J|A|C|S

A R T I C L E S

A Novel RNA Motif Based on the Structure of Unusually Stable 2′,5′-Linked r(UUCG) Loops

Alexei Yu. Denisov,[†] Rami N. Hannoush,[‡,§] Kalle Gehring,[†,‡] and Masad J. Damha*,[‡]

Contribution from the Department of Biochemistry and Montreal Joint Centre for Structural Biology, McGill University, Montreal, QC, H3G 1Y6, Canada and Department of Chemistry, Otto Maass Chemistry Building, McGill University, 801 Sherbrooke Street West, Montreal, QC, H3A 2K6, Canada

Received May 19, 2003; E-mail: masad.damha@mcgill.ca

Abstract: We have recently shown that hairpins containing 2′,5′-linked RNA loops exhibit superior thermodynamic stability compared to native hairpins comprised of 3′,5′-RNA loops [Hannoush, R. N.; Damha, M. J. *J. Am. Chem. Soc.* **2001**, *123*, 12368−12374]. A remarkable feature of the 2′,5′-r(UUCG) tetraloop is that, unlike the corresponding 3′,5′-linked tetraloop, its stability is virtually independent of the hairpin stem composition. Here, we determine the solution structure of unusually stable hairpins of the sequence 5′-$G_1G_2A_3C_4$-($U_5U_6C_7G_8$)-$G_9(U/T_{10})C_{11}C_{12}$-3′ containing a 2′,5′-linked RNA (UUCG) loop and either an RNA or a DNA stem. The 2′,5′-linked RNA loop adopts a new fold that is completely different from that previously observed for the native 3′,5′-linked RNA loop. The 2′,5′-RNA loop is stabilized by (a) U5·G8 *wobble* base pairing, with both nucleotide residues in the *anti*-conformation, (b) extensive base stacking, and (c) sugar−base and sugar−sugar contacts, all of which contribute to the extra stability of this hairpin structure. The U5:G8 base pair stacks on top of the C4:G9 loop-closing base pair and thus appears as a continuation of the stem. The loop uracil U6 base stacks above U5 base, while the cytosine C7 base protrudes out into the solvent and does not participate in any of the stabilizing interactions. The different sugar pucker and intrinsic bonding interactions within the 2′,5′-linked ribonucleotides help explain the unusual stability and conformational properties displayed by 2′,5′-RNA tetraloops. These findings are relevant for the design of more effective RNA-based aptamers, ribozymes, and antisense agents and identify the 2′,5′-RNA loop as a novel structural motif.

Introduction

RNA hairpins exhibit structural motifs that display an array of important biological functions. They serve as nucleation sites for tertiary RNA folding,[1,2] as protein-binding sites,[3,4] as recognition signals for interaction with other nucleic acids,[5,6] and as templates for reverse transcription termination.[7] In addition, they protect mRNA against nuclease degradation,[8] a property that has been exploited to alter the rate of degradation of antisense oligonucleotides in cells.[9]

Hairpin motifs found in ribosomal RNA generally contain four loop residues. The consensus sequences (GNRA) and (UNCG) (N = any nucleotide; R = purine base) occur quite frequently in 16S and 23S ribosomal RNAs, and UUCG and GCAA are especially favored.[10] Messenger RNAs with the sequence repeats 5′-C(UUCG)G-3′ are unusually stable ("extra-stable"), and their folded structure prevents reverse transcriptase read through.[7,11,12] Loop structure and its contribution to the thermodynamic stability of hairpins have been described in detail[1,7,11,13] as a means of understanding RNA tertiary folding. For instance, the loop (UUCG) displays the extrastability by virtue of very specific hydrogen bonding and stacking interactions.[14−17]

Very little is known about structural RNA motifs containing 2′,5′-linked nucleotide loops. Recently, we showed that the RNA

[†] Department of Biochemistry and Montreal Joint Centre for Structural Biology, McGill University.

[‡] Department of Chemistry, McGill University.

[§] Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA.

(1) Varani, G. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 379−404.
(2) Tinoco, I., Jr.; Puglisi, J. D.; Wyatt, J. R. *Nucleic Acids Mol. Biol.* **1990**, *4*, 206−226.
(3) Heaphy, S.; Dingwall, C.; Ernberg, I.; Gait, M. J.; Green, S. M.; Karn, J.; Lowe, A. D.; Singh, M.; Skinner, M. A. *Cell* **1990**, *60*, 685−693.
(4) Leulliot, N.; Varani, G. *Biochemistry* **2001**, *40*, 7947−7956.
(5) Perez-Ruiz, M.; Sievers, D.; Garcia-Lopez, P. A.; Berzal-Herranz, A. *Antisense Nucleic Acid Drug Dev.* **1999**, *9*, 33−42.
(6) Lee, A. J.; Crothers, D. M. *Structure* **1998**, *6*, 993−1005.
(7) Tuerk, C.; Gauss, P.; Thermes, C.; Groebe, D. R.; Gayle, M.; Guild, N.; Stormo, G.; d'Aubenton-Carafa, Y.; Uhlenbeck, O. C.; Tinoco, I., Jr.; Brody, E. N.; Gold, L. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1364−1368.
(8) Adams, C. C.; Stern, D. B. *Nucleic Acids Res.* **1990**, *18*, 6003−6010.
(9) Khan, I. M.; Coulson, J. M. *Nucleic Acids Res.* **1993**, *21*, 2957−2958.

(10) Woese, C. R.; Winker, S.; Gutell, R. R. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8467−8471.
(11) Antao, V. P.; Lai, S. Y.; Tinoco, I., Jr. *Nucleic Acids Res.* **1991**, *19*, 5901−5905.
(12) Antao, V. P.; Tinoco, I., Jr. *Nucleic Acids Res.* **1992**, *20*, 819−824.
(13) Serra, M. J.; Lyttle, M. H.; Axenson, T. J.; Schadt, C. A.; Turner, D. H. *Nucleic Acids Res.* **1993**, *21*, 3845−3849.
(14) Cheong, C.; Varani, G.; Tinoco, I., Jr. *Nature* **1990**, *346*, 680−682.
(15) Varani, G.; Cheong, C.; Tinoco, I., Jr. *Biochemistry* **1991**, *30*, 3280−3289.
(16) Allain, F. H. T.; Varani, G. *J. Mol. Biol.* **1995**, *250*, 333−353.
(17) Comolli, L. R.; Ulyanov, N. B.; Soto, A. M.; Marky, L. A.; James, T. L.; Gmeiner, W. H. *Nucleic Acids Res.* **2002**, *30*, 4371−4379.

