Stability of Chimeric DNA/RNA Cytosine Tetrads: Implications for *i*-Motif Formation by RNA

Delphine Collin[†] and Kalle Gehring*

Contribution from the Department of Biochemistry and Montreal Joint Centre for Structural Biology, McIntyre Medical Science Building, McGill University, 3655 Drummond, Montréal, QC, Canada, H3G 1Y6

Received September 24, 1997

Abstract: UV melting curves of mono- and di-2′-hydroxylated cytosine tetrads were recorded. The substitution of one DNA residue by RNA decreased the melting temperature (T_m) by roughly 6.5 °C. In sequences with two RNA residues, strong positional effects were observed. Juxtaposition of the two D-ribose sugars in the tetrad incurred an additional loss of 6 °C in tetrad stability and decreased the T_m by 19 °C from that of the parent compound, d(TCCCCC)$_4$. The effect of increasing the number of cytosines on tetrad stability was also measured. Addition (or removal) of one cytosine-protonated cytosine base-pair from the tetrad increased (or decreased) the T_m by roughly 9 °C. Thus, the penalty of adding an RNA residue is equal to or greater than the benefit of the additional base-pair. This implies that the stability of RNA cytosine tetrads should not rise with increasing length and that thermal transitions previously observed in poly(rC) correspond to duplex structures.

Introduction

The first intercalated four-stranded nucleic acid structure (*i*-motif) was reported in 1993.[1] In the structure, four strands of the DNA oligomer, d(TC$_5$), associate pairwise at low pH to form two parallel duplexes with cytosine-protonated cytosine (C·C$^+$) base-pairs. The two duplexes are themselves associated in an antiparallel fashion through mutual intercalation of the ten C·C$^+$ base-pairs. The resulting complex resembles two ladders with criss-crossed rungs and was called the *i*-motif to emphasize the intercalation. The antiparallel arrangement of the ladders means that base-pair contacts are all interduplex and follow an order 1-**6**-2-**5**-3-**4**-4-**3**-5-**2**-6-**1** where bases from one duplex are underlined. Viewed down the helical axis, the tetrad has two large and two narrow grooves. Intimate sugar−sugar contacts exist in the narrow groove between residues belonging to different (antiparallel) strands. Since the first report, a number of similar oligocytidine sequences have been shown to form stable *i*-motif tetrads[2−4] and several X-ray structures reported.[5−8]

Surprisingly, small RNA oligomers do not seem to form a stable *i*-motif.[9,10] Differences in structural stability between homopolymer forms of poly(dC) and poly(rC) at low pH were first described[11] in the 1960s as the transition temperatures reported for poly[r(C)][12−14] at low pH were significantly lower than the T_m for poly[d(C)].[15] The origin of this difference between RNA and DNA complexes with C·C$^+$ base-pairs has not been systematically investigated, and it is not known whether longer RNAs can associate into the *i*-motif. Here, we report a study of the relative thermal stabilities of DNA and mixed DNA/RNA tetrads. The results allow us to quantify the effects of adding 2′ hydroxyls to the *i*-motif and to make melting point predictions for longer DNA, mixed DNA/RNA, and RNA sequences.

Material and Methods

Nomenclature. The sequences studied are shown in Table 1. The term *i*-motif refers to an arrangement of intercalated C·C$^+$ base-pairs.

Oligomer Synthesis and Purification. Oligonucleotides were synthesized according to standard solid-phase phosphoramidite chemistry on a 1 μmol scale with a Millipore Cylone Plus DNA synthesizer. β-D-Arabinocytidine and 2′-*O*-methylcytidine phosphoramidites were purchased from Chemgenes (Waltham, MA) and PerSeptive Biosystems (Framingham, MA), respectively. Deoxyoligonucleotides were cleaved from the CPG column resin and deprotected by incubation for 24 h at room temperature in concentrated (30%) ammonium hydroxide. For oligonucleotides with D-ribose sugars, cleavage and deprotection were performed by heating the resin in a sealed tube with 1.5 mL of dry

[†] Permanent address: Laboratoire de Résonance Magnétique Nucléaire, Institut de Chimie des Substances Naturelles, Centre National de la Recherche Scientifique, Gif-sur-Yvette, Cedex 94128, France.

