J|A|C|S
COMMUNICATIONS

Published on Web 01/29/2003

Thermodynamic Coupling of the Loop and Stem in Unusually Stable DNA Hairpins Closed by CG Base Pairs

Ellen M. Moody and Philip C. Bevilacqua*

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

Received December 19, 2002; E-mail: pcb@chem.psu.edu

Hairpins are common secondary structural elements in RNA and DNA. RNA hairpins play roles in initiating folding and forming tertiary structure and protein binding sites,[1] and DNA hairpins are involved in regulating replication and transcription.[2] Earlier studies demonstrated a dependence of stability on hairpin loop and closing base pair identity.[3] In particular, for certain DNA hairpin loops, a CG closing base pair provides enhanced stability.[4,5] For example, changing the closing base pair from CG to GC in d(ggacGNABgtcc)[6] hairpins gives a large destabilization with $\Delta\Delta G_{37}^\circ = +1.9$ kcal/mol and $\Delta T_M \approx -17$ °C.[5] Likewise, in d(cGNNAg) hairpins, this change gives a $\Delta\Delta G_{37}^\circ$ of +2.0 kcal/mol and a ΔT_M of ~ -21 °C.[4,5] For comparison, a CG to GC change following an AT base pair is predicted to have a $\Delta\Delta G_{37}^\circ$ of only 0.16 kcal/mol.[7] Thus, for these motifs, the thermodynamic contribution of the CG closing base pair to hairpin stability cannot be explained by the current nearest-neighbor rules. Similar thermodynamic contributions of CG closing base pairs have been observed for certain RNA hairpins as well.[8] The goal of the present study is to probe the origin of the unusual stability of these hairpins.

We report the use of three-carbon (C3) spacers[9] (Figure 1) to investigate the expandability of DNA hairpin loops and possible coupling between the loop and closing base pair. C3-spacers provide a simple way to interrupt potential interactions between adjacent nucleotides and probe possible coupling in the molecule. C3-spacers provide the backbone length of an additional nucleotide without the possible interactions of a base (Figure 1). The model hairpins for these studies are d(cGCACg) and d(cGCAg),[5,10] which conform to the exceptionally stable d(cGNABg) and d(cGNAg) motifs, and contain sheared GA base pairs in their loops (Figure 2).

Previously, the d(cGNABg) and d(cGNNAg) tetraloop hairpins were shown to be expanded d(cGNAg) triloops.[5] The extent of expandability of the loop was probed here by first inserting C3-spacers throughout the triloop hairpin, d(cGCAg). Only modest thermodynamic effects were observed at positions 2−4 of the loop ($\Delta\Delta G_{37}^\circ = -0.02$ to -0.23 kcal/mol; $\Delta T_M = -1.2$ to $+2.7$ °C)[11] (Figure 3A). However, insertion of a C3-spacer between the closing base pair and nucleotide 1 of the loop (position 1) was strongly destabilizing with $\Delta\Delta G_{37}^\circ = +1.57$ and $\Delta T_M = -15.8$ °C. These results show that the triloop can be readily expanded at positions 2 and 3 to create a d(cGNNAg) loop, and at position 4 to create a d(cGNABg) loop, but not at position 1.

Similar trends for expandability were found upon insertion of C3-spacers into the tetraloop hairpin, d(cGCACg) (Figure 3B), with modest thermodynamic effects at positions 2−5 of the loop ($\Delta\Delta G_{37}^\circ = -0.34$ to $+0.60$ kcal/mol; $\Delta T_M = -0.5$ to -7.9 °C), and a large destabilization at position 1 ($\Delta\Delta G_{37}^\circ = +1.61$ kcal/mol; $\Delta T_M = -15.4$ °C). The C3-spacers were somewhat destabilizing for positions 4 and 5 of the tetraloop, while they were slightly stabilizing for positions 3 and 4 of the triloop. The difference may be because of the more unfavorable entropy change for closure of a larger loop and relief of strain upon C3-spacer insertion into the triloop. To test loop expandability further, two C3-spacers were added in tandem at the 3′ end of the triloop. Addition of the second C3-spacer was destabilizing by 1.38 kcal/mol (Figure 3A), likely due to increased entropic cost for forming the loop. Consistent with this idea, the triloop with the tandem C3-spacers, d(cGCA(C3)$_2$g), was similar in stability to the tetraloop with the single C3-spacer, d(cGCAC(C3)g), with a difference in ΔG_{37}° of 0.19 kcal/mol.