Table 1. Thermodynamic Parameters of Hairpins[a]

code	hairpin	T_m (°C)	% H	$\Delta H°$ (kcal/mol)	$\Delta S°$ (eu)	$\Delta G°_{37}$ (kcal/mol)
RRR	GGAC(UUCG)GUCC	71.8	8.5	−53.4	−154.8	−5.4
RRR	GGAC(*UUCG*)GUCC	69.3	9.6	−55.6	−162.1	−5.3
DDD	ggac(uucg)gtcc	56.2	11.3	−36.6	−111.1	−2.1
DRD	ggac(UUCG)gtcc	54.6	11.5	−36.0	−109.8	−1.9
DRD	ggac(*UUCG*)gtcc	61.4	12.6	−39.9	−119.4	−2.9

[a] Adapted from ref 18. Measurements were made in 0.01 M Na$_2$HPO$_4$ and 0.1 mM Na$_2$EDTA, pH 7.0; oligonucleotide concentration was ∼4.5 μM. Values represent the average of at least five independent measurements. Error in T_m is within ±1 °C. Errors in thermodynamic parameters are within ±7.5% for $\Delta H°$ and $\Delta S°$ and ±0.20 kcal/mol for $\Delta G°_{37}$. Percentage hypochromicity (%H) was calculated from UV absorbances of the hairpin (A_0) and fully denatured species (A_f) using the following equation: $\%H = (A_f - A_0)/A_f$.
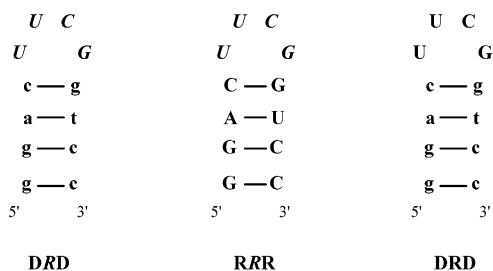

Figure 1. Schematic representation of the three hairpin sequences under study. Hairpin **DRD** contains a 2′,5′-RNA loop and DNA stem; hairpin **RRR**, the same 2′,5′-RNA loop but with an RNA stem; and hairpin **DRD** serves as a control sequence and contains the native 3′,5′-RNA loop with a DNA stem. Capital letters represent RNA residues, small letters represent DNA residues, and capital italicized letters represent 2′,5′-RNA residues.

sequence GGAC(*UUCG*)GUCC (where the italicized residues are 2′,5′-linked nucleotides) exhibits exceptional stability.[18,19] The stability of this hairpin molecule (Table 1) was identical to that of the native 3′,5′-linked hairpin. This is remarkable in view of the structural and conformational differences between 2′,5′-RNA and 3′,5′-RNA[20] and the fact that 2′,5′-linked RNA is inferior to native RNA with respect to duplex stability.[21−23] We also noted that the extrastability imparted by the (*UUCG*) loop was independent of the composition of the stem.[18,19] Hairpins containing a (*UUCG*) loop and DNA:DNA, RNA:RNA, or 2′,5′-RNA:2′,5′-RNA stem duplexes were extrastable relative to hairpins of purportedly normal thermodynamic stability. This contrasted the behavior of the native (UUCG) loop, which was extrastable only when the stem was duplex RNA (Table 1). To provide structure-based reasoning for the exceptional stability of hairpins with 2′,5′-linked loops and elucidate their general features, we investigated the three-dimensional structure of the 2′,5′-RNA (UUCG) loop within the context of RNA and DNA stem hairpins via high-resolution NMR. We report here the solution structure of 2′,5′-RNA loop hairpins **RRR** and **DRD**, simultaneously with NMR analysis of a 3′,5′-RNA loop hairpin **DRD** (Figure 1).

Experimental Section

Sample Preparation. Oligonucleotides in this work were synthesized as described previously[18] using an Applied Biosystems (381A) synthesizer and utilizing LCAA-controlled pore glass (500 Å) as solid support. Monomer coupling times were 10 min (RNA or 2′,5′-RNA monomers) and 2 min (DNA monomers). Extended coupling times were used for rG 2′- or 3′-*O*-phosphoramidite (15 min) and dG monomers (3 min). The concentrations of monomers were 0.15−0.17 M (RNA) and 0.1 M (DNA). The activator solution consisted of 0.5 M 4,5-dicyanoimidazole/acetonitrile for DNA, 3′,5′-RNA, and 2′,5′-RNA synthesis.[24] Following chain assembly, the CPG support was treated with aqueous ammonia/ethanol (3:1−1.5 mL total volume) for 48 h at room temperature. After centrifugation, the supernatant was collected, and the solid support was washed with ethanol. The supernatant and ethanol washings were combined and evaporated to dryness under vacuum. The pellet obtained was treated with NEt$_3$/3HF at room temperature for 48 h.[25] The reaction was quenched by addition of deionized double distilled water, and the resulting solution was lyophilized to dryness under vacuum. The oligomers were purified by anion-exchange HPLC (Protein Pak DEAE-5PW column-Waters, 22.5 mm × 150 mm) using a linear gradient of 0−20% LiClO$_4$ in H$_2$O (1 M) with a flow rate of 5 mL/min at 55 °C. The oligonucleotides were then desalted by using reversed-phase chromatography on a Sep-Pak cartridge.[26] The overall isolated yields as well as % purity are given in the Supporting Information (Table A).

The samples were then dissolved in 0.3 mL of 100% D$_2$O or 9:1 H$_2$O/D$_2$O, v/v, (for imino-proton spectra) to a final concentration of 1.5 mM. The solutions contained 0.1 mM EDTA sodium salt, and the final pH was adjusted to 7.0 with 100 mM NaOH.