* Corresponding author. Telephone: (514) 496-2558. Fax: (514) 496-5143. E-mail: Kalle.Gehring@bri.nrc.ca.

(1) Gehring, K.; Leroy, J.-L.; Guéron, M. *Nature* **1993**, *363*, 561−565.
(2) Ahmed, S.; Kintanar, A.; Henderson, E. *Nature Struct. Biol.* **1994**, *1*, 83−88.
(3) Benevides, J. M.; Kang, C.; Thomas, G. J. *Biochemistry* **1996**, *35*, 5747−5755.
(4) Leroy, J.-L.; Gehring, K.; Kettani, A.; Guéron, M. *Biochemistry* **1993**, *23*, 6019−6031.
(5) Berger, I.; Kang, C. H.; Fredian, A.; Ratliff, R.; Moyzis, R.; Rich, A. *Nat. Struct. Biol.* **1995**, *2*, 416−424.
(6) Kang, C. H.; Berger, I.; Lockshin, C.; Ratliff, R.; Moyzis, R.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3874−3878.
(7) Kang, C. H.; Berger, I.; Lockshin, C.; Ratliff, R.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11636−11640.
(8) Chen, L.; Cai, L.; Zhang, X.; Rich, A. *Biochemistry* **1994**, *33*, 13540−13546.
(9) Robidoux, S.; Damha, M. J. *J. Biomolec. Struct. Dynamics* **1997**, *14*, 804.
(10) Lacroix, L.; Mergny, J.-L.; Leroy, J.-L.; Helene, C. *Biochemistry* **1996**, *35*, 8715−8722.
(11) Michelson, A. M.; Massoulié, J.; Guschlbauer, W. *Prog. Nucl. Acid Res., Mol. Biol.* **1967**, *6*, 83−141.
(12) Akirimiski, E. O.; Sander, C.; Ts'o, P. O. P. *Biochemistry* **1963**, *2*, 340−344.
(13) Hartman, K. A., Jr.; Rich, A. *J. Am. Chem. Soc.* **1965**, *87*, 2033−2039.
(14) Rich, A.; Davies, D. R.; Crick, F. H. C.; Watson, J. D. *J. Mol. Biol.* **1961**, *3*, 71−86.
(15) Inman, R. B. *J. Mol. Biol.* **1964**, *9*, 624−637.

S0002-7863(97)03346-5 CCC: $15.00
Published on Web 04/15/1998

Table 1. Melting Temperatures of Cytosine-Tetrad Oligomers

Cpd.	Sequence[a] (1 2 3 4 5 6)	Tm[b]	Tm[c]	Cpd.	Sequence[a] (1 2 3 4 5 6 7 8)	Tm[b]	Tm[c]
1	T C C C C C	49	48	**10**	U^{R}C^{R}C^{R}C^{R}C^{R}C^{R}	< 4	
2	TC^{R}C C C C	41	40.5	**11**	T C C C C C C C	64.5	64
3	T C C^{R}C C C	41	40	**12**	T C C C C C C	59	56
4	T C C C^{R}C C	45	44.5	**13**	T C C C C	39	36.5
5	T C C C C^{R}C	40.5	39	**14**	T C C C	17.5	15.5
6	T C C^{R}C C^{R}C	35	34	**15**	T C C C^{R}C^{R}C C	43	41
7	TC^{R}C C C^{R}C	30	27	**16**	T CC^{A}C C C	49	
8	T C C^{R}C^{R}C C	29.5	26.5	**17**	T CC^{M}C C C	< 4	
9	T C C C^{R}C^{R}C	41	40				

[a] Sequence of oligomer studied. 2′-Hydroxylated (RNA), D-arabinose, and 2′-*O*-methyl D-ribose residues are shaded and indicated respectively by superscript R, A, or M. [b] Melting temperature determined at a heating rate of 0.25 °C/min. [c] Melting temperature determined at a heating rate of 0.1 °C/min.

