Hoogsteen DNA Duplexes of 3′−3′- and 5′−5′-Linked Oligonucleotides and Triplex Formation with RNA and DNA Pyrimidine Single Strands: Experimental and Molecular Modeling Studies

Ekambar R. Kandimalla* and Sudhir Agrawal*

Hybridon, Inc., One Innovation Drive, Worcester, Massachusetts 01605

Received June 21, 1996; Revised Manuscript Received September 26, 1996[⊗]

ABSTRACT: DNA oligonucleotide sequences containing two parallel complementary strands attached through 3′−3′ and 5′−5′ linkages were synthesized. These oligonucleotides form Hoogsteen base-paired parallel-stranded (PS) hairpin duplexes under appropriate conditions [Kandimalla, E. R., Agrawal, S., Venkataraman, G., & Sasisekharan, V. (1995a) *J. Am. Chem. Soc. 117*, 6416−6417]. UV melting experiments show that these Hoogsteen hairpin duplexes have a lower thermal stability than that of the Watson−Crick (WC) hairpin duplex (antiparallel) of the same sequence. The circular dichroism (CD) spectrum of the Hoogsteen duplex is different from the canonical B-DNA WC duplex spectrum. The formation of the Hoogsteen duplex is pH-dependent since protonation of cytosine requires lower pH conditions. Studies with oligonucleotides of different loop sizes revel that three- and two-base loops are optimum for the formation of stable Hoogsteen duplexes with 3′−3′ and 5′−5′ linkages, respectively. The guanine residues in the loop stabilize the duplex as a result of G-G interactions as confirmed by molecular modeling studies. The new PS Hoogsteen duplexes form stable triplexes with complementary (antiparallel to the purine domain) single-stranded RNA and DNA pyrimidine sequences in Py·Pu:Py (pyrimidine third strand−purine WC strand:pyrimidine WC strand) motif. The thermal stability of the resulting triplexes is much higher than that of the conventional triplex (binding of a Hoogsteen pyrimidine third strand to a WC duplex) of the same sequence. The CD spectra of the new triplexes are similar to those of conventional triplexes, suggesting that no conformational change occurs as a result of 3′−3′ or 5′−5′ linkage. A molecular modeling study was carried out to examine the stereochemical feasibility of the Hoogsteen duplexes and formation of triplexes with single-stranded pyrimidine complementary strands.

DNA is highly polymorphic in nature and exists in at least three different families of A-, B-, and Z-DNA duplexes (Saenger, 1984). A common feature of these three types of duplexes is the antiparallel (AP)[1] orientation of the two constituent strands. Several other unusual and higher-order structures have been added to the ever-growing families of DNA in the last two decades. These new structures include triplexes (Cheng & Pettitt, 1992; Maher, 1992; Thoung & Helene, 1993, and references cited therein), tetraplexes (Blackburn, 1991), and reversed Watson−Crick (rWC) parallel-stranded (PS) duplexes (van de Sande et al., 1988). In addition, a structure involving a Hoogsteen hydrogen-bonding pattern is known to exist between the purine strand of a duplex DNA and a pyrimidine third strand of a triplex in the Py·Pu:Py motif (Cheng & Pettitt, 1992; Thoung & Helene, 1993, and references cited therein).

Until recently, scientists believed that Hoogsteen hydrogen bonding is possible only in triple helices (Wilkins et al., 1970) and monomeric crystals (Hoogsteen, 1963) or with modified bases containing bulky groups that prevent Watson−Crick (WC) base pairing (Ikehara et al., 1972; Ishikawa et al., 1972; Hakoshima et al., 1981). We have reported the existence of a pH-dependent Hoogsteen hydrogen-bonded PS duplex, adding a new structure to the DNA families (Liu et al., 1993; Kandimalla et al., 1995a). The WC duplex d(A-G)$_6$·d(C-T)$_6$ undergoes transition to the PS Hoogsteen duplex d(A-G)$_6$·d(C^+-T)$_6$ at a lower pH and is more stable than the corresponding WC duplex (Lavelle & Fresco, 1995). An intercalator, benzopyridoindole derivative, interacts specifically with the Hoogsteen PS duplex and stabilizes the Hoogsteen duplex structure (Escude et al., 1996).