NMR Spectroscopy. The NMR spectra were recorded on a Bruker DRX-500 spectrometer equipped with a ^{1}H/^{13}C/^{31}P triple resonance (*x, y, z*) gradient probe operating at a 500.13 MHz proton frequency. Proton chemical shifts were measured relative to internal DSS, and phosphorus resonances were indirectly referenced to 85% H$_3$PO$_4$.[27,28]

NOESY experiments were performed in D$_2$O at 15 °C for all three hairpins and additionally at 25 °C for the hairpin **DRD** using mixing times t_m of 70, 200, and 300 ms. The volumes of cross-peaks of NOESY spectra were calculated using XWINNMR (Bruker). NOESY experiments in 9:1 H$_2$O/D$_2$O (v/v) were performed at 5 °C with a mixing time of 180 ms. DQF-COSY spectra were collected with phosphorus decoupling (final data size of 8 K × 2 K points). Proton MLEV-17 TOCSY experiments were performed with a mixing time of 84 ms. ^{1}H,^{13}C-correlation HMQC spectra were recorded using GARP heteronuclear decoupling ($^1J_{CH}$ = 180 Hz). Inverse H,P-HetCOSY spectra were collected with final spectral sizes of 2 K × 1 K data points.

Structural Modeling. The starting coordinates of **DRD** and **RRR** hairpins under study were generated using Sybyl 6.5 software (Tripos Inc.) from both *A*- and *B*-type DNA structures. The X-PLOR 3.843 package[29] with a nucleic acid all-hydrogen force field was used for hairpin molecular modeling.

At the initial stage, 100 starting structures with a "randomized" loop (half with canonical *A*-type and half with *B*-type hairpin stem) were generated by molecular dynamics without experimental constraints. Subsequent stages of simulated annealing with NOE distance and torsion angle constraints were similar to those described previ-

(18) Hannoush, R. N.; Damha, M. J. *J. Am. Chem. Soc.* **2001**, *123*, 12368−12374.
(19) Hannoush, R. N.; Damha, M. J. *Nucleosides Nucleotides* **2001**, *20*, 1201−1204.
(20) Premraj, B. J.; Patel, P. K.; Kandimalla, E. R.; Agrawal, S.; Hosur, R. V.; Yathindra, N. *Biochem. Biophys. Res. Commun.* **2001**, *283*, 537−543.
(21) Wasner, M.; Arion, D.; Borkow, G.; Noronha, A.; Uddin, A. H.; Parniak, M. A.; Damha, M. J. *Biochemistry* **1998**, *37*, 7478−7486.
(22) Kierzek, R.; He, L.; Turner, D. H. *Nucleic Acids Res.* **1992**, *20*, 1685−1690.
(23) Giannaris, P. A.; Damha, M. J. *Nucleic Acids Res.* **1993**, *21*, 4742−4749.

(24) Vargeese, C.; Carter, J.; Yegge, J.; Krivjansky, S.; Settle, A.; Kropp, E.; Peterson, K.; Pieken, W. *Nucleic Acids Res.* **1998**, *26*, 1046−1050.
(25) Gasparutto, D.; Livache, T.; Bazin, H.; Duplaa, A. M.; Guy, A.; Khorlin, A.; Molko, D.; Roget, A.; Teoule, R. *Nucleic Acids Res.* **1992**, *20*, 5159−5166.
(26) Damha, M. J.; Ogilvie, K. K. *Methods Mol. Biol.* **1993**, *20*, 81−114.
(27) Trempe, J. F.; Wilds, C. J.; Denisov, A. Y.; Pon, R. T.; Damha, M. J.; Gehring, K. *J. Am. Chem. Soc.* **2001**, *123*, 4896−4903.
(28) Denisov, A. Y.; Noronha, A. M.; Wilds, C. J.; Trempe, J. F.; Pon, R. T.; Gehring, K.; Damha, M. J. *Nucleic Acids Res.* **2001**, *29*, 4284−4293.
(29) Brünger, A. T. *X-PLOR 3.1, A system for X-ray crystallography and NMR*; Yale University Press: New Haven, CT, 1992.

ously.[15,16,30] Global fold was reached by restrained simulated annealing at 5000 K, and the 20 most energetically favored structures with minimal number of structural violations were selected at this stage. Subsequently, the following gentle refinement was accomplished from these structures by molecular dynamics simulation of 12 ps (5 ps at 1000 K, 4 ps of cooling to 300 K, and then 3 ps at 300 K). During the last stage, NOE distance and hydrogen bond force constants were gradually built up to final values of 30−40 kcal/(mol Å^2), and the backbone torsion angle constants, to 60 kcal/(mol rad^2). A distance-dependent dielectric constant was used to mimic the solvent. Final ensemble of the 10 best structures (including structures which converged from both *A*- and *B*-type starting models) have been deposited into the Protein Data Bank (PDB) under ID codes 1ME0 for hairpin **D**R**D** and 1ME1 for hairpin **R**R**R**. Global helical parameters for hairpin stem base pairs were calculated using the CURVES 5.2 program.[31]

Distance restraints were derived from NOESY spectra at different mixing times by cross-peak volume integration, using the d^{-6} distance relationship[32] and average cross-peak volume values for calibration for H5−H6 in cytidines and uridines ($d = 2.45$ Å) and for Me−H6 in thymidines ($d = 2.70$ Å). The distance constraints were given with 10% of lower and 15% of upper bounds. Sugar puckering was determined by the PSEUROT 3B program[33] from vicinal coupling constants. The five torsion angles for hairpin sugars were constrained (with ±10° bounds) according to *south* or *north* sugar conformations determined from *J*-couplings. Backbone torsion angle constraints were set in both hairpins. The β torsion angles were constrained using the information about H5′/H5″−P and H4′−P cross-peaks in H,P-HetCOSY spectra. The β angles were found to be in the *trans* conformation (180 ± 60°) as determined by symmetry and the relatively low intensity of the H5′/H5″−P cross-peaks as well as detectable $^4J_{H4'-P}$ W-pathway coupling constants.[32,34,35] The γ angles were constrained (60 ± 40°) using the sums of $J_{H4'H5'}$ and $J_{H4'H5''}$ which were available from phosphorus decoupled DQF-COSY spectra and the NOE H1′/H6/H8−H4′ cross-peak line widths.[35] The ϵ angles have been estimated from the vicinal $^3J_{H3'-P}$ or $^3J_{H2'-P}$ coupling constants (Table 2). These coupling constants lie in the range of 7−9 Hz for all RNA stem and loop nucleotides, and ϵ angles (or C3′−C2′−O2′−P torsions for 2′,5′-RNA) were constrained to 240 ± 50° from the Karplus equation.[35] In contrast, the small $^3J_{H3'-P}$ values observed for the DNA stem (3−5 Hz) suggest a ϵ value of 170 ± 50°. Finally, glycosidic angles were not constrained in structure calculations but were fixed indirectly in *anti*-conformation by intranucleotide aromatic-sugar distance constraints.

