Inability of RNA To Form the i-Motif: Implications for Triplex Formation

Laurent Lacroix,[‡] Jean-Louis Mergny,*[,‡] Jean-Louis Leroy,[§] and Claude Hélène[‡]

Laboratoire de Biophysique, Muséum National d'Histoire Naturelle INSERM U201, CNRS UA481, 43, rue Cuvier 75005 Paris, France, and Groupe de Biophysique, Ecole Polytechnique, URA D1254 CNRS, 91128 Palaiseau, France

Received January 17, 1996; Revised Manuscript Received April 3, 1996[⊗]

ABSTRACT: At slightly acidic or even neutral pH, oligodeoxyribonucleotides which include stretches of cytidines form a tetrameric structure involving C·C$^+$ base pairs in a so-called i-motif. Such a structure, which is very stable at pH 6, is still detectable at neutral pH. This motif, whether intramolecular or intermolecular, was shown to act as an undesired, competing structure for triplex formation. Depending on the sequence and the experimental conditions, triple helix formation was inhibited or completely abolished. Thus, the sequence recognition repertoire of pyrimidine DNA third strands is more restricted than previously admitted; only T·A rich-oligopurine-oligopyrimidine sequences are amenable to triplex formation. On the other hand, cytosine-rich RNA oligoribonucleotides and their 2′-*O*-methyl derivatives were unable to form a stable autoassociated structure above pH 6. Concomitantly, they were shown to form a triplex with (C·G)-rich targets at pH 6. Thus, not only was the affinity for the duplex increased by DNA to RNA substitution, but the possibility of formation of an inactive form of the third strand was greatly reduced. Thus, i-motif formation is favored for cytosine-rich oligodeoxynucleotides, whereas triplex formation is favored for oligoribonucleotides. These properties make RNA pyrimidine oligonucleotides an attractive choice for triplex formation on a (C·G)-rich target, extending the practical sequence repertoire of pyrimidine triplexes to (C·G*C$^+$)-rich sequences.

Triple helices were first observed in 1957 for polyribonucleotides (Felsenfeld et al., 1957) and then for polydeoxynucleotides and hybrids (Riley et al., 1966; Morgan & Wells, 1968). These polymer interactions suggested an approach to double-stranded DNA recognition called oligonucleotide-directed triple helix formation. Intermolecular triple helix formation has been implicated as a possible means of controlling gene expression at the transcriptional level (Le Doan et al., 1987; Moser & Dervan, 1987; Maher et al., 1989; Maher, 1992; Thuong & Hélène, 1993). In the first class of triple helices, the pyrimidine motif, the third strand binds parallel to the purine strand of the duplex by Hoogsteen hydrogen bonds, forming T·A*T and C·G*C$^+$ triplets (de los Santos et al., 1989; Rajagopal & Feigon, 1989). The cytosines of the third strand must be protonated at the N3 position in order to establish a C·G*C$^+$ base triplet. The pK_a of the imino group of cytosine is well below 7, making these triplexes pH-dependent. Despite this handicap, DNA triple helices of the first motif have been reported at neutral pH. Several laboratories have investigated triple helices with cytosine analogs which do not require protonation (Ono et al., 1992; Krawczyk et al., 1992; Jetter & Hobbs, 1993).

The second motif involves T or A in the third strand bonding to A in the duplex and G in the third strand bonding to G in the duplex (Cooney et al., 1988; Beal & Dervan, 1991; Pilch et al., 1991). These triple helices often require a high divalent cation concentration, and the third strand, which is G-rich, is prone to form self-associated structures involving the formation of G tetrads (Sen & Gilbert, 1988). Such complexes reduce the ability of the triplex-forming oligonucleotide (TFO)[1] to bind to its intended target (Cheng & Van Dyke, 1993; Milligan et al., 1993; Olivas & Maher, 1995a). This tendency led to the design of guanine analogs unable to form G quartets but still having the potential to establish one or more hydrogen bonds with a C·G base pair (Olivas & Maher, 1995b; Rao et al., 1995). Purine oligonucleotides also have the potential to form another competing structure, a G·A parallel duplex, that can also interfere with triplex formation (Noonberg et al., 1995).

The possibility of an undesired, self-association of a pyrimidine third strand has been generally neglected, except in a few examples (Pilch et al., 1990; Pilch & Shafer, 1993). The formation of C·C$^+$ base pairs was originally reported 30 years ago in crystals of acetylcytosine (Marsh et al., 1962) and later in polydeoxy- or polyribonucleotides (Akinrimisi et al., 1963; Langridge & Rich, 1963; Fasman et al., 1964; Inman, 1964; Guschlbauer, 1967). The proposed hemiprotonated base pair is displayed in Figure 1. Solution studies of polydeoxycytidylic acid have pointed to the importance of hemiprotonation. Thus, the C·C$^+$ base pair and the C·G*C$^+$ base triplet both require protonation of one cytosine at the N3 position (Figure 1). Overall, the results showed that DNA sequences containing a significant stretch of cytosines could adopt complex pH-dependent conformations. Recently, the structure and stoichiometry of the complex formed by the hexamer d-T**CCCCC** has been described (Gehring et al., 1993; Leroy et al., 1993). The so-called *i-motif* is a tetramer of equivalent strands, which presents the novel feature of intercalated C·C$^+$ base pairs of two-parallel-stranded duplexes (Figure 1). Individual parallel-stranded duplexes are right-handed and underwound, and the two duplexes are "zipped together" in an antiparallel fashion.

* To whom correspondence should be addressed. Fax: (33-1) 40 79 37 05. E-mail: faucon@mnhn.fr.

[‡] CNRS UA481.

[§] URA D1254 CNRS.

[⊗] Abstract published in *Advance ACS Abstracts,* June 1, 1996.

[1] Abbreviations: TFO, triplex-forming oligonucleotide; MES, 2-(*N*-morpholino)ethanesulfonic acid.