Pyrimidine linear (Xodo et al., 1990; Giovannangeli et al., 1991; D'Souza & Kool, 1992; Kandimalla & Agrawal, 1993, 1994; Noll et al., 1994) or circular (Kool, 1991; Prakash & Kool, 1992; Wang & Kool, 1994) oligonucleotides form triplexes with complementary purine single strands (Figure 1B), in addition to the conventional way of forming triplexes by adding a pyrimidine (Roberts & Crothers, 1992; Han & Dervan, 1993; Escude et al., 1993) third strand to an already existing duplex (Figure 1A). These oligonucleotides are generally named foldback triplex-forming (FTF) and circular (C) oligonucleotides (Os). The triplexes formed by FTFOs or COs are more stable than conventional triplexes (Prakash & Kool, 1992; Kandimalla & Agrawal, 1994). Recently, we reported triple-helix formation of a pyrimidine RNA/DNA strand with a PS Hoogsteen hairpin duplex (Figure 1C) (Kandimalla et al., 1995a). The resulting triplex showed a greater thermal stability than the conventional triplexes (Figure 1A) (Kandimalla et al., 1995a).

* Authors to whom correspondence should be addressed. Phone: 508-752-7000. Fax: 508-751-7692.

[⊗] Abstract published in *Advance ACS Abstracts,* November 15, 1996.

[1] Abbreviations: AP, antiparallel; CD, circular dichroism; FTFO, foldback triplex-forming oligonucleotide; PS, parallel-stranded; rWC, reversed Watson−Crick; T_m, melting temperature; UV, ultraviolet; WC, Watson−Crick.

S0006-2960(96)01505-X CCC: $12.00

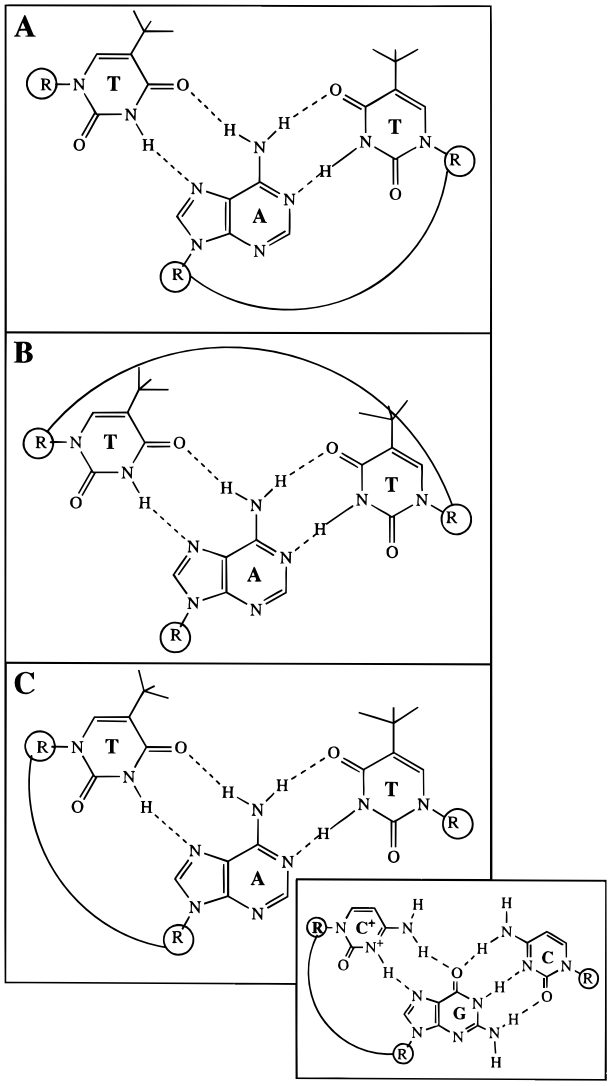

FIGURE 1: Three possible ways of strand attachment with hydrogen bonding showing for the T·A:T triplet. (A) Conventional triplex in which WC strands are covalently attached and the Hoogsteen strand is added to the hairpin WC duplex. (B) Foldback triplex in which WC and Hoogsteen pyrimidine strands are covalently attached and the purine strand is added. (C) Current model in which the Hoogsteen and WC purine strands are covalently attached and the WC pyrimidine strand is added. The C^+·G:C triplet is shown in the inset.