Results

Resonance Assignments and Preliminary Structural Conclusions. Assignment of nonexchangeable proton resonances in NOESY spectra of hairpins **D**R**D**, **R**R**R**, and **D**R**D** was carried out in the standard manner employed for right-handed duplexes (Figure 2A and B).[28,32] DQF-COSY, TOCSY, and H,C-HMQC spectra based on different ^{13}C chemical shifts for C2′/C3′, C4′, and C5′ were also used to assign sugar protons. The adenine H2 resonances were confirmed by NOESY interstrand H2(A3)−H1′(C11) cross-peaks. The sugar signals of H1′/H2′/H2″ for the dG9 residue in hairpin **D**R**D** were strongly broadened at room temperature. Assignment of stem imino protons was made from NOESY spectra acquired in H$_2$O/

(30) James, J. K.; Tinoco, I., Jr. *Nucleic Acids. Res.* **1993**, *21*, 3287−3293.
(31) Lavery, R.; Sklenar, H. *CURVES 5.2, Helical analysis of irregular nucleic acids*; Laboratory de Biochimie Theorique, CNRS URA 77: Paris, 1997.
(32) Wüthrich, K. *NMR of proteins and nucleic acids*; John Wiley & Sons: New York, 1986.
(33) de Leeuw, F. A. A. M.; Altona, C. *Quantum chemistry program exchange, No.463: PSEUROT 3B*; Indiana University: Bloomington, IN, 1983.
(34) Saenger, W. *Principles of nucleic acids structure*; Springer-Verlag: New York, 1984.
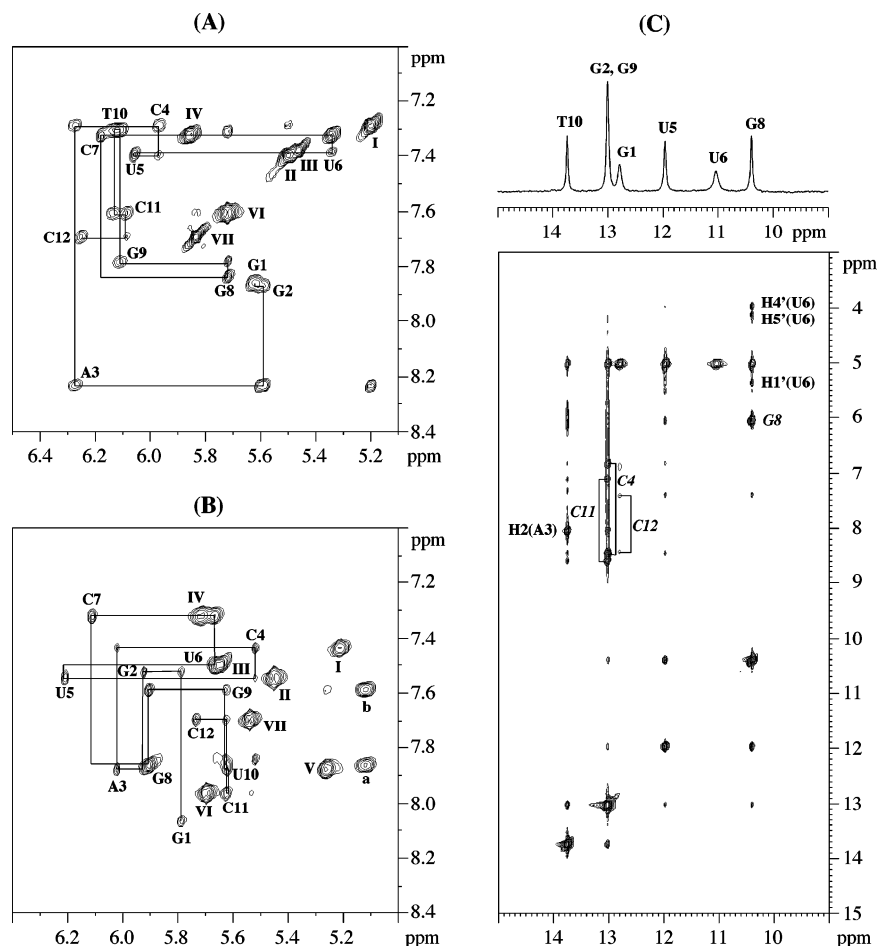(35) Kim, S. G.; Lin, L. J.; Reid, B. R. *Biochemistry* **1992**, *31*, 3564−3574.

Table 2. Coupling Constants (Hz)^a and Calculated Sugar Puckers for Hairpins under Study

residue	hairpin	$J_{1'2'}$	$J_{1'2''}$	$J_{2'3'}$	$J_{2''3'}$	$J_{3'4'}$	$J_{H2'/H3'-P}$	south/north^b
G1	D*R*D	9.7	5.3	5.5	<2	2.5	4	10/0
	R*R*R	<2		4		8	8	0/10
	D*R*D	9.5	5.0	5.5	n.d.	n.d.	5 ± 2	10/0
G2	D*R*D	9.5	5.2	5.5	<2	<2	4	10/0
	R*R*R	<2		n.d.		n.d.	8 ± 2	0/10
	D*R*D	sum = 13.8		n.d.	n.d.	n.d.	5 ± 2	7/3
A3	D*R*D	9.8	5.4	6	<2	2.5	4	10/0
	R*R*R	<2		n.d.			8 ± 2	0/10
	D*R*D	8.5	5.0	6	4	5	6	7/3
C4	D*R*D	9.7	5.5	6	<2	3.5	4	10/0
	R*R*R	<2		4		n.d.	8 ± 2	0/10
	D*R*D	sum = 10.5		n.d.	n.d.	n.d.	6 ± 2	1/9
U5	D*R*D	3.8		7		5	8	4/6
	R*R*R	3.0		7		n.d.	8 ± 2	3/7
	D*R*D	<2		5		9	7	0/10
U6	D*R*D	6.0		5		3	8	9/1
	R*R*R	6.0		5		4	8	9/1
	D*R*D	8.4		5		<2	8	10/0
C7	D*R*D	5.7		6		3	7	8/2
	R*R*R	5.5		6		4	7	8/2
	D*R*D	8.1		4.5		<2	8	10/0
G8	D*R*D	<2		4.5		8.5	8	0/10
	R*R*R	3.5		7		6	7	3/7
	D*R*D	<2		5.5		9	4	0/10
G9	D*R*D	9.7	5.5	6	<2	<2	3	10/0
	R*R*R	<2		n.d.		n.d.	8 ± 2	0/10
	D*R*D	n.d.	n.d.	n.d.	n.d.	4 ± 2	n.d.	
T10	D*R*D	9.7	5.5	6	<2	<2	4	10/0
U10	R*R*R	<2		n.d.		n.d.	8 ± 2	0/10
T10	D*R*D	4.6	7.2	8	5	6	5	4/6
C11	D*R*D	9.3	5.2	6.5	3	4	5	9/1
	R*R*R	<2		4		n.d.	8 ± 2	0/10
	D*R*D	4.8	7.5	8	5	6	6	4/6
C12	D*R*D	sum = 13.7		n.d.	n.d.	n.d.	n.d.	7/3
	R*R*R	2.0		5		n.d.		0/10
	D*R*D	sum = 13.3		n.d.	n.d.	n.d.		6/4