S0006-2960(96)00107-9 CCC: $12.00

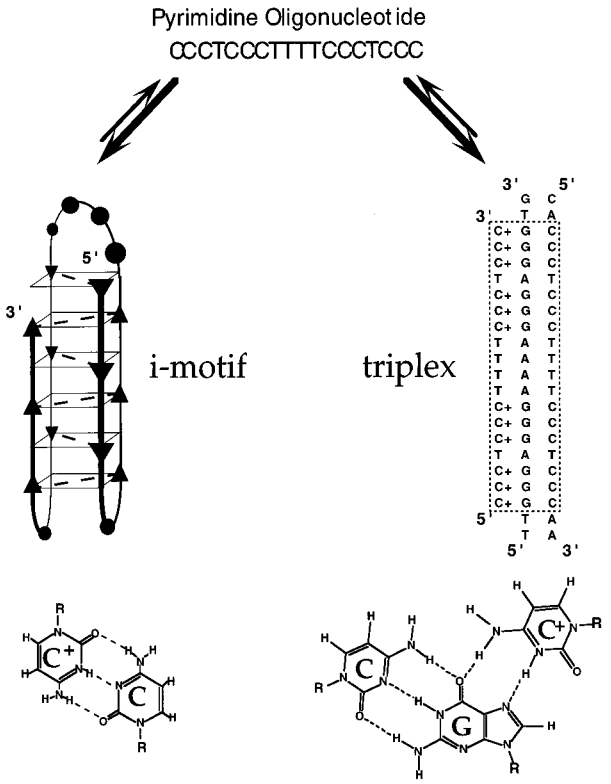

FIGURE 1: (Upper Part) i-Motif *vs* triplex formation. The intramolecular folding of the 18d oligonucleotide is presented on the left part (each thymine is symbolized by a circle and each cytosine by a triangle). i-Motif refers to an arrangement of intercalated C·C$^+$ base pairs (Gehring et al., 1993; Leroy et al., 1993). The triplex is presented on the right part (boxed region). The complete sequence of the purine strand of the duplex is 5′CATTCGTACGTTGGGAGGGAAAAGGGAGGGTGCAAGTAATAC3′, and the pyrimidine strand of the duplex is 5′GTATTACTTGCACCCTCCCTTTTCCCTCCCAACGTACGAATG3′. (Lower Part) The hemiprotonated C·C$^+$ base pair and the C·G*C$^+$ base triplet. Both the C·C$^+$ base pair and the C·G*C$^+$ base triplet involve protonation at the N3 position of one cytosine.

Crystallographic data on the d-**CCCC** and d-**CCC**T oligodeoxynucleotides confirmed this structure (Chen et al., 1994; Kang et al., 1994).

One of the problems of triple helix formation is that only oligopurine-oligopyrimidine sequences are recognized by the third strand. This sequence repertoire is even more restricted in practice; pyrimidine oligonucleotides were mainly targeted to T·A rich (≥50%) double-stranded targets. In the few articles dealing with C·G-rich targets (50−60%), the pyrimidine cytosine-rich third strand was usually short (9−12 bases) (Manzini et al., 1990; Pilch et al., 1990; Xodo et al., 1991; Völker et al., 1993; Hüsler & Klump, 1994,1995b). Adjacent C·G*C$^+$ base triplets were shown to be unstable, because of an electrostatic repulsion between the protonated cytosines (Kiessling & Dervan, 1992). Several laboratories have investigated cytosine analogs which do not require protonation in order to overcome this problem (Ono et al., 1992; Krawczyk et al., 1992; Jetter & Hobbs, 1993). In all other cases, the pyrimidine third strands contained a majority of thymines with a few interspaced cytosines, as a consequence of the Hoogsteen code (T recognizes a T·A base pair; C$^+$ recognizes a C·G base pair). In these conditions, intramolecular folding based on cytosine self-pairing is limited. A quick glance at the litterature shows that most reports of triplexes on G·C rich duplexes were obtained with G-rich oligodeoxynucleotides (the second motif). Besides the pH independence of these triplexes, and the electrostatic repulsion between adjacent C·G*C$^+$ triplets, was there any practical reason for this preference? Is it possible that the i-motif could interfere with triplex formation, by trapping the third strand in a stable folded conformation? To answer these questions, we started to investigate the factors that could affect the balance between self-association and triplex formation (Figure 1), especially the nature of the sugar backbone. Pyrimidine triplexes can be formed with an RNA third strand instead of a DNA third strand (Escudé et al., 1992, 1993; Roberts & Crothers, 1992; Han & Dervan, 1993; Wang & Kool, 1994; Mcdonald & Maher, 1995). In general, these triplexes are more stable than the corresponding triplexes with a DNA third strand. A further gain in stability can be obtained with 2′-*O*-methyl oligonucleotides (Escudé et al., 1992; Shimizu et al., 1992,1994). In this study, we have determined the role of the sugar backbone on the relative stabilities of the i-motif and the triple helix; RNA oligos and 2′-*O*-methyl oligoribonucleotides are attractive not only for their triplex-forming potential but also for their limited tendency to form undesired structures.

MATERIALS AND METHODS

Nomenclature. The terms stoichiometry, multimer, *n* mer, etc., refer to the association of several molecules and not to properties of a structure resulting from intramolecular folding. i-Motif refers to an arrangement of intercalated C·C$^+$ base pairs (Gehring et al., 1993; Leroy et al., 1993). We adopted the following convention for all the oligonucleotides. The last letter allows us to differentiate between DNA (d), RNA (r), and 2′-*O*-methyl (m) oligonucleotides with the same length and cytosine content.

Oligonucleotides. Unmodified oligodeoxynucleotides were synthesized by Eurogentec (Belgium) on the 0.2 μmol scale. Oligoribonucleotides and their 2′-*O*-methyl derivatives were synthesized by Eurogentec or Genset (France). They were suspended in 200 μL of bidistilled water, precipitated with 5 volumes of ethanol in the presence of 0.2 M sodium acetate, and washed with ethanol. Concentrations of all oligodeoxynucleotides were estimated by UV absorption at 30 °C, in a pH 8.0 buffer, using the sequence-dependent absorption coefficients computed from Cantor and Warshaw (1970). All concentrations were expressed in strand molarity.

UV Absorption Spectrophotometry. All experiments were performed as in Noonberg et al. (1995).

Gel Filtration. The size, and hence stoichiometry, of the *n* mers ($n \geq 1$) was determined by high-pressure gel filtration chromatography performed at room temperature with Beckman equipment. The column was a Synchropack GPC 100 (250 mm × 4.6 mm inside diameter) from Interchim, calibrated with oligodeoxynucleotides as previously described (Leroy et al., 1993). Oligonucleotides were eluted with a 20 mM sodium acetate, 300 mM NaCl buffer, at pH 4.7 or 4.2. Lower salt concentrations increased DNA−matrix interactions and decreased the resolution. The optimal flow rate was determined to be 0.4 mL/min, giving an elution time of 6 ± 2 min for a typical experiment.

Nondenaturing Gel Electrophoresis. Oligodeoxynucleotides were labeled with T4 polynucleotide kinase (New England Biolabs) and [^{32}P]ATP (ICN). The absence of smaller, contaminating species was first checked by 20%

polyacrylamide denaturating gel electrophoresis. Radiolabeled oligodeoxynucleotides (10 nM strand concentration) were incubated at 4 °C in the presence or in the absence of a large excess of an unlabeled oligodeoxynucleotide (2−200 μM strand concentration) in a pH 5.6, 50 mM MES buffer or in a pH 8.3, 50 mM Tris-borate EDTA (TBE) buffer. After a variable incubation time (from 0.5 to 100 h), samples were loaded on a nondenaturing 12% polyacrylamide gel. For triplex experiments, the pyrimidine strand of the duplex, 42 bases long with a 10 nM strand concentration, was ^{32}P-labeled, and preincubated for 1 h at room temperature with a 20% M excess of its complementary strand. A variable amount of TFO (18d, 18r, or 18m) was then added to reach a final third strand concentration of 0−30 μM. The triplex was incubated overnight in a MES buffer, with or without magnesium, and then loaded on a 12% nondenaturing polyacrylamide gel. Migration at 13 °C lasted 2−4 h (3 W) at 4 °C. The running buffer was recirculated during electrophoresis. Gels were dried and analyzed on a phosphorimager instrument (Molecular Dynamics). The migration of C-rich oligodeoxynucleotides was compared with that of control oligodeoxynucleotides (oligothymidylates T_{22}, T_{16}, T_{14}, or T_9).