ethanol saturated with ammonia gas for 12 h at 55 °C. After the separation from the resin, the supernatant was lyophilized, and 600 μL of tetrabutylammonium fluoride solution (1 M) was added to remove the *tert*-butyl dimethylsilyl group present at the 2′ position of D-ribose ring. The reaction was carried out at room temperature for 24 h before quenching with 600 μL of triethylamine acetate (1 M) pH 7.0 and desalting on Sephadex G-10. Concentrations of oligonucleotides were determined using estimated molar extinction coefficients for the single stranded forms at pH 7.0 and at pH 4.6.[16,17]

NMR Spectroscopy. All NMR experiments were performed on a Bruker DRX 500 spectrometer equipped with a triple resonance, three-axis gradient probe. Sample concentrations were 0.4−0.7 mM in 50 mM sodium citrate buffer, pH 4.6, containing either 10% or 100% D_2O. Exchangeable proton spectra were recorded at 5 °C with a 15 000 Hz spectral width using a hybrid jump-and-return[18] WATERGATE[19] water suppression sequence with interpulse delays of 50 and 100 μs (for an excitation maximum at 14.7 ppm). Two-dimensional NOESY spectra of the nonexchangeable proton resonances were recorded with 4000 Hz spectral width and TPPI quadrature detection[20] at 20 °C. The residual water signal was selectively saturated during the 200 ms mixing time and the 1.4 s recycle delay. Proton resonances were referenced to the water resonance at 4.7 ppm. Spectral data were processed with the GIFA 4.10 program[21] on a Silicon Graphics Indigo 2 workstation.

UV Absorbance Melting Experiment. Ultraviolet absorption measurements were performed on a Cary 3E UV/visible spectrophotometer, equipped with thermostated cell holder. Denaturations were monitored at 270 nm by increasing the temperature with ramps of 1, 0.5, 0.25, and 0.1 °C/min. Buffer conditions were 50 mM sodium citrate, pH 4.6. Oligonucleotides concentrations were 11.5 ± 0.5 μM, which were determined from optical densities at 270 nm, pH 4.6, 20 °C of 0.52 OD for **1**−**10**, **16**, and **17** and 0.71 OD, 0.60 OD, 0.44 OD, 0.35 OD, and 0.60 OD for **11**−**15**, respectively.

Results and Discussion

NMR Spectra of 2′-Hydroxylated *i*-Motif DNA. All of the sequences were studied by 1D and 2D NMR spectroscopy at 5 °C and pH 4.6 in H_2O to confirm the presence [or absence in the case of r(UC_5)] of the *i*-motif before monitoring the structural transition by UV spectroscopy. The 1D spectrum of **6** shows protonated cytosine imino resonances at 15−16 ppm and a single thymine imino resonance at 11 ppm (Figure 1). Strong nonsequential imino−imino and H1′−H1′ NOE cross-peaks confirm the formation a single *i*-motif structure.[1,4] Immediately following synthesis, all the oligomers showed multiple thymine imino resonances, indicating the presence of multiple low pH conformers. After 1 week at 4 °C, the samples matured to give the spectrum of a single *i*-motif tetrad with one thymine imino peak. Oligomers **7** and **8** were an exception, and multiple, minor species were detected even after long periods of storage.

UV Melting Curves. Table 1 shows the compounds studied and their various melting points as established by UV spectroscopy (Figure 2). Thermal transitions were monophasic but kinetically irreversible as dissociation of the tetrads showed unusually slow kinetics. A study of the melting curves as a function of the heating rate confirmed that the transition temperatures (T_m) were slightly overestimated in our experiments. For the all DNA *i*-motif, **1**, the measured T_m varied between 56 °C at 2 °C/min and 48 °C at 0.1 °C/min (Figure 3). When measured at 0.25 and 0.1 °C/min, approximately the same error of 1 to 3 °C was present for all the sequences studied. Thus, we conclude that the observed T_m values can be used as a measure of the relative stability of the different *i*-motifs.

It was initially surprising to find that the difference between T_m's measured at the two heating rates was independent of the number of base-pairs in the tetrad. It appears that the intrinsically faster kinetics of the shorter sequences, **13** and **14**, is balanced by their lower transition temperatures. Indeed, for oligomers with five cytosines, **1**−**9**, the largest differences were measured with the two sequences, **7** and **8**, that had the lowest T_m's.