^{a} Error is ±0.5 Hz for $J_{1'2'}$ and $J_{1'2''}$ and is ±1 Hz for other coupling constants if it is not defined specially; n.d. = not determined; sum = ($J_{1'2'}$ + $J_{1'2''}$). ^{b} Fractions (ratio) of sugar conformers.

D$_2$O by using the cross-peaks of complementary base pairs which include NH(G) with NH$_2$(C) or NH(T,U10) with H2(A3) (Figure 2C). Signals of NH$_2$(C) were easily identified from their strong cross-peaks with H5(C) of the same nucleotide residue. Phosphorus signals were determined from strong H3′−P cross-peaks in H,P−HetCOSY spectra. Proton and phosphorus chemical shifts for all hairpins are given in the Supporting Information (Table B).

The 2′,5′-RNA (UUCG) loops in both **D**R**D** and **R**R**R** hairpins share similar chemical shifts and general NOE contacts (Figure 3A) demonstrating that they adopt a common, unusually unique structure that is distinct from that of the native 3′,5′-RNA loop. A strong NOE was detected between U5 and G8 imino protons, and medium−weak NOEs of U5 and G8 imino signals with G9 imino proton as well as G8 imino proton with H1′/H4′/H5′(U6) were found (Figure 2C). These confirm that U5 and G8 loop residues form a *wobble* base pair[34] that appears as a continuation of the stem. Unlike the native RNA loop structure,[14−16] the glycosidic bond of G8 in 2′,5′-RNA loop adopts the *anti*-conformation. Furthermore, the aromatic to sugar H1′/H2′ NOE, typical of helical strands, shows that uracil U6 base stacks above U5 base in the 2′,5′-RNA loop (Figure 3A). Sequential intraloop contacts between U6−C7−G8 residues include few general NOEs: H1′(U6)−H6/H5/H3′(C7) and H3′(C7)−H8(G8). However, despite all structural similarities,

Figure 2. Expanded plots of NOESY spectra at 500 MHz: (A and B) hairpins **DRD** and **RRR** in D₂O, 15 °C, mixing time t_m = 200 ms. The assignment of oligonucleotide protons are shown by solid lines and nucleotide name with number. **I−VII**: H5−H6(C4,U5,U6,C7,U10,C11,C12) cross-peaks respectively. The letter marks **a** and **b**: H2′−H8(G8) and H2′(G8)−H8(G9) cross-peaks; (C) hairpin **DRD** in H₂O/D₂O, 5 °C, with a 1D spectrum of imino protons at the top; cross-peaks for amino protons of cytidine residues and G8 are labeled by marks in italic font.
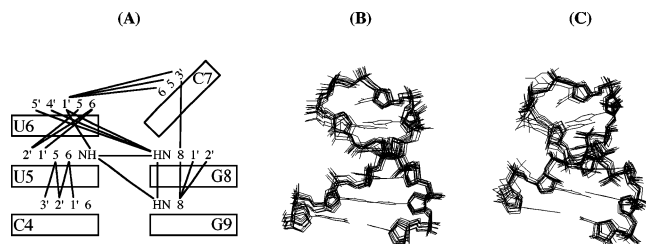

Figure 3. General inter-residue NOE contacts observed within the superstable 2′,5′-RNA loops in **DRD** and **RRR** hairpins in panel A. Superimposition of 10 final individual structures (sugar−phosphate heavy atoms only) and average minimized structure (with nucleotide bases) for hairpin **DRD** in panel B and hairpin **RRR** in panel C.

the 2′,5′-RNA loops in hairpins **DRD** and **RRR** still show a minor difference which manifests itself in the G8 residue of hairpin **RRR** sometimes existing in the *syn*-conformation. Indeed, the intraresidual H8−H1′ and H3′−H8 (or H2′−H8) NOE cross-peaks of G8 are of strong intensity at any mixing time at low temperatures (Figure 2B), which cannot be realized for a single loop conformation. This fact indicates that there exists some degree of mobility in the G8-residue of hairpin **RRR**.

Hairpin **DRD** (with a 3′,5′-RNA loop) exhibits unusual values in H2′(U5), H4′/H5′(C7), H3′(G8), and H1′/5′P(G9) chemical shifts which are strikingly similar to those observed for the extrastable all-RNA hairpin (**RRR**).[14−16] With many other

spectral features in common between loops of **DRD** and **RRR** such as typical NOE cross-peaks (imino(G8)−H1′/H2′(U5), H2′(U5)−H5(C7), etc), *syn*-conformation of G8, and other spectral data (Table 2), we propose that the structure of the native 3′,5′-RNA (UUCG) loop is conserved in both RNA and DNA stem hairpins. This follows the same pattern for the 2′,5′-RNA loop which also shows a conserved uniquely folded structure in both **DRD** and **RRR** hairpins.

Sugar Ring Conformation. Sugar conformations were determined from ³J_{HH} data obtained from DQF-COSY experiments (Table 2). Coupling constants $J_{1'2'}$ and $J_{1'2''}$ were determined from H1′−H2″ or H1′−H2′ cross-peak splittings, whereas $J_{2'3'}$, $J_{2''3'}$, and $J_{3'4'}$ were extracted from the sums of coupling constants as previously described.[32,36] The mole fractions of conformers (*south/north*) were determined using the PSEUROT program[33] with a fixed puckering amplitude of 37° or using the equation[37] for cases of *J*-coupling sum measured only: %(*south*) = $(J_{1'2'} + J_{1'2''} - 9.8)/5.9$.