Nuclear Magnetic Resonance. The NMR experiments were performed on a 360 MHz home-built spectrometer. The jump and return sequence with maximum sensitivity at 13.5 ppm was used for water suppression (Plateau & Guéron, 1982).

RESULTS

Comparison of Ribo- and Deoxyribonucleotides. Dissociation of the i-motif leads to a hyperchromism at 265 nm (Ahmed et al., 1994; Leroy et al., 1994; Manzini et al., 1994; Mergny et al., 1995). The 18d oligodeoxynucleotide (d-**CCC**T**CCC**TTTT**CCC**T**CCC**) had a T_m of 54 °C at pH 5.6. No signal was observed in a pH 7.8 cacodylate buffer, showing that this structure was pH-dependent. This denaturation profile was characterized by a sharp increase of absorbance at 265 nm, whereas an inverted transition was observed at 295 nm, due to cytosine protonation (Lavelle & Fresco, 1995; Mergny et al., 1995). Both curves (at 265 and 295 nm) could be analyzed as an all-or-none intramolecular phenomenon, and thermodynamic parameters determined from the two profiles were identical within experimental error. The fraction of folded 18 mer *vs* temperature is presented in Figure 2. Thermodynamic parameters are recapitulated on Table 1.

The melting profile of the unmodified 18d oligodeoxynucleotide was compared in the same experimental conditions with the profile of an oligoribonucleotide of identical length and corresponding sequence (with uracil replacing thymine, r-**CCC**U**CCC**UUUU**CCC**U**CCC**). As shown in Figure 2, the 18r oligo formed a less stable structure, as shown by a T_m of 25 °C instead of 54 °C. A decrease of absorbance at 295 nm was observed upon dissociation of the complex, suggesting that C·C$^+$ base pairs were involved in the folding. A 2′-*O*-methyl oligo of identical sequence (18m) melted at an even lower temperature, and folding was not complete even at 0 °C. The thermodynamical parameters determined from the denaturation profiles were also very different (Table 1), showing that the folding of the RNA oligoribonucleotide was much less enthalpy-driven.

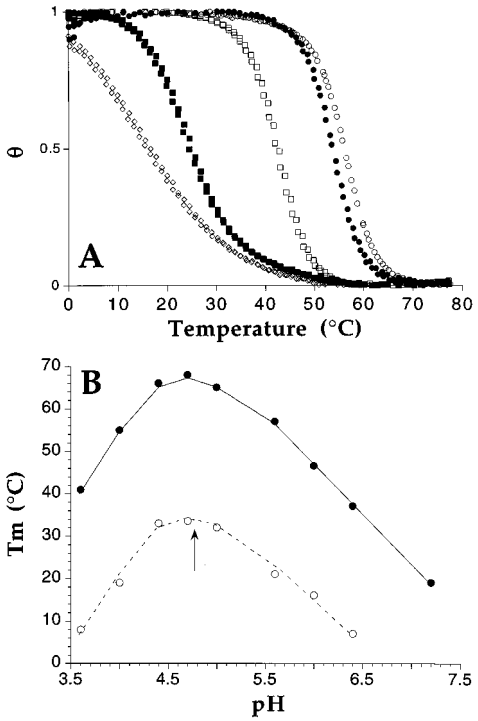

FIGURE 2: (A) Calculated fraction of the folded states determined form the denaturation profiles (T_m) obtained at pH 6.8 for 18 mer oligonucleotides at 295 nm: (●) 18d (d-**CCC**T**CCC**TTTT**CCC**T**CCC**), (○) 18dU (d-**CCC**U**CCC**UUUU**CCC**U**CCC**), (□) 18Chim (r-**CCC**d-T**CCC**TTTT**CCC**T**CCC**), (■) 18r (r-**CCC**U**CCC**UUUU**CCC**U**CCC**), (◇) 18m (2′-OMe-**CCC**U**CCC**UUUU**CCC**U**CCC**). (B) Effect of pH on T_m value for the 18d (●) and 18r (○) oligonucleotides in a 10 mM acetate or cacodylate buffer. The pK_a of cytosine is indicated by a vertical arrow.

Table 1: Thermodynamic Parameters for Intramolecular Folding of the 18 mers[a]

	T_m (°C)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (cal mol^{-1} K^{-1})	$\Delta G°$ (37 °C, kcal/mol)
18d	54	−68.2	−208	−2.67
18dU	56	−73.3	−223	−2.16
18ch	43	−57.7	−183	−1.01
18r	25	−32.5	−109	+1.18
18m	16	nd	nd	+1.56

[a] T_m values are given at ±1 °C, $\Delta G°$ values at ±0.1 kcal/mol, and $\Delta H°$ values at ±3 kcal/mol. (nd) The denaturation profile of the 18m oligo did not allow a precise determination of these parameters. The results are the average of three independent experiments. These values were calculated from the denaturation profiles at 265 nm (see Figure 2A and Materials and Methods).

To determine if the decrease of stability for the RNA oligoribonucleotide was the result of T to U substitution, we synthesized an 18 mer oligodeoxynucleotide containing dU instead of dT. This 18 mer (18dU, d-**CCC**U**CCC**UUUU**CCC**U**CCC**) had a T_m of 56 °C in the same experimental conditions. Thus, replacement of thymine by uracil did not decrease the stability of the folded oligodeoxynucleotide but even stabilized the i-motif, as shown by a small but significant increase in T_m ($\Delta T = +2$ °C). Therefore, the difference between 18r and 18d with respect to i-motif formation is not the result of T to U substitution and must arise from the presence of a 2′-hydroxyl group on RNA.

To confirm this result, we studied a chimeric oligonucleotide containing RNA and DNA moieties. The first three nucleotides (5′ to 3′ direction) contained a 2′-hydroxyl (RNA), whereas the other 15 nucleotides were oligodeoxy-

ribonucleotides. Its sequence was r(**CCC**)d(T**CCC**TTTT**C-CC**T**CCC**). Thus, intramolecular folding of this oligo would lead to an i-motif with three DNA and one RNA backbones. As shown in Figure 2A, this oligo had an intermediate stability between those of 18r and 18d (T_m = 43 °C), showing that the substitution of only three deoxyriboses by riboses had a significant destabilizing effect.