In the present study, we have examined in detail PS DNA hairpin duplexes with Hoogsteen hydrogen bonding independent of triplex structure. The new structures are obtained by synthesizing oligonucleotides containing two complementary strands, one of which is an all-purine strand and the other of which is an all-pyrimidine strand, attached by a 3′−3′ or 5′−5′ linkage (Kandimalla et al., 1995a) in order to facilitate parallel orientation of the two strands. Such covalent attachment of the two strands prevents formation of mismatched, slipped structures, which have been reported for linear strands designed to form PS structures (Liu et al., 1993). We show that the PS Hoogsteen duplexes readily form triplexes with a pyrimidine strand, which is complementary in antiparallel orientation to the purine strand of the PS Hoogsteen duplex. We use molecular modeling to generate stereochemically feasible structures for the PS Hoogsteen-bonded duplex. We also show that the guanines that are a part of the loop linking the two strands are capable of forming hydrogen bonds. Further, we systematically studied the length of the spacer loop required for both the 3′−3′ and the 5′−5′ linkages. The third pyrimidine strand in the triple helix is capable of fitting in the groove of the PS structure without altering the conformation of the PS duplex or the loop connecting the Hoogsteen strands.

EXPERIMENTAL PROCEDURES

Oligonucleotide Synthesis. Oligonucleotides were synthesized on a Milligen 8700 DNA synthesizer (Bedford, MA) using phosphoramidite chemistry. β-Cyanoethyl *N*,*N*-diisopropylphosphoramidites and 2′-(*tert*-butyldimethylsilyl)-3′-β-cyanoethyl *N*,*N*-diisopropylphosphoramidites were purchased from Millipore for DNA and RNA syntheses, respectively. The 3′-DMT-5′-β-cyanoethyl *N*,*N*-diisopropylphosphoramidites and 5′-monomer-attached CPG solid support were obtained from Glen Research Laboratories (Sterling, VA). The 5′−5′-attached oligonucleotide synthesis was carried out on a regular CPG solid support (3′-attached nucleoside) until the point of 5′−5′ attachment, and from there onward, synthesis was continued using 3′-DMT-5′-phosphoramidite monomers with longer detritylation and coupling times. The 3′−3′-linked oligonucleotide was synthesized on 5′-monomer-attached CPG using 3′-DMT-5′-phosphoramidite monomers with longer detritylation and coupling times until the point of 3′−3′ attachment, and from there onward, synthesis was continued with normal phosphoramidite monomers to the required length. After synthesis, all the oligonucleotides were deprotected, purified on reverse phase HPLC (C_{18}), detritylated, and desalted using C_{18} Sep-pack cartridges (Waters, Milford, MA). The purity of the oligonucleotides was checked by denaturing polyacrylamide gel electrophoresis (PAGE).

Oligoribonucleotide (RNA) was deprotected with a 3:1 mixture of ammonium hydroxide and ethanol at 55 °C for about 15 h and then with 1 M tetrabutylammonium fluoride at room temperature for another 15 h. RNA was then purified on 20% denaturing PAGE, eluted from the gel, and desalted using a C_{18} Sep-pack cartridge (Waters). Oligonucleotides were quantified by measuring absorbance at 260 nm using extinction coefficients, which were calculated by the nearest neighbor method (Puglisi & Tinoco, 1989).

UV Thermal Melting Experiments. UV thermal melting experiments were conducted in either 100 mM sodium acetate, pH 5.0 buffer containing 10 mM magnesium chloride or 100 mM sodium acetate, pH 7.6 buffer containing 10 mM magnesium chloride and 1 mM spermine (Kandimalla et al., 1995a). Thermal melting curves were recorded at 260 nm at a heating rate of 0.5 °C/min with a Perkin-Elmer Lambda 2 spectrophotometer. The oligonucleotide concentration was 1.2 μM per each strand. The midpoint of the thermal melting curves (T_m) was determined from first-derivative curves obtained by plotting dA_{260nm}/dT versus T (temperature). Each value is an average of two independent experiments. The uncertainty in the T_m values is about 1.0 °C.

Circular Dichroism (CD) Experiments. Oligonucleotides were mixed in 10 mM cacodylate (pH 5.0), 100 mM sodium chloride, and 10 mM magnesium chloride buffer. The concentration of each strand was 1.0 μM. The solutions were heated at 90 °C for 10 min, allowed to come to room temperature slowly, and stored at 4 °C until CD measurements were carried out. The CD spectra were recorded on a Jasco J-710 spectropolarimeter attached to a peltier thermal

controller using a 0.5 cm quartz cell. The samples were equilibrated at the required temperature for 15 min before recording the spectra. Each spectrum is an average of eight scans subtracted with the buffer blank, which is also an average of eight scans and obtained at the same scan speed (100 nm/min). All the spectra are noise-reduced using the software supplied by Jasco, Inc., and the molar ellipticities were calculated using the same software.