Data analysis shows that the ribose sugars of U6 and C7 residues in all three hairpins adopt a *south* (C2′-*endo*/C1′-*exo*) pucker, while U5 and G8 exist predominantly in a *northern* (C3′-*endo*) pucker, a pattern that is similar to that observed in

---

(36) van Wijk, J.; Huckriede, B. D.; Ippel, J. H.; Altona, C. *Methods Enzymol.* **1992**, *211*, 286−306.

(37) Rinkel, L. J.; Altona, C. *J. Biomol. Struct. Dyn.* **1987**, *4*, 621−649.

Table 3. Structural Statistics for the 10 Final Individual Structures of D*R*D and R*R*R Hairpins

parameter	D*R*D	R*R*R
number of NOE distance restraints	152	120
intranucleotide	92	64
internucleotide	60	56
loop UUCG residues	49	47
torsion angle restraints	94	94
hydrogen bond restraints	13	13
NOE violation (>0.2 Å)	0	0−1
backbone angle violation (>5°)	0−1	0−1
RMSD for all heavy atoms (Å)		
relative to the average structure:		
for stem residues[a]	0.4−0.5	0.3−0.5
for loop UU residues	0.4−0.5	0.3−0.4
for loop CG residues	0.6−0.7	0.7−0.8
average RMSD from covalent geometry:		
bond lengths (Å)	0.0102	0.0092
angles (deg)	1.56	1.58
impropers (deg)	0.74	0.55

[a] Except for terminal G1−C12 base pair (RMSD 0.5−0.8 Å).

RNA hairpins with 3′,5′-linked (UUCG) or (GCAA) loops.[15,16,38] The DNA stem sugar residues in **D*R*D** exist predominantly in the *south* conformation (typical *B*-form), while those of the RNA stem in **R*R*R** adopt the *north* pucker as for *A*-type helices. By contrast, the majority of deoxyribose sugars in **D*R*D** exist as a mixture of *north*/*south* conformations, suggesting a significant degree of flexibility within the stem residues. This is further corroborated by the average values of $^3J_{H3′−P}$ coupling constants for the DNA stem in this hairpin (5−6 Hz, Table 2). It suggests an incompatibility between the native 3′,5′-RNA (UUCG) loop and the DNA stem which might be responsible for the loss in thermal stability observed for the **D*R*D** hairpin (Table 1). DNA-stem flexibility observed for **D*R*D** prevented us from obtaining a detailed spatial structure for this hairpin.

Structure Refinement and Analysis. To determine the detailed spatial structures of hairpins **D*R*D** and **R*R*R**, NMR-restrained molecular dynamics calculations were performed. The structural statistics of X-PLOR refinement are presented in Table 3, and superimposition of the best final individual structures for both hairpins is shown in Figure 3B and C. Starting from *A*- and *B*-type structures (an overall RMSD of 1.8 Å for stem heavy atoms) with randomly organized loop conformations, both hairpins **D*R*D** and **R*R*R** were refined to an RMSD of 0.3−0.5 Å for stem residues (except of terminal base pair) as well as U5 and U6 nucleotides. The loop C7 residue and sugar−phosphate part of G8 nucleotide in both hairpins are the most flexible with an RMSD of 0.6−0.8 Å. This reflects big variations of backbone torsions between C7 and G8 in hairpin **R*R*R** (Table 4). The stacking between C4:G9 and U5:G8 base pairs was confirmed in both hairpin structures and helps explain the remarkable stability of the 2′,5′-linked RNA loop. In both **D*R*D** and **R*R*R**, U6 base stacks with U5 base, thus contributing to stabilization of the loop and consequently the overall hairpin structure.

The final individual structures of hairpins **D*R*D** and **R*R*R** were averaged, and energy minimized for the purpose of comparing helical parameters and stereoviews of both loop structures. As shown in Figure 4, both loop structures are very similar. The base of C7 is farther away from the G8 residue in hairpin **R*R*R** (Figure 4). This probably affects the G8 base and

(38) Heus, H. A.; Pardi, A. *Science* **1991**, *253*, 191−194.

makes it more mobile in hairpin **R*R*R** compared to the same residue in hairpin **D*R*D**. The minor groove width of the RNA stem is 2−2.5 Å bigger in comparison with that of the DNA stem (Table 5). As expected, the minor groove width and helical parameters of the RNA stem are closer to *A*-type DNA, while those of the DNA stem hairpin are closer to *B*-type DNA.

Discussion

The ability of 2′,5′-linked nucleic acids to form hairpins, and particularly ordered loop structures, is largely unexplored. We have recently shown that hairpin structures containing a 2′,5′-linked RNA loop are more thermodynamically stable than hairpins comprised of 3′,5′-RNA loops.[18,19] Inspection of the thermodynamic parameters revealed that a major determinant of loop stabilization is the enthalpy (versus entropy) of hairpin formation, which indicates better stacking and pairing interactions. The 2′,5′-linked loop can promote double helix formation between complementary strands that are normally incapable of stable hybridization. For example, the 2′,5′-linked r(UUCG) loop can be used to induce the formation of a DNA:2′,5′-RNA hybrid stem.[18] Another interesting feature of the 2′,5′-r(UUCG) tetraloop is that, relative to the native 3′,5′-linked RNA tetraloop,[14−16] its stability is less dependent on sugar stem composition (DNA: DNA, RNA:RNA, 2′,5′-RNA:2′,5′-RNA, etc). For example, the T_m range of the 2′,5′-r(UUCG) loop-hairpins is 16 °C, while that of the native 3′,5′-r(UUCG) loop-hairpins is 26 °C.[19]

Here, we show that the stable 2′,5′-RNA (UUCG) loop (abbreviated **R**) folds in a unique pattern that is conserved in both **D*R*D** and **R*R*R** hairpins. Its structure is completely different from that of the native 3′,5′-linked C(UUCG)G RNA loop[14−16] and constitutes a novel RNA structural motif. Specifically, the 2′,5′-linked r(UUCG) loop is stabilized by a U5:G8 wobble base pair and a stacking interaction between U5 and U6 (Figures 4 and 5). The U5:G8 base pair stacks above the C4:G9 loop-closing base pair and thus appears as a continuation of the stem. Unlike the native 3′,5′-RNA loop structure, G8 in the 2′,5′-RNA loop is found in the *anti*-conformation while the U6 base participates in loop stabilization. Furthermore, the amino group of the C7 base in **R*R*R** contacts the phosphate group linking U5 and U6, whereas the corresponding C7 residue in **R*R*R** is exposed to the solvent, outside the loop (Figure 5).