Folding of the 18r and 18d oligos involved the formation of hemiprotonated C·C$^+$ base pairs, as shown by the variation of absorbance at 295 nm. To exclude the possibility that the difference in thermal stability was the result of a change in pK_a, we recorded the absorption spectra of the 18d and 18r oligos at different pHs and at a high temperature (75 °C) where all structures were unfolded (not shown). From absorbance spectra, the fraction of protonated cytosines could be calculated at each pH. No significant difference was seen between the 18d and 18r oligonucleotides. Thus, the difference of thermal stability at pH 6 is not the result of a difference in pK_a. pK_a values reported in the litterature for polyribo- and polydeoxyribonucleotides are in good agreement with this observation.

Another argument in favor of the formation of C·C$^+$ base pairs is the pH dependence of the complexes. As shown in Figure 2B, the melting temperature of the 18d oligodeoxynucleotide was maximum at a pH value close to the pK_a of cytosine [4.8 in a low-salt buffer (Mergny et al., 1995)]. In the 5.2−7.2 pH range, the T_m of the structure was a linear function of pH, as previously observed (Leroy et al., 1994; Mergny et al., 1995). An increase in pH of 1 unit led to a decrease in the melting temperature of 20 °C or more. Such a pH dependence was previously observed for cytosine-rich polyribonucleotides (Akinrimisi et al., 1963), polydeoxyribonucleotides (Inman, 1964), as well as oligodeoxynucleotides (Leroy et al., 1994). The T_m *vs* pH curve for the 18r oligoribonucleotide was comparable in shape to the one obtained with 18d, but shifted to much lower temperatures (Figure 2B). The ΔT_m between 18d and 18r was almost independent of pH and never less than 28 °C. The optimal pH was in both cases 4.8, close to the pK_a of cytosines, in good agreement with the formation of hemiprotonated base pairs. Again, the maxima of the T_m *vs* pH curve argue against a large change of cytosine pK_a between DNA and RNA.

The pH value also had an impact on the kinetics of intramolecular folding. At neutral pH (6.8−7.2), 18d did not melt in a reversible fashion. Such a phenomenom could not be observed for RNA oligoribonucleotides, as no transition was obtained near neutral pH.

Evidence for Intramolecular Folding. The 18m, 18d, and 18r oligos showed a concentration-independent UV denaturation profile, indicative of intramolecular pairing. Gel filtration experiments also confirmed the stoichiometry of the complexes. In the 0.1−10 μM strand concentration range, the 18d, 18r, and 18m oligos were eluted as a single peak on a gel exclusion chromatography column, in agreement with a size of 20 ± 3 nucleotides, only compatible with a monomer species. The electrophoretic mobility of 18d and 18r oligonucleotides was also compared with that of oligodeoxythymidylates on a nondenaturing gel. At basic pH (8.3) and high temperature (45 °C), no anomaly of migration was observed (not shown). 18d comigrated to a position close to that of the T_{18} control oligodeoxynucleotide.

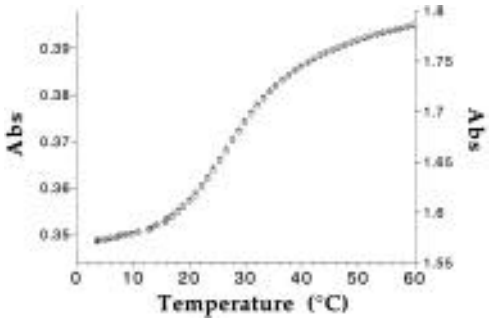

FIGURE 3: Concentration independence of the melting profile at 265 nm of the 21r oligoribonucleotide (r-**CCC**UAA**CCC**UAA**C-CC**UAA**CCC**) in a pH 5.6, 10 mM cacodylate buffer: (+) 21r, 2 μM strand concentration (absorbance shown on left scale); and (○) 21r, 9 μM strand concentration (absorbance shown on right scale).

Oligoribonucleotides had a slower migration than the corresponding oligodeoxynucleotides. This is in clear contrast with the abnormally fast, concentration-independent migration observed at pH 5.6 (Mergny et al., 1995). The 18r oligoribonucleotide was slightly retarded compared to the 18d oligodeoxynucleotide. The migration of the retarded band was not altered by the addition of unlabeled 18r, even in a large excess (up to 30 μM).

Extension to Other Cytosine-Rich Sequences. Is the large difference of stability between RNA and DNA limited to a specific sequence or a more general phenomenon? To answer this question, we analyzed two other classes of cytosine-rich sequences: (i) d-**CCC**TAA**CCC**TAA**CCC**-TAA**CCC** (21d) and its corresponding RNA r-**CCC**UAA**C-CC**UAA**CCC**UAA**CCC** (21r) and (ii) d-T**CC**T**CC**TTTT**C-CT**C**C**T (16d) with the corresponding RNA r-U**CC**U**CC**-UUUU**CC**U**CC**U (16r). Formation of the i-motif has already been demonstrated for the first DNA sequence, which corresponds to the C-rich strand of vertebrate telomers (Ahmed et al., 1994; Leroy et al., 1994; Kang et al., 1995). In both cases, the RNA strand had a significantly lower T_m. The T_m values for 21d and 21r were 49 and 28 °C, respectively; the T_m values for 16d and 16r were 38 and 24 °C, respectively, at pH 5.6. Thus, the difference of T_m between the DNA and RNA oligonucleotides was 29 °C for the 18d/18r pair, 21 °C for the 21d/21r pair, and 14 °C for the 16d/16r pair. Depending on the sequence, the difference of stability (ΔT_m) was not constant, but always largely in favor of the DNA oligonucleotide. Intramolecular folding was confirmed by the concentration independence of the melting profile (Figure 3 for the 21r oligoribonucleotide). At pH 6.8, the difference between DNA and RNA was even more important, as no transition at all was observed for any RNA oligoribonucleotide. This is in contrast with the DNA oligonucleotides, which melted at 25 °C (18d), 21 °C (21d), and 15 °C (16d). All these results confirm that RNA oligoribonucleotides do not form C·C$^+$ base pairs, except in acidic conditions, far from physiological conditions. Thus, DNA oligodeoxynucleotides have a larger pH window for i-motif formation.

Nature of the Folded Form. Intramolecular folding of cytosine-rich DNA oligonucleotides is now well-established. No structural data are available on the folding of the RNA counterpart. The NMR spectra of 18r and of 16r at −7 °C in H_2O at pH 4.2 exhibit broad NMR lines indicative of the presence of multiple species or of sample aggregation.

Sample heating above the melting temperature followed by fast cooling, a procedure which may improve the spectral quality by destroying undesired aggregation, left the spectrum unchanged.