The large destabilization for insertion of the C3-spacer at position 1 of both the triloop and the tetraloop hairpins strongly supports an interaction between the 5′ end of the loop and the CG closing base pair. To investigate this interaction, the closing base pair was changed to the other Watson−Crick possibilities. Hairpins with the d(GCAC) loop and AT, GC, and TA closing base pairs were less stable than d(cGCACg) by 2.38, 1.88, and 1.62 kcal/mol, respectively (Table 1). Nearest-neighbor rules predict $\Delta\Delta G_{37}^\circ$ values for AT, GC, and TA closing base pairs in this stem of only 0.44, 0.16,

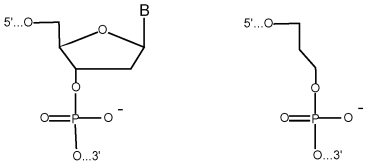

Figure 1. The C3-spacer (right) has the same framework as the backbone of a nucleotide (left), but it does not contain a sugar or base (B).

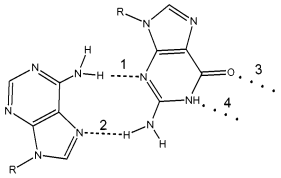

Figure 2. Diagram of the sheared GA base pair of the loop showing the hydrogen bonds, 1 and 2 (dashed lines), between the G and A. Also shown are two potential interactions, 3 and 4 (dotted lines), from the major groove face of the G to the CG closing base pair of the stem, which would lie below. Dashed lines are not used for interactions 3 and 4 because it is not known if they involve hydrogen bonds (see text).

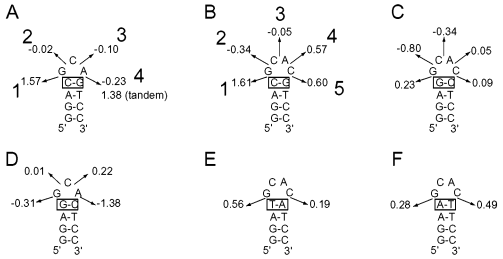

Figure 3. Tetraloops and triloops containing C3-spacers with $\Delta\Delta G_{37}^\circ$ values relative to those of unsubstituted loops.[11] Numbering for substitution is denoted in panels A and B. Negative $\Delta\Delta G_{37}^\circ$ values indicate stabilization, and positive values indicate destabilization. Closing base pairs are boxed. "Tandem" indicates the $\Delta\Delta G_{37}^\circ$ for insertion of two C3-spacers at this position relative to 1. All oligonucleotides are DNA.

10.1021/ja029831q CCC: $25.00

Table 1. Thermodynamic Parameters for Hairpin Formation

sequence[a]	$\Delta G_{37}°$ (kcal mol^{-1})	T_M (°C)	$\Delta\Delta G_{37}°$ (kcal mol^{-1})[b]	ΔT_M (°C)
d(cGCACg)	−2.96	67.5		
d(gGCACc)	−1.08	50.2	1.88	−17.3
d(tGCACa)	−1.34	52.1	1.62	−15.4
d(aGCACt)	−0.58	45.7	2.38	−21.8
d(c(2AP)CACg)	−2.04	58.6	0.92	−8.9
d(c(DAP)CACg)	−1.83	58.9	1.13	−8.6
d(gGCACc)	−1.08	50.2		
d(g(2AP)CACc)	−0.44	43.5	0.64	−6.7
d(g(DAP)CACc)	−0.29	43.0	0.79	−7.2
d(cGCAg)	−3.69	76.1		
d(gGCAc)	−0.63	51.3	3.06	−24.8

[a] 2AP is 2-aminopurine, and DAP is 2,6-diaminopurine. [b] $\Delta\Delta G_{37}°$ and ΔT_M values are referenced to the entry at the top of each grouping.

and 0.56 kcal/mol, respectively. These data indicate that loop-closing base pair coupling is specific for a CG closing base pair.

