have a greater number of interactions than DNA loops, also exhibit cooperativity in their energetics.

Acknowledgment. We thank Mariko Nakano for collecting initial data on several inosine-containing oligonucleotides and David Proctor for comments on the manuscript. This work was supported by National Science Foundation 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.

JA038897Y