Effect of 2′-Hydroxyl. The presence of ribose residues decreased the stability of the *i*-motif as measured by the transition temperature. Addition of a first RNA residue to **1** decreased the melting point by 7.5, 8, 3.5, or 9 °C depending on position. It is not clear why substitution at residue C_4 (in **4**) is better tolerated, but it may be related to the 2-fold symmetry axis of the tetrad present at C_4. The increased tolerance of a hydroxyl at this position can also be seen in a comparison of the T_m's of the dihydroxylated tetrads, **6** and **9**. Using the average T_m of the monohydroxylated tetrads, we estimate loss of approximately 6.5 °C in *i*-motif stability for addition of the first ribose residue.

Positional effects were more marked in the tetrads with two ribose residues. For **6** and **9**, the T_m was decreased by an average of 11 °C from the T_m of **1**, while for **7** and **8** the T_m was decreased by 21 °C. This is explained by the existence of two different types of sugar−sugar contacts in the *i*-motif of d($TCCCCC)_4$. For sugar residues 1-6, 2-5, and 3-4, short distances exist between the deoxyribose H2″ protons. We term these residues back-to-back (Figure 4). For residue pairs 6-2, 5-3, and 4-4, a face-to-face orientation of the sugars exists in which the H2″ protons are well separated. The additional 7 to 13 °C of destabilization observed in **7** and **8** is due to the juxtaposition of 2′-hydroxyl groups in back-to-back positions. Last, we note that the RNA sequence, **10**, did not form a tetrad at 5 °C.

The positional effects in dihydroxylated tetrads may explain the presence of multiple conformers detected for tetrads **7** and **8**. In the parent tetrad d($TCCCCC)_4$, the order of intercalated bases is 1-**6**-2-**5**-3-**4** (for half of the symmetric structure). Less stable, shifted tetrads could form with the 3′ cytosines capping the structures (to give a base order of **6**-1-**5**-2-**4**-3) or with nonintercalated bases (such as 1-2-**6**-3-**5**-4). These alternatively folded (intercalated) tetrads could be the kinetic intermediates

(16) Voet, A. D.; Gratzer, W. B.; Cox, R. A.; Doty, P. *Biopolymers* **1963**, *1*, 193.
(17) Puglisi, J. D.; Tinoco, I., Jr. In *Methods in Enzymology*; Dahlberg, J. E., Abelson, J. N., Eds.; Academic Press: San Diego, CA, 1989; Vol. 180, pp 304−325.
(18) Plateau, P.; Guéron, M. *J. Am. Chem. Soc.* **1982**, *104*, 7311−7312.
(19) Piotto, M.; Saudek, V.; Skelnár, V. *J. Biomol. NMR* **1992**, *2*, 661−665.
(20) Marion, D.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967−974.
(21) Pons, J.-L.; Malliavin, T. E.; Delsuc, M. A. *J. Biomol. NMR* **1996**, *8*, 445−452.

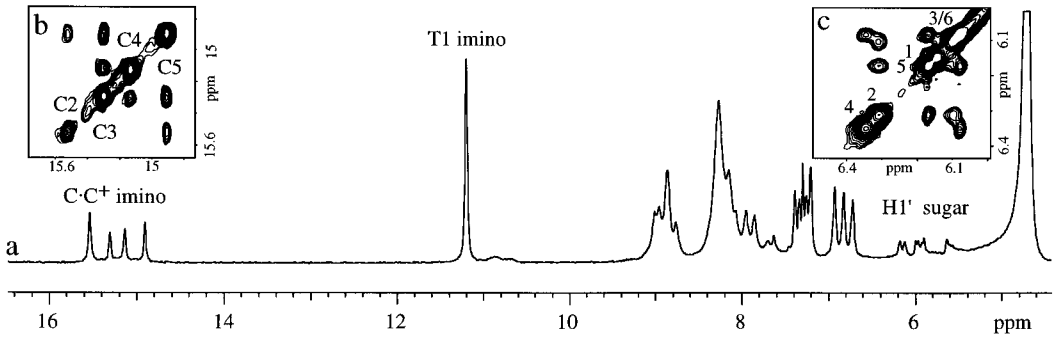

Figure 1. NMR spectra of a dihydroxylated *i*-motif DNA, d(TCC^{r}CC^{r}C)$_4$. (a) 1D exchangeable proton spectrum of **6** showing four protonated cytosine imino resonances at 15 ppm and the single thymine imino at 11 ppm. The imino proton of C6 is not detected due to chemical exchange with solvent. (b) Extract of a NOESY spectrum (180 ms mixing time) showing nonsequential C+ imino cross-peaks. (c) Extract of a NOESY spectrum (200 ms) in D_2O showing characteristic H1′−H1′ NOEs.