To investigate the loop-closing base pair coupling further, we repeated the C3-spacer cycle in the GC closing base pair background (Figure 3C). In contrast to the hairpin with a CG closing base pair, adding C3-spacers to the d(gGCACc) hairpin did not lead to significant destabilization at any position ($\Delta\Delta G_{37}° = +0.23$ to -0.80 kcal/mol; $\Delta T_M = -2.7$ to $+6.0$ °C). Similar results were observed for the triloop hairpin with a GC closing base pair ($\Delta\Delta G_{37}° = +0.22$ to -0.31 kcal/mol; $\Delta T_M = -1.8$ to -9.5 °C), except that a C3-spacer inserted at the 3′ end of the triloop was stabilizing (Figure 3D).[12] Moreover, hairpins with C3-spacers inserted at positions 1 and 5 in d(tGCACa) and d(aGCACt) loops behaved similarly to hairpins with a GC closing base pair ($\Delta\Delta G_{37}° = +0.19$ to $+0.56$; $\Delta T_M = -0.9$ to -7.2 °C) (Figure 3E, F). Together, these results further support significant coupling of the loop and stem for a CG closing base pair only. Interestingly, d(g(C3)GCACc) and d(c(C3)GCACg) have a $\Delta\Delta G_{37}°$ of only 0.5 kcal/mol, as compared to 1.9 kcal/mol for the unmodified loops. The difference of 0.5 kcal/mol is close to the difference of 0.16 predicted by a nearest-neighbor model.[7] This similarity suggests that the additional stability of the CG closing base pair is due largely to its interaction with loop position 1.

Next, we changed the identity of non-hydrogen bonded loop functional groups, which might be free to participate in interactions with the closing base pair. The G of the sheared GA base pair (Figure 2) was substituted with 2-aminopurine (2AP) or 2,6-diaminopurine (DAP). 2AP eliminates the NH1 imino proton and 6-carbonyl group of G; DAP eliminates the imino proton and has an amino group at position 6. Importantly, these changes retain the potential for hydrogen bonds 1 and 2 of the sheared GA base pair (Figure 2). Substitutions of 2AP and DAP for G of the d(cGCACg) loop were strongly destabilizing with $\Delta\Delta G_{37}°$ values of $+0.92$ and $+1.13$ kcal/mol, respectively (Table 1). The same substitutions with a GC closing base pair were less destabilizing, with $\Delta\Delta G_{37}°$ values of $+0.64$ and $+0.79$ kcal/mol. These results suggest that the imino proton and carbonyl of the G at position 1 of the loop may help mediate the loop/CG closing base pair interactions, depicted as "3" and "4" in Figure 2.

The results herein support coupling between the CG closing base pair of the stem and the 5′ end of d(GNAB) and d(GNA) loops. C3-spacer results indicate that interaction of the CG closing base pair with the G at the first position of the loop accounts for most of the extra hairpin stability. The 2AP and DAP results suggest that the coupling is mediated in part by the carbonyl and imino functional groups of the G of the loop. These functional groups do not participate in loop−loop base pairing, suggesting they may make favorable vertical interactions with the CG closing base pair. Non-hydrogen bonded functional groups have been implicated in mediating stacking interactions in other cases. For example, vertical alignment of the carbonyl-4 of a uridine dangling end with the amino group of an adjacent C was found to contribute ~0.5 kcal/mol to $\Delta G_{37,stack}°$.[13] Alternatively, partial pyramidalization of amino and ring nitrogens may allow for hydrogen bonding between the loop and closing base pair.[14] Further studies will be needed to resolve these issues. Thermodynamic coupling of the loop and closing base pair of a hairpin may diminish dynamic behavior and affect its ability to interact with proteins and engage in tertiary structure formation.

Acknowledgment. We thank Doug Turner and members of the Bevilacqua lab for helpful comments. This work was supported by NSF Grant MCB-9984129, a Fellowship from the Alfred P. Sloan Foundation and a Camille Dreyfus Teacher-Scholar Award to P.C.B., and a Paul Berg award to E.M.M.