As it might be expected in hemiprotonated C·C^+ pairs, the internal and external amino protons (9 and 8 ppm, respectively) are midway between the cytidine amino protons of neutral G·C pairs (8.5 and 6.8 ppm) and the amino protons of the protonated cytidine of Hoogsteen C^+·G pairs (10 and 9 ppm). The observation that the chemical shift of amino protons is independent of pH in a range (pH 4.2−5.3) where protonation of the N3 position would shift by 1 ppm the amino proton of the free cytidine (Raszka, 1974) is also in favor of the formation of hemiprotonated C·C^+ pairs. The 18r sample exhibits several broad C imino proton peaks in the 17−14 ppm region. By comparison with the uridine imino proton peak, we find that the imino proton cluster integrates for only one proton per strand. The 16r oligomer shows an extremely broad imino proton peak. The imino proton spectra of these RNA samples are not incompatible with the existence of hemiprotonated C·C^+ pairs. It may indicate fast proton exchange from short-lived C·C^+ pairs. Due to the low stability of the RNA complexes, we could not confirm whether this form was the i-motif. Spectrophotometric data showed that the folding was indeed unstable, pH-dependent (see Figure 2B), and intramolecular (Figure 3). The inverted profile at 295 nm strongly suggests that folding is accompained by partial cytosine protonation, in good agreement with the formation of hemiprotonated cytosine base pairs. However, this could not be taken as proof of RNA i-motif formation, as other types of structure could also involve C·C^+ base pairs.

To get a better understanding of the RNA folded form, we decided to investigate a shorter RNA fragment, i.e. r-U**CCCCC**. This fragment contains only one stretch of five cytosines; thus, four fragments are required to form an i-DNA motif. It was previously shown by gel filtration chromatography that d-T**CCCCC**, the DNA counterpart of r-U**CCCCC**, forms a tetramer (Gehring et al., 1993). The stoichiometry of the r-U**CCCCC** sample was also determined by high-pressure gel filtration chromatography in the same experimental conditions. We observed two components on the chromatography profile at 0 °C. One was eluted as a monomer (7.5 ± 2.5 residues) and the other as a molecule twice as large (15.6 ± 4 residues). We measured the monomer and dimer concentrations on the chromatography profiles of a series of samples prepared at different concentrations. The plot displayed in Figure 4 shows that the fraction of dimer varied as the square of the monomer concentration. The line of slope 2 drawn through the data points corresponds to a dissociation constant of 1.5×10^{-2} M. At room temperature, the chromatography profile of r-U**CCCCC** exhibits only one component which is eluted at a position depending on the sample concentration. This indicates that the monomer−dimer interconversion process is faster than the time required by the chromatography, i.e. 7 min.

i-Motif vs Triple-Helix. Pyrimidine oligodeoxynucleotides and oligoribonucleotides have been shown to bind with a high affinity to their double-stranded oligopyrimidine-oligopurine target. This property was mainly demonstrated in the case of T·A*T (T·A*U)-rich triple helices (C·G content of the DNA target below 50%). Is it still true for a cytosine-rich oligonucleotide, bound to a C·G rich target? The sequence chosen for triplex formation is presented in Figure 1. Binding of a third strand would lead to a triplex composed of 12 C·G·C^+ and 6 T·A*T (or T·A*U) triplets.

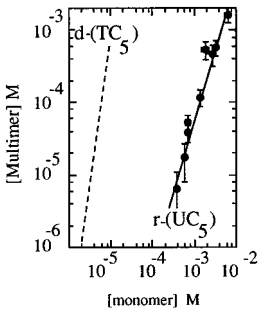

FIGURE 4: Characterization of the stoichiometry of the r-U**CCCCC** multimer by gel filtration chromatography at 0 °C. The multimer and monomer concentrations were measured on the elution profile of the sample whose concentration ranged from 10^{-2} to 4×10^{-4} M. The full line with a slope of 2 corresponds to a dimer−monomer equilibrium with a dissociation constant $K_d = 1.5 \times 10^{-2}$ M. The solution conditions were as follows: 0.3 M NaCl, 20 mM sodium acetate, and pH 4.2. The titration of the d-T**CCCCC** tetramer−monomer equilibrium (dashed line with a slope of 4) is only shown for comparison. The dissociation constant of the (d-T**CCCCC**)$_4$ tetramer is $(2.3 \times 10^{-6}$ M$)^3$ (Gehring et al., 1993).

In a sample containing equall amounts (1:1 strand stoichiometry) of the target duplex with a pyrimidine oligodeoxynucleotide, a transition was obtained upon the temperature being raised (Figure 5). This transition was absent in the case of the duplex alone, which melted at 70 °C. This transition was also seen at 295 nm, where cytosine protonation/deprotonation can be followed (Mergny et al., 1995). The T_m values obtained at 265 and 295 nm were the same. However, a similar transition was obtained with the third strand alone, in the absence of any target duplex (Figure 5A). This control shows that the transition is the result of the folding of the third strand, rather than of triplex formation. This transition was pH-dependent but concentration-independent (not shown). The T_m was slightly reduced upon addition of 10 mM $MgCl_2$. All these results strongly argue against triplex formation; the thermal transition is the result of an intramolecular pH-dependent folding of the third strand itself and not of triplex formation.

On the other hand, the RNA third strand (18r) did not give rise to a thermal transition when studied alone (Figure 5B). Upon addition of the duplex, a transition was observed; this melting curve was not reversible, as the cooling and heating curves could not be superimposed. This hysteresis phenomenon is the result of well-known slow kinetics of association and dissociation for triplexes (Rougée et al., 1992). In this case, $MgCl_2$ stabilized the complex ($\Delta T_m = +3$ °C at 10 mM $MgCl_2$).

Triplex formation was confirmed by a gel retardation assay. No retarded band was obtained when the 18d oligodeoxynucleotide was added to the duplex (0−10 μM strand concentration, Figure 6A). This was in clear contrast with the retarded band obtained in presence of increasing amounts of 18m (Figure 6B). Apparent K_d values of 0.06 and 0.1 μM were determined for the 18m and 18r oligoribonucleotides, respectively. Again, the gel retardation experiment showed that triplex formation is possible with a cytosine-rich RNA third strand but not with a DNA oligonucleotide; in none of the experimental conditions tested (pH

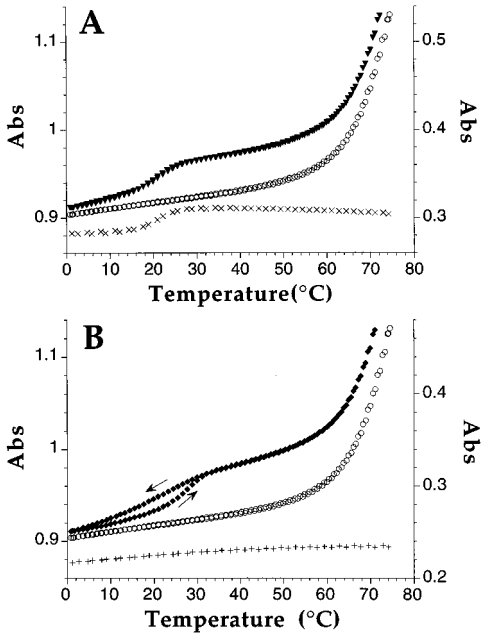

FIGURE 5: Triple helix formation *vs* i-motif folding. Heating and cooling profiles of the 18d and 18r oligonucleotides alone or in the presence of a 42-base pair duplex, in a pH 6.4, 0.14 M KCl, 10 mM $MgCl_2$ cacodylate buffer. (A) (○) Duplex alone, 1.5 μM strand concentration (absorbance on left scale), (▼) duplex (1.5 μM) + 18d oligo (1.8 μM) (absorbance on left scale with a 0.24 unit offset), and (×): 18d oligo alone (1.8 μM; absorbance on right scale). Each curve is the result of the superimposition of the cooling and heating profiles, showing that all phenomenons are reversible. (B) (○) Duplex alone, 1.5 μM strand concentration (absorbance on left scale), (◆) duplex (1.5 μM) + 18r oligo (1.8 μM) (absorbance on left scale with a 0.33 unit offset), and (+) 18r oligo alone (1.8 μM; absorbance on right scale). A hysteresis is obtained with the sample containing the duplex mixed with the RNA third strand; the heating profile is shifted toward higher temperatures than the cooling profile.