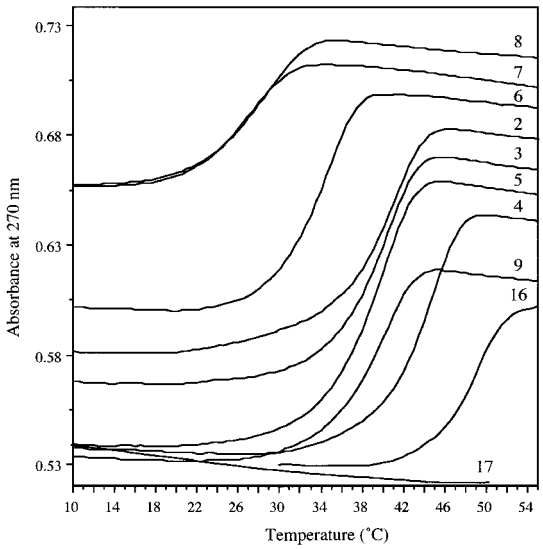

Figure 2. UV melting curves of DNA, RNA, and mixed DNA/RNA chimeras, identified as in Table 1. Melting curves of the monosubstitued D-arabinose and 2′-*O*-methyl oligomers were obtained at a heating rate of 0.25 °C/min; all other curves were recorded at 0.1 °C/min. The curves were offset vertically for clarity.

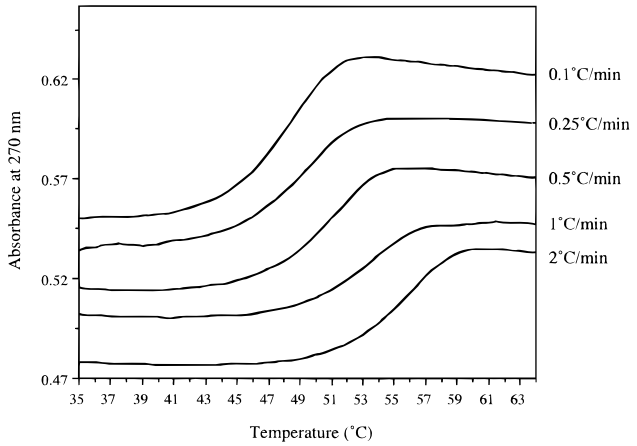

Figure 3. Melting curves of the DNA oligomer d(TC5) recorded at different heating rates. Melting temperatures were 56 °C at 2 °C/min, 53 °C at 1 °C/min, 51 °C at 0.5 °C/min, 49 °C at 0.25 °C/min, and 48 °C at 0.1 °C/min. The curves were slightly offset vertically to clarify the presentation.

detected following oligomer transfer to low pH conditions. In **7** and **8**, back-to-back juxaposition of 2′-hydroxyls occurs only in the unshifted tetrad with 1-**6**-2-**5**-3-**4** intercalation and would not occur in the alternatively folded tetrads. This destabilizes