Supporting Information Available: Tables of $\Delta H°$, $\Delta S°$, $\Delta G_{37}°$, and T_M (PDF). All sequences are DNA. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Tinoco, I., Jr.; Bustamante, C. *J. Mol. Biol.* **1999**, *293*, 271−281 (b) Williams, D. J.; Hall, K. B. *J. Mol. Biol.* **2000**, *297*, 1045−1061. (c) Proctor, D. J.; Schaak, J. E.; Bevilacqua, J. M.; Falzone, C. J.; Bevilacqua, P. C. *Biochemistry* **2002**, *41*, 12062−12075.

(2) (a) Varani, G. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 379−404. (b) Glucksmann-Kuis, M. A.; Dai, X.; Markiewicz, P.; Rothman-Denes, L. B. *Cell* **1996**, *84*, 147−154.

(3) (a) Senior, M. M.; Jones, R. A.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6242−6246. (b) Paner, T. M.; Amaratunga, M.; Doktycz, M. J.; Benight, A. S. *Biopolymers* **1990**, *29*, 1715−1734. (c) Vallone, P. M.; Paner, T. M.; Hilario, J.; Lane, M. J.; Faldasz, B. D.; Benight, A. S. *Biopolymers* **1999**, *50*, 425−442.

(4) Sandusky, P.; Wooten, E. W.; Kurochkin, A. V.; Kavanaugh, T.; Mandecki, W.; Zuiderweg, E. R. *Nucleic Acids Res.* **1995**, *23*, 4717−4725.

(5) Nakano, M.; Moody, E. M.; Liang, J.; Bevilacqua, P. C. *Biochemistry* **2002**, *41*, 14281−14292.

(6) Lowercase denotes nucleotides in the stem, and uppercase denotes the loop; N = A, C, G, or T; and B = C, G, or T. All oligonucleotides have the same three beginning (gga) and ending (tcc) nucleotides, so, in general, only loop and closing base pair nucleotides are provided.

(7) (a) SantaLucia, J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1460−1465. (b) Allawi, H. T.; SantaLucia, J., Jr. *Biochemistry* **1997**, *36*, 10581−10594. (c) These parameters are at 1 M NaCl, while the parameters reported here are at ~10 mM Na^+ to minimize duplex formation; melting experiments for representative d(GNNA) and d(GNAB) loop sequences at 1 M NaCl gave similar $\Delta\Delta G_{37}°$ and ΔT_M values for CG to GC or AT closing base pair changes (Denkenberger, T. M.; Bevilacqua, P. C., unpublished data).

(8) (a) Antao, V. P.; Lai, S. Y.; Tinoco, I., Jr. *Nucleic Acids Res.* **1991**, *19*, 5901−5905. (b) Shu, Z.; Bevilacqua, P. C. *Biochemistry* **1999**, *38*, 15369−15379.

(9) Oligonucleotides were from IDT and were deblocked and desalted by the manufacturer. Oligonucleotides were dialyzed against $P_{10}E_{0.1}$ = [10 mM sodium phosphate, 0.1 mM Na_2EDTA], pH 7 for 3−4 h using a molecular weight cutoff of 1000. Oligonucleotides containing DAP were synthesized by the Nucleic Acids Facility on the Pennsylvania State University campus using reagents purchased from Glen Research.

(10) Hirao, I.; Kawai, G.; Yoshizawa, S.; Nishimura, Y.; Ishido, Y.; Watanabe, K.; Miura, K. *Nucleic Acids Res.* **1994**, *22*, 576−582. (b) Yoshizawa, S.; Kawai, G.; Watanabe, K.; Miura, K.; Hirao, I. *Biochemistry* **1997**, *36*, 4761−4767.

(11) Thermodynamic parameters were determined by UV melting as described.[5] Tables of $\Delta H°$, $\Delta S°$, $\Delta G_{37}°$, and T_M are available in the Supporting Information. In general, the error in $\Delta\Delta G_{37}° < \pm 0.3$ kcal/mol, and the error in $\Delta T_M < \pm 1.5$ °C.

(12) This stabilization may be because absence of coupling does not allow a GC closing base pair to form for a triloop. This is supported by $\Delta\Delta G_{37}°$ for a CG to GC change of +3.06 kcal/mol.

(13) Burkard, M. E.; Kierzek, R.; Turner, D. H. *J. Mol. Biol.* **1999**, *290*, 967−982.

(14) Sponer, J.; Leszczynski, J.; Hobza, P. *Biopolymers* **2001**, *61*, 3−31.

JA029831Q