5−7, 0−20 mM $MgCl_2$) was any triplex seen with the 18d oligonucleotide.

DISCUSSION

DNA, RNA, and the i-Motif. In this report, we have observed a considerable difference between RNA and DNA oligonucleotides with respect to self-association. Cytosine-rich oligoribonucleotides were unable to form a self-associated structure above pH 6. The difference of stability was shown by a difference in T_m of 29 °C for the 18d/18r pair. It is interesting to note that a comparison of the data accumulated three decades ago on polyribonucleotides (Akinrimisi et al., 1963; Fasman et al., 1964; Hartman & Rich, 1965; Guschlbauer, 1967) and polydeoxynucleotides (Inman, 1964) supported the idea that the folded form of poly[r(C)] was less stable than the one obtained with poly[d(C)]. 2′-*O*-methyl oligoribonucleotides were even more resistant to self-association. What could explain this difference? First, the difference of stability between RNA and DNA cannot be explained by a difference of cytosine pK_a; no large change was observed with riboses instead of deoxyriboses in the backbone. Two chemical features distinguish DNA from RNA: a 2′-hydroxyl group on the sugar moiety and a 5-methyl group on thymine which is absent on uracil. We have shown that the substitution of dT to dU does not destabilize the i-motif. A slight stabilization is observed with uracil-subsituted oligodeoxynucleotides, showing that the methyl group of thymine does not play a role in the stabilization of the motif. Thus, the difference between RNA and DNA must arise from the presence/absence of a 2′-OH group. The structure of an i-DNA has been solved by NMR and crystallography for a fairly large number of sequences: d-T**CC** (Leroy & Guéron, 1995), d-^{m}**CC**T (Leroy & Guéron, 1995), d-T**CCC** (Leroy et al., 1993), d-**CCC**T (Kang et al., 1994), d-**CCC**TAA (Kang et al., 1995), d-**CCC**AAT (Berger et al., 1995), d-**CCCC** (Chen et al., 1994), and d-T**CCCCC** (Gehring et al., 1993). The three-dimensional structure of such motifs does not reveal any potential steric clash around the 2′ position of the sugar moiety. The destabilization induced by RNA is not the result of an obvious steric hindrance. More subtle effects must account for the large difference of stability, and we can only speculate on the reasons of this destabilization. The 2′-OH group plays an important role in the sugar conformation of an RNA strand in a duplex and induces a preferred N-type sugar pucker but may also be involved in direct or water-mediated intrastrand and interstrand interactions.

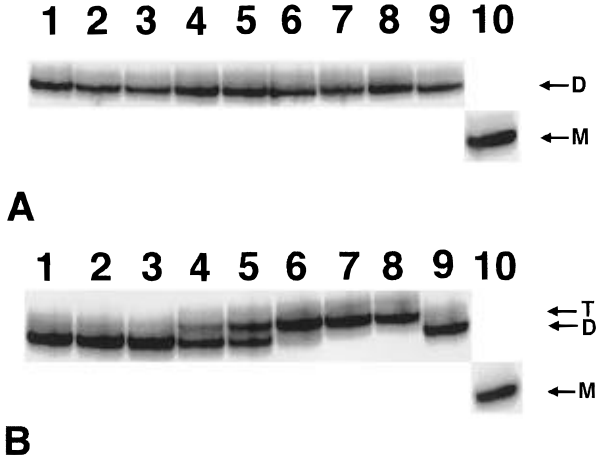

FIGURE 6: Gel retardation assay. The pyrimidine strand of the duplex (5′GTATTACTTGCACCCTCCCTTTTCCCTCCCAACG-TACGAATG3′) is radioactively labeled. The relative positions of the single strand (M), duplex (D), and triplex (T) species are indicated. The incubation and running buffers were identical: 50 mM MES, 100 mM NaCl, and 10 mM $MgCl_2$ at pH 5.6. (A) Lanes 1 and 9, duplex alone; lanes 2−8, increasing concentrations of the 18d oligo (0.1, 0.3, 0.6, 1, 3, 6, and 10 μM, respectively); and lane 10, labeled strand alone. (B) Lanes 1 and 9, duplex alone; lanes 2−8, increasing concentrations of the 18m oligo (0.01, 0.03, 0.06, 0.1, 0.3, 0.6, and 1 μM, respectively); and lane 10, labeled strand alone.

We have no proof that the unstable folded structure of an RNA strand is indeed the i-motif. For RNA, we have shown that the folding is intramolecular, leads to a comparable hypochromism at 265 nm and hyperchromism at 295 nm as the DNA oligodeoxynucleotide, is pH-dependent, and involves hemiprotonated C·C^+ base pairs. These data, although compatible with i-motif formation, are certainly not proof of the existence of an RNA i-motif. The thermodynamical parameters for 18r and 18d are very different (Table 1). The observation that RNA folding is much less enthalpy-driven might indicate a different structure. More strongly against an RNA i-motif is the preliminary observation that the NMR spectrum of r-U**CCCCC** is very different from that of d-T**CCCCC** (not shown). Moreover, as shown by gel filtration experiments, the structured form of r-U**CCCCC** involves only two strands. This is in clear constrast with

the four-strand stoichiometry expected for i-RNA formation. Thus, at least for this short RNA oligoribonucleotide, the structured form is very unlikely to be the i-motif, but rather a duplex. We did not investigate in further detail the nature of the folded structure for a large set of RNA molecules as, contrary to DNA which still can form the i-motif at pH 7, the i structure is not observed for RNA at physiological pH and is thus very unlikely to occur *in vivo*.