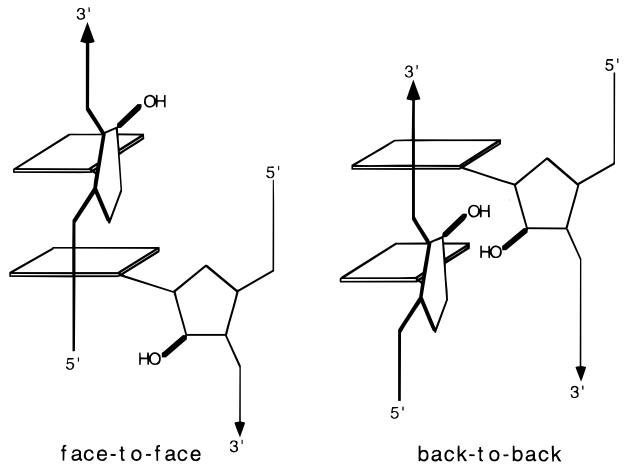

Figure 4. Detail of *i*-motif cytosine tetrads showing ribose residues in face-to-face (left panel) or back-to-back orientations (right panel). Each panel presents a pair of residues viewed from the narrow groove. Because of the antiparallel sense of the strands, the cytidine residues are alternately face-to-face and back-to-back. The ribose 2′-hydroxyls are separated by approximately 0.35 nm in the back-to-back orientation and by 0.65 nm in the face-to-face orientation. Strand polarity is indicated by 3′ and 5′ labels.

the former relative to the latter and could lead to the multiple tetrad forms detected by NMR spectroscopy.

The results from UV spectroscopy were confirmed by a study of migration on nondenaturing polyacrylamide gels (available as Supporting Information). Selected samples, **1**−**3**, **5**−**7**, and **10**, were incubated for 15 min at three different temperatures (4, 35, and 46 °C) before migration at 15 °C. At 4 °C, all species except **10** migrated as a tetrad. At 35 °C, **7** and **10** migrated as single strands, and **2**, **3**, **5**, and **6** migrated as a mixture of single strands and tetrads as expected from their UV melting points. At 46 °C, **6**, **7**, and **10** migrated as single strands, **2**, **3**, and **5** as mixtures, and only **1** migrated uniquely as a tetrad. The ability to resolve the tetrad and single stranded forms by electrophoresis underlines the slow kinetics of tetrad assembly and disassembly.

A Rule for Thermal Stability. Comparison of the T_m's of mono- and di-2′-hydroxylated sequences allowed the establishment of a rule for melting point predictions. For each RNA residue added, the melting point decreases by roughly 6.5 °C. Juxtaposition of the two ribose sugars in a back-to-back orientation incurs an additional loss of 3 °C in tetrad stability so that the melting temperature can be estimated by the following equation: $T_m^{\text{tetrad}} = T_m^{\text{DNA tetrad}} - (n\ 6.5\ °C) - (m\ 3\ °C)$ where n is the number of 2′-hydroxylated bases and m is

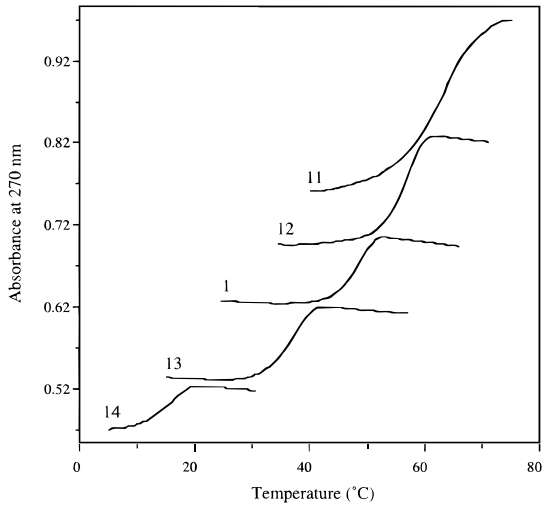

Figure 5. Melting curves of DNA tetrads with between 3 and 7 cytidines per strand. Sequences **1** and **11**−**14** are identified as in Table 1. Melting curves were determined at a heating rate of 0.1 °C/min and have been offset vertically for clarity.

the number of 2′-hydroxylated bases back-to-back. For r(UC_5), $n = 6$ and $m = 6$ so we expect a melting point of about −9 °C which agrees with our inability to observe a melting transition for the RNA oligomer, **10**.

We examined melting curves for varying length DNA sequences (Figure 5) containing between 3 and 7 cytosine DNA residues (Table 1). For each deoxycytidine added, the melting point increased. The magnitude of this stabilization effect decreased with the length of oligomer and was successively 21, 11.5, 8, and 8 °C. Addition or removal of a deoxycytidine from **1** changed the T_m by an average of 9 °C. This is very close to the change in T_m that results from the addition of a sterically hindered 2′-hydroxyl group (6.5 °C + 3 °C). Thus, addition of cytidine (versus deoxycytidine) should change the stability of an *i*-motif very little. Comparison of the T_m's of **13** and the monohydroxylated compounds (which lack back-to-back interactions), **2**, **3**, and **5** shows this to be true. The slight improvement in stability observed for the monohydroxylated compounds is due to the absence of sterically hindered back-to-back interactions. Similarly, comparison of the T_m of the dihydroxylated tetrad, **15**, with the T_m's of shorter tetrads, **2**−**5** and **13**, shows that little stability is gained by addition of one or more ribose residues.