The proposed conformation for the 2′,5′-RNA loop is in strong agreement with the pattern of chemical base modification[18] and correlates with thermal melting data. Changing the loop base sequence from C(*UUCG*)G to C(*UACG*)G in **R*R*R** results in significantly diminished thermal stability ($\Delta T_m = 7$ °C, $\Delta\Delta G°_{37} = 1.9$ kcal/mol). This is further supported by the NMR structure, which suggests the loss of a stacking interaction between U5 and U6 bases. However, changing C(*UUCG*)G ($T_m = 69.3$ °C) to C(*UUUG*)G ($T_m = 68.0$ °C) does not affect the thermal stability of the hairpin, which is consistent with the notion that the C7 base protrudes out into the solvent and does not participate in any loop stabilizing interactions. Mutating the G8 residue to a uracil (to yield the homopolymeric C(*UUUU*)G loop, $T_m = 60.5$ °C) strongly destabilizes the native 2′,5′-RNA loop structure by abolishing U5:G8 *wobble* base pairing. The same features were demonstrated for the 2′,5′-RNA loop hairpins with a DNA stem (**D*R*D**).[18]

The 2′,5′-RNA loop has two unpaired nucleotides, that is, U6 and C7, that adopt the *south* (C2′-*endo*) conformation. A

Table 4. Ranges of Backbone[a] and Glycosidic Torsion Angles in the 10 Final Individual Structures of **D***R***D** and **R***R***R** Hairpins

| residue	hairpin	α	β	γ	ε	ζ	χ
G1	**D***R***D**				160 ± 4	−84 ± 4	−119 ± 7
	R*R***R**				−157 ± 11	−69 ± 8	−170 ± 6
G2	**D***R***D**	−83 ± 4	−161 ± 4	57 ± 3	174 ± 3	−81 ± 6	−126 ± 3
	R*R***R**	−74 ± 10	−177 ± 11	61 ± 6	−171 ± 2	−80 ± 6	−158 ± 6
A3	**D***R***D**	−78 ± 8	−169 ± 5	56 ± 3	177 ± 7	−80 ± 4	−119 ± 4
	R*R***R**	−98 ± 11	−156 ± 8	57 ± 4	−146 ± 8	−56 ± 8	−158 ± 6
C4	**D***R***D**	−76 ± 6	−175 ± 4	55 ± 3	171 ± 6	−86 ± 6	−120 ± 5
	R*R***R**	−67 ± 7	168 ± 9	58 ± 2	−163 ± 3	−69 ± 3	−161 ± 2
U5	**D***R***D**	−86 ± 7	−168 ± 9	56 ± 3	−100 ± 7	−94 ± 5	−131 ± 8
	R*R***R**	−83 ± 7	−176 ± 6	61 ± 3	−93 ± 6	−90 ± 4	−141 ± 3
U6	**D***R***D**	−67 ± 5	174 ± 3	58 ± 2	−157 ± 3	−97 ± 4	−121 ± 2
	R*R***R**	−74 ± 6	176 ± 8	58 ± 4	−170 ± 4	−92 ± 4	−115 ± 3
C7	**D***R***D**	180 ± 5	117 ± 2	77 ± 3	−74 ± 3	−80 ± 7	−122 ± 2
	R*R***R**	172 ± 4	119 ± 4	74 ± 6	−121 ± 53	21 ± 81	−117 ± 4
G8	**D***R***D**	−54 ± 5	161 ± 8	61 ± 2	−126 ± 17	−68 ± 6	−127 ± 6
	R*R***R**	−124 ± 52	171 ± 21	64 ± 3	−119 ± 4	−93 ± 12	−109 ± 9
G9	**D***R***D**	−90 ± 14	−158 ± 20	59 ± 5	169 ± 7	−80 ± 3	−121 ± 8
	R*R***R**	−68 ± 10	163 ± 12	58 ± 3	−164 ± 6	−73 ± 4	−167 ± 7
T10	**D***R***D**	−83 ± 7	−162 ± 10	56 ± 2	174 ± 3	−85 ± 4	−120 ± 5
U10	**R***R***R**	−83 ± 7	−163 ± 6	56 ± 5	−141 ± 11	−76 ± 7	−154 ± 5
C11	**D***R***D**	−74 ± 6	−175 ± 4	57 ± 2	170 ± 2	−81 ± 3	−120 ± 3
	R*R***R**	−66 ± 5	150 ± 7	65 ± 4	−167 ± 3	−74 ± 5	−156 ± 4
C12	**D***R***D**	−79 ± 4	−166 ± 4	55 ± 2			−115 ± 3
	R*R***R**	−73 ± 8	−179 ± 9	62 ± 3			−150 ± 6

[a] Backbone for 2′,5′-links includes C2′ and O2′ atoms instead of regular C3′ and O3′ atoms (epsilon torsion is between C3′−C2′−O2′−P atoms).

(A)

(B)
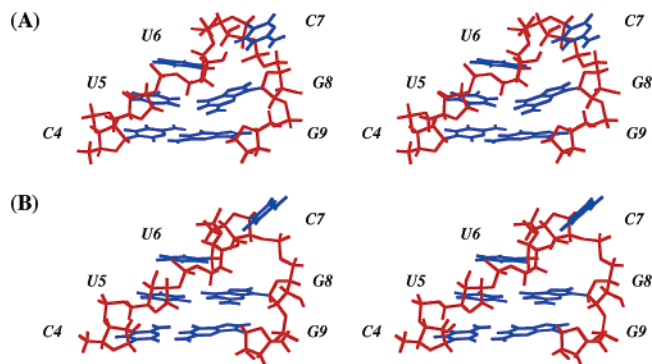
Figure 4. Stereoview of the loop regions for average minimized structures of hairpins **D***R***D** and **R***R***R** (in panels A and B, respectively).