Protonation of the Cytosines and Biological Significance. There are some similarities between the C·G*C^+ base triplet, the C·C^+ base pair, and the G·C^+ base pair of a Hoogsteen duplex; all of them require protonated cytosines at the N3 position. As the pK_a of this group is low (between 4.2 and 4.8, depending on temperature and ionic strength), formation of these structures at neutral pH is accompanied by a protonation of the cytosines (Lavelle & Fresco, 1995), and the stability of these complexes is pH-dependent (Hüsler & Klump, 1995a). The stability of C·C^+ base pairs is indeed maximal at pH = pK_a, when half of the cytosines on an unstructured oligodeoxynucleotide are protonated. At pH 7, where a negligible fraction of the cytosines are protonated in the single-stranded structure, one proton has to be taken up for the formation of each hemiprotonated base pair or each C·G*C^+ base triplet (Lavelle & Fresco, 1995). This base pair or base triplet is, of course, less stable at pH 7 than at pH 5. Still, many oligodeoxynucleotides showed a thermal transition at neutral pH (Mergny et al., 1995). The fact that the DNA i-motif might be formed at neutral pH enhances the possibility of it being found under physiological conditions, especially if some proteins specifically bind to this structure and stabilize it *in vivo*. On the other hand, such folding is not likely to play any role in the secondary structure of RNA. Many RNA sequences found in Genbank were compatible, in theory, with intramolecular i-motif formation. In some cases, such as the Xist RNA, long cytosine-rich regions are present. This study shows unambiguously that such a putative RNA i-motif does not exist under physiological conditions. But the *absence* of such a structure is by itself an interesting observation for the antisense and antigene strategies.

Interference of Self-Association with Oligonucleotide-Based Strategies. We have shown that the folding of the third strand can interfere with triplex formation. To design a triplex-forming oligodeoxynucleotide, the choice of the target sequence is of course the first constraint, as triplex formation is mostly limited to oligopurine-oligopyrimidine stretches. Any sequence containing at least four repeats of two cytosines (or more) is compatible with i-motif formation. Many of the potential binding sites for triplex-forming oligodeoxynucleotides belong to this family; the promoter regions of important genes contain C·G-rich repeating sequence motifs (Firulli et al., 1994; Manzini et al., 1994). These sequences will be deleterious to triplex formation with DNA because the third strand is likely to self-associate and thus to be blocked in an inactive conformation.

A pyrimidine oligodeoxynucleotide, 18 bases long, that contained 12 cytosines and 6 thymines was not able to bind to its double-stranded target, even at acidic pH and in the presence of magnesium ions. This low affinity might be the result of the presence of several rows of destabilized adjacent C·G*C^+ base triplets in the triplex (Kiessling & Dervan, 1992) but also a consequence of self-association of the third strand. i-Motif formation was preferred over triplex formation in all the experimental conditions tested. A change of pH could not favor triplex over i-DNA. Both structures have the same qualitative pH dependence above cytosine pK_a. The fact that only half of the cytosines must be protonated in the i-motif [*vs* all in the third strand of a triplex (Lavelle & Fresco, 1995)] would even suggest that an increase of pH above pK_a would lead to a moderate destabilization of the i-motif, compared to the triplex. One would expect that a very low pH (4 or less), below cytosine pK_a, might shift the equilibrium toward triplex formation, as the i-motif is itself destabilized under strong acidic conditions (Figure 2B). The required physicochemical conditions are so remote from the physiological ones that we did not investigate such a possibility. The addition of 10−20 mM magnesium ions, known to favor triplex formation, was insufficient to promote triplex formation in the 18d case. The third strand was trapped in a folded configuration which is not compatible with triplex formation. To confirm that this preference was not limited to very specific sequences, highly favorable to i-motif formation, we performed a competition experiment on a shorter triplex, with a reduced number of C·G*C^+ triplets. The sequence of the third strand was d-T**CC**T**C**-**C**TTTT**CC**T**CC**T (16d). Again, this oligonucleotide was trapped in a folded structure in a large number of experimental conditions (Mergny et al., 1995), and triplex formation was only possible at a high magnesium concentration (20 mM) and at a high dication/monocation ratio (not shown). The sequence repertoire for the triplex strategy using oligodeoxynucleotides is thus even more restricted than generally admitted. Fortunately, this limitation can be overcome by using RNA and 2′-*O*-methyl oligoribonucleotides, which were able to bind to the C·G-rich target with submicromolar dissociation constants at slightly acidic pH.

It should be noted that this DNA to RNA change in the third strand chemistry, which is doubly advantageous in the pyrimidine motif, is of little use in the purine motif. As a matter of fact, G quartets are effectively formed with RNA strands (Cheong & Moore, 1992), and thus, guanine-rich RNA third strands could also be trapped in tetraplex structures. But the main reason why oligoribonucleotides are of little interest for the purine triplexes is that they do not form stable triple helices (Semerad & Maher, 1994). Formation of undesired structures should be taken into account not only for triplex formation but also in antisense applications. For example, ISIS 3466, which is complementary to gp120 messenger RNA (Perlaky et al., 1993), has a sequence compatible with i-motif formation, and we have shown that such oligonucleotides, whether phosphodiester or phosphorothioate, effectively self-associate *in vitro* (unpublished results).

Many research groups are designing oligonucleotide analogs to improve the efficiency of such molecules as antisense or antigene agents. Such chemical modifications could also have an impact on i-motif stability or more generally on any type of competing self-structure. In the present study, we investigated substitutions at the 2′ position of the sugar moiety. RNA and 2′-*O*-methyl oligos do not form a stable i-motif and are promising agents to form triple helices on C·G-rich DNA targets, at least in slightly acidic conditions. Thus, the choice of an oligodeoxynucleotide for triplex formation must take into account not only its ability to form a triplex but also its ability to form an undesired, competing structure such as G quartets (Olivas & Maher,

1995a), GA-parallel duplexes (Noonberg et al., 1995), or i-DNA. The observation that RNA oligonucleotides can form stable triple helices opens the possibility of using DNA vectors to synthesize RNA transcripts within cells.

REFERENCES

Ahmed, S., Kintanar, A., & Henderson, E. (1994) *Nat. Struct. Biol. 1*: 83−88.

Akinrimisi, E. O., Sander, C., & Tso, P. O. P. (1963) *Biochemistry 2*, 340−344.

Beal, P. A., & Dervan, P. B. (1991) *Science 251*, 1360−1363.

Berger, I., Kang, C., Fredian, A., Ratliff, R., Moyzis, R., & Rich, A. (1995) *Nat. Struct. Biol. 2*, 416−425.

Brown, D. M., Gray, D. M., & Patrick, M. H. (1985) *Biochemistry 24*, 1676−1683.

Cantor, C. R., & Warshaw, M. M. (1970) *Biopolymers 9*, 1059−1077.

Chen, L., Cai, L., Zhang, X. H., & Rich, A. (1994) *Biochemistry 33*, 13540−13546.

Cheng, A.-J., & Van Dyke, M. W. (1993) *Nucleic Acids Res. 21*, 5630−5635.

Cheong, C. J., & Moore, P. B. (1992) *Biochemistry 31*, 8406−8414.

Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J., & Hogan, M. E. (1988) *Science 241*, 456−459.

de los Santos, C., Rosen, M., & Patel, D. (1989) *Biochemistry 28*, 7282−7289.