This would not be the case for C·C^+ base-paired duplex structures that do not exhibit the extensive sugar−sugar contacts characteristic of the *i*-motif. Thus it follows that for longer RNA sequences the relative stability of duplex over tetrad (*i*-motif) structures should increase. Putative duplex forms of **1** and **10** have been reported[1,10] and we observed minority species in the NMR spectra of **7** and **8** that may correspond to duplex forms or other types of tetrads. In conclusion, it is most reasonable to assume that poly(rC) complexes previously studied[12,13,22−25] were RNA duplexes and not *i*-motif tetrads.

Other Modifications. We studied two other sequences, **16**, with a β-D-arabinose sugar at position three and, **17**, with a 2′-*O*-methyl-β-D-ribose at the same position. The transition temperature of **16** was exactly the same as the T_m of the parent compound, while **17** failed to show a thermal transition (Figure 2). Robidoux and Damha[9] first showed that incorporation of D-arabinose, the C2′ epimer of ribose, has no effect on the stability of a cytosine tetrad. This is because the hydroxyl in D-arabinose is located in the wide groove of the tetrad and not in the narrow groove. On the other hand, incorporation of a single 2′-*O*-methyl group in the narrow groove completely inhibits tetrad formation and decreases the T_m by >40 °C.

Conclusions

Why is RNA unable to form a stable *i*-motif? It cannot be explained by sugar conformation because the local geometry of each residue in the cytosine tetrad is more like RNA than DNA. The glycosidic angles χ are generally high anti conformation and sugar puckers are mostly C3′-*endo*.[1,5−8,26] Both conformations are typical for RNA A-type duplexes.[27] Instead, steric hindrance between 2′-hydroxyls in the narrow groove is most responsible for the absence of an RNA *i*-motif. This is clearly shown by the positional dependence of *i*-motif stability in the dihydroxylated tetrads, by the absence of effect of a 2′-hydroxyl substitution in tetrads containing arabinose and by the complete intolerance of 2′-*O*-methyl modifications.

Multiple factors contribute to the stability of intercalated cytosine tetrads. For example, the decreased linear electrostatic charge from protonation of cytosine facilitates the association of the four strands. Cytosine tetrads are also stabilized by the formation of three hydrogen bonds at each base-pair. However, these considerations apply equally to duplex C·C^+ structures and do not explain the formation of an intercalated structure. Part of the differential stability of the *i*-motif comes from the close contacts between sugars in the narrow grooves of the tetrad.[1,28] These give rise to favorable van der Waals energies in molecular dynamics simulations and, as they exclude water, may represent hydrophobic contacts between deoxyribose sugars. Thus, even in the absence of steric hindrance, substitution of the 2′ proton by a hydroxyl makes the sugar more hydrophilic, favoring hydration and destabilizing the tetrad.

Acknowledgment. This work was supported by a Canadian Medical Research Council operating grant and salary support to K.G. and a Canadian government foreign exchange student fellowship to D.C.

Supporting Information Available: Three polyacrylamide elecrophoresis gels of oligomers incubated at 4, 35, and 46 °C (1 page). See any current masthead page for ordering and Web access instructions.

JA973346R

(22) Langridge, R.; Rich, A. R. *Nature* **1963**, *198*, 725−728.
(23) Guschlbauer, W. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *55*, 1441−1448.
(24) Brahms, J.; Maurizot, J. C.; Michelson, A. M. *J. Mol. Biol.* **1967**, *25*, 465−480.
(25) Adler, A.; Grossman, L.; Fasman, G. D. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *57*, 423−430.
(26) Leroy, J.-L.; Guéron, M. *Structure* **1995**, *3*, 101−120.
(27) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: Berlin, 1984.
(28) Berger, I.; Egli, M.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12116−21.