Table 5. Ranges[a] of Minor Groove Widths and Helical Parameters for Stem Base Pairs in Average Minimized Structures of **D***R***D** and **R***R***R** Hairpins and in Canonical *A*- and *B*-Types of DNA

structure	minor groove width (Å)	X-displace-ment (Å)	inclination (deg)	rise (Å)	helical twist (deg)
D*R***D**	7.7 ± 0.2	−2.6 ± 0.2	7.7 ± 0.8	2.9 ± 0.2	34.8 ± 2.1
R*R***R**	10.1 ± 0.2	−2.4 ± 0.8	13.0 ± 1.8	2.6 ± 0.3	39.8 ± 2.4
B-type DNA	5.9	−0.7	−6.0	3.4	36.1
A-type DNA	11.1	−5.4	19.3	2.6	32.7

[a] Average values with corresponding deviations for the set of stem base pairs are shown.

similar sugar pucker is observed for the corresponding 3′,5′-RNA loop residues (U6 and C7), yet the sugar−phosphate distances of individual nucleotides in this motif are quite different. A C2′-*endo* nucleotide has a "compact" shape in 2′,5′-RNA but is "extended" in 3′,5′-RNA.[20,39] These differences arise from a switch in the equatorial to axial placement of phosphate groups linking C2′−O2′ versus C3′−O3′ bonds on a C2′-*endo* oriented sugar frame.[39] On the other hand, U5 and G8 residues

(39) Premraj, B. J.; Yathindra, N. *J. Biomol. Struct. Dyn.* **1998**, *16*, 313−328.
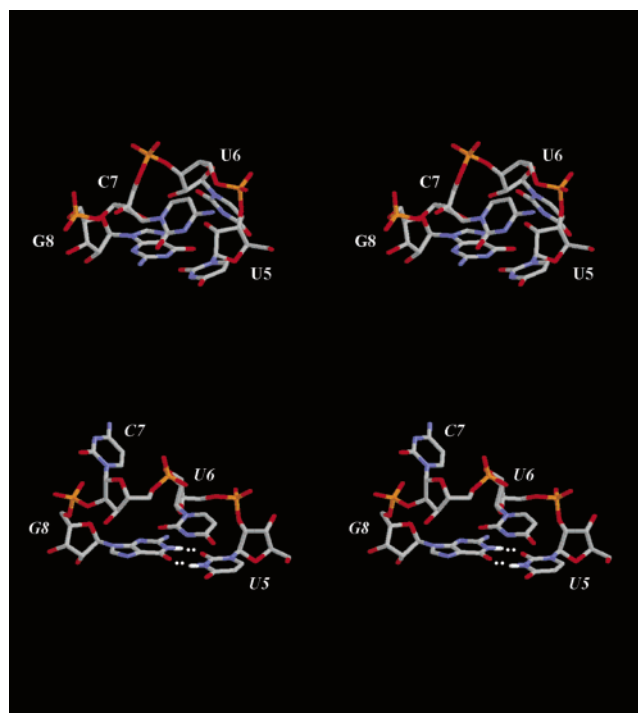
Figure 5. Stereoview of the native 3′,5′-RNA (top) and 2′,5′-RNA (bottom) loops in **R***R***R** (PDB ID code = 1HLX) and **R***R***R** hairpins. Hydrogen bonds between *wobble* base paired U5 and G8 in **R***R***R** are shown.

are characterized by *north* (C3′-*endo*) sugar puckers in both motifs and, therefore, adopt the "compact" conformation in the 3′,5′-RNA loop, whereas they are "extended" in the 2′,5′-RNA loop. To summarize, the 2′,5′- and 3′,5′-loop residues adopt, respectively, the U5(**extended**)−U6(compact)−C7(compact)−G8(**extended**) and U5(compact)−U6(**extended**)−C7(**extended**)−G8(compact) geometries. In both cases, the "extended" residues make it possible to bridge the stem to the loop without imposing unfavorable steric constraints.

Despite a great degree of similarity between the 2′,5′-loops of both **R***R***R** and **D***R***D** hairpins, some differences are evident

in the backbone, particularly in the ϵ, ζ, α torsional angles between C7 and G8 and in the ζ, α, β torsional angles between G8 and G9 residues (Table 4). Also the calculated helical twist between the U5:G8 and C4:G9 base pairs is larger in the average structure of **D*R*D** (48°) than in **R*R*R** (39°). The other helical parameters for these base pairs are virtually identical. These differences in backbone angles and helical parameters between **R*R*R** and **D*R*D** allow the 2′,5′-r(UUCG) loop to maintain its folded (extrastable) structure (Table 5).

Our data show that the 3′,5′-RNA loop in **DRD** exhibits the same NOEs as those observed for **R*R*R**.[14−16] This suggests that the loop structure remains the same irrespective of whether the stem is duplex DNA or duplex RNA. Why, then, is **D*R*D** extrastable but **DRD** is not? A close inspection of the sugar *J*-couplings reveals that the deoxyribose residues in **D*R*D** are highly preorganized in a rigid *south* pucker characteristic of *B*-DNA. This contrasts strongly with the dynamic sugar conformations observed in the DNA strands of **DRD**. As a consequence, the thermodynamic stability of **D*R*D** is higher than that for **DRD** because preorganization of the deoxyribose sugar would entropically favor folding or duplex formation. These findings draw attention to the inter-relation between loop structure and dynamics of the helical stem.

In conclusion, the 2′,5′-RNA (UUCG) loop displays a uniquely folded and well-ordered structure that is distinct from that of its natural 3′,5′-linked counterpart. The unique folding of the 2′,5′-RNA (UUCG) loop along with its inherent stacking interactions gives it a compact shape that confers unusual thermodynamic stability. Recent findings show that 2′,5′-RNA loops are significantly more resistant toward nuclease degradation compared to 3′,5′-RNA loops. In addition, the 2′,5′-RNA loop structure is recognized by HIV-1 reverse transcriptase (RT) and can act as a potent inhibitor of RNase H activity of HIV-1 RT (IC_{50} 30−50 μM) when present within the appropriate hairpin stem [Hannoush, R. N.; Min, K.-L.; Carriero, S.; Damha, M. J. In preparation]. This discovery may help in the design of new RNA-based aptamers or ribozymes and identifies the 2′,5′-linked r(UUCG) loop as a novel RNA structural motif.

Supporting Information Available: A table of isolated yields from solid-phase synthesis and a table containing the chemical shifts of proton and phosphorus resonances for all three hairpins. This material is available free of charge via the Internet at http://pubs.acs.org.

JA036207K