Escudé, C., Sun, J. S., Rougée, M., Garestier, T., & Hélène, C. (1992) *C. R. Acad. Sci., Ser. III 315*. 521−525.

Escudé, C., François, J. C., Sun, J. S., Ott, G., Sprinzl, M., Garestier, T., & Hélène, C. (1993) *Nucleic Acids Res. 21*, 5547−5553.

Fasman, G. D., Lindblow, C., & Grossman, L. (1964) *Biochemistry 3*, 1015−1021.

Felsenfeld, G., Davies, D. R., & Rich, A. (1957) *J. Am. Chem. Soc. 79*, 2023.

Firulli, A. B., Maibenco, D. C., & Kinniburgh, A. J. (1994) *Arch. Biochem. Biophys. 310*, 236−242.

Gehring, K., Leroy, J. L., & Guéron, M. (1993) *Nature 363*, 561−565.

Guschlbauer, W. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 1441−1448.

Han, H., & Dervan, P. B. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3806−3810.

Hartman, K. A., & Rich, A. (1965) *J. Am. Chem. Soc. 87*, 2033.

Hüsler, P. L., & Klump, H. H. (1994) *Arch. Biochem. Biophys. 313*, 29−38.

Hüsler, P. L., & Klump, H. H. (1995a) *Arch. Biochem. Biophys. 317*, 46−56.

Hüsler, P. L., & Klump, H. H. (1995b) *Arch. Biochem. Biophys. 322*, 149−166.

Inman, R. B. (1964) *J. Mol. Biol. 9*, 624−637.

Jetter, M. C., & Hobbs, F. W. (1993) *Biochemistry 32*, 3249−3254.

Kang, C. H., Berger, I., Lockshin, C., Radliff, R., Moyzis, R., & Rich, A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11636−11640.

Kang, C. H., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R., & Rich, A. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 3874−3878.

Kiessling, L. L., & Dervan, P. B. (1992) *Biochemistry 31*, 2829−2834.

Krawczyk, S. H., Milligan, J. F., Wadwani, S., Moulds, C., Froehler, B. C., & Matteucci, M. D. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 3761−3764.

Langridge, R., & Rich, A. (1963) *Nature 198*, 725−728.

Lavelle, L., & Fresco, J. R. (1995) *Nucleic Acids Res. 23*, 2692−2705.

Le Doan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J.-L., Thuong, N. T., Lhomme, J., & Hélène, C. (1987) *Nucleic Acids Res. 15*, 7749−7760.

Leroy, J. L., & Guéron, M. (1995) *Structure 3*, 101−120.

Leroy, J. L., Gehring, K., Kettani, A., & Guéron, M. (1993) *Biochemistry 32*, 6019−6031.

Leroy, J. L., Guéron, M., Mergny, J. L., & Hélène, C. (1994) *Nucleic Acids Res. 22*, 1600−1606.

Maher, L. J., III (1992) *Bioessays 14*, 807−815.

Manzini, G., Xodo, L. E., Gasparotto, D., Quadrifoglio, F., van der Marel, G. A., & van Boom, J. H. (1990) *J. Mol. Biol. 213*, 833−843.

Manzini, G., Yathindra, N., & Xodo, L. E. (1994) *Nucleic Acids Res. 22*, 4634−4640.

Marsh, R. E., Bierstedt, R., & Eichhorn, E. L. (1962) *Acta Crystallogr. 15*, 310−316.

Mcdonald, C. D., & Maher, L. J., III (1995) *Nucleic Acids Res. 23*, 500−506.

Mergny, J. L., Lacroix, L., Han, X., Leroy, J. L., & Hélène, C. (1995) *J. Am. Chem. Soc. 117*, 8887−8898.

Milligan, J. F., Krawczyk, S. H., Wadwani, S., & Matteucci, M. D. (1993) *Nucleic Acids Res. 21*, 327−333.

Morgan, A. R., & Wells, R. D. (1968) *J. Mol. Biol. 37*, 63−80.

Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645−650.

Noonberg, S. B., François, J. C., Garestier, T., & Hélène, C. (1995) *Nucleic Acids Res. 23*, 1956−1963.

Olivas, W. M., & Maher, L. J., III (1995a) *Biochemistry 34*, 278−284.

Olivas, W. M., & Maher, L. J., III (1995b) *Nucleic Acids Res. 23*, 1936−1941.

Ono, A., Ts'O, P. O. P., & Kan, L. S. (1992) *J. Org. Chem. 57*, 3225−3230.

Perlaky, L., Saijo, Y., Busch, R. K., Bennett, C. F., Mirabelli, C. K., Crooke, S. T., & Busch, H. (1993) *Anti-Cancer Drug Des. 8*, 3−14.

Pilch, D. S., & Shafer, R. H. (1993) *J. Am. Chem. Soc. 115*, 2565−2571.

Pilch, D. S., Brousseau, R., & Shafer, R. H. (1990) *Nucleic Acids Res. 18*, 5743−5750.

Pilch, D. S., Levenson, C., & Shafer, R. H. (1991) *Biochemistry 30*, 6081−6087.

Plateau, P., & Guéron, M. (1982) *J. Am. Chem. Soc. 104*, 7310−7311.

Rajagopal, P., & Feigon, J. (1989) *Nature 339*, 637−640.

Rao, T. S., Durland, R. H., Seth, D. M., Myrick, M. A., Bodepudi, V., & Revankar, G. R. (1995) *Biochemistry 34*, 765−772.

Raszka, M. (1974) *Biochemistry 13*, 4616−4622.

Riley, M., Maling, B., & Chamberlin, M. J. (1966) *J. Mol. Biol. 20*, 359−389.

Roberts, R. W., & Crothers, D. M. (1992) *Science 258*, 1463−1466.

Rougée, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannangéli, C., Garestier, T., & Hélène, C. (1992) *Biochemistry 31*, 9269−9278.

Semerad, C. L., & Maher, L. J., III (1994) *Nucleic Acids Res. 22*, 5321−5325.

Sen, D., & Gilbert, W. (1988) *Nature 334*, 364−366.

Shimizu, M., Konishi, A., Shimada, Y., Inoue, H., & Ohtsuka, E. (1992) *FEBS Lett. 302*, 155−158.

Shimizu, M., Koizumi, T., Inoue, H., & Ohtsuka, E. (1994) *Bioorg. Med. Chem. Lett. 4*, 1029−1032.

Thuong, N. T., & Hélène, C. (1993) *Angew. Chem., (Int. Ed. Engl.) 32*, 666−690.

Völker, J., Botes, D. P., Lindsey, G. G., & Klump, H. H. (1993) *J. Mol. Biol. 230*, 1278−1290.

Wang, S. H., & Kool, E. T. (1994) *Nucleic Acids Res. 22*, 2326−2333.

Xodo, L. E., Manzini, G., Quadrifoglio, F., van der Marel, G. A., & van Boom, J. H. (1991) *Nucleic Acids Res. 19*, 5625−5631.

BI960107S