CD Mixing Curves. The oligonucleotide (pyrimidine target and the hairpin parallel duplex oligonucleotides) solutions were prepared in 88:12, 75:25, 63:37, 50:50, 37:63, 25:75, and 12:88 molar ratios in 10 mM cacodylate (pH 5.0), 100 mM sodium chloride, and 10 mM magnesium chloride buffer. Two samples containing only one of the two strands were prepared (100:0 or 0:100) in the same buffer. The final concentration of the oligonucleotide strands was 2 μM. The sample solutions were prepared as described above. The CD measurements were carried out as described above after equilibrating each sample at 20 °C for 15−20 min.

Molecular Modeling. Molecular modeling was performed on a Silicon Graphics Iris Indigo workstation using the InsightII program (version 2.3.1 Biosym-MSI, Burlington, MA). The coordinates of the purine strand were generated using the linked atom least-squares program (Smith & Arnott, 1978), and the helical parameters were the same as used earlier for Hoogsteen-paired duplexes ($n = 12$, $h = 3.26$) (Raghunathan et al., 1994). The coordinates of the pyrimidine strand of the Hoogsteen duplex were generated (Raghunathan et al., 1994) by rotating the coordinates of the purine strand about the helix axis by an angle of 69.5°. Energy minimizations were performed using the CVFF force field in Discover (Biosym-MSI) with 100 steps of steepest descent followed by 500 steps of the conjugate gradient method in a vacuum (dielectric of 1.0), primarily to relieve bad stereochemical contacts in the loop structure (Kandimalla et al., 1995c). Then the backbone torsions and the glycosidic linkage torsions were manually adjusted to facilitate the G−G hydrogen-bonding scheme (Saenger, 1984; Mohanty & Bansal, 1993, 1994), and the loop was energy-minimized to relieve any unfavorable contacts. The coordinates of the WC base-paired pyrimidine chain for the triple helix were generated as described earlier (Raghunathan et al., 1994).

RESULTS AND DISCUSSION

Design of Oligonucleotides. Hoogsteen base pairing takes place between a purine and a pyrimidine strand that run parallel to each other. Two hydrogen bonds are formed between A and T and two between G and C^+. The new oligonucleotides contain a 16-nucleotide pyrimidine domain and a complementary 16-nucleotide purine domain parallel to the pyrimidine domain. The two oligonucleotide sequences are attached through a 3′−3′ or a 5′−5′ linkage to facilitate the parallel orientation. Adjacent to the 3′−3′ or 5′−5′ linkage, a pentanucleotide purine loop is incorporated to function as a spacer. The oligonucleotides are designed to contain G-C bases in order to avoid formation of the rWC parallel duplex (Rentzeperis et al., 1992). Oligonucleotides used in the study are shown in Table 1. Oligonucleotides **1**, **2**, **7**, **3**+**5**, and **6**+**5** were characterized by native polyacrylamide gels (Kandimalla et al., 1995a).

Thermal Stability of the Hoogsteen Hairpin Duplexes. Thermal melting results are summarized in Table 2. Oligonucleotides **1** and **2** alone showed sharp, cooperative, and monophasic melting transitions (data not shown) in pH 5.0 buffer, indicating formation of stable secondary structures. The thermal stability of linear WC (**3**+**5**) and Hoogsteen (**6**+**5**) duplexes of the same sequences as that of oligonucleotides **1** and **2** was also studied (Table 2). The thermal stability of the hairpin duplexes of **1** and **2** is about 5.5 °C higher than that of the linear PS duplex of oligonucleotides **3**+**5** in pH 5.0 buffer. The WC hairpin duplex of oligonucleotide **7**, which has the same sequence as that of PS oligonucleotides **1** and **2**, shows a T_m of 67.4 °C. The unusually high T_m for the WC hairpin duplex could be due to stacking of guanines in the loop (Hirao et al., 1992) (see

Table 1: Oligonucleotide Sequences

Parallel-Stranded Hairpin Sequences

1 A G G TT T TT TCT TTC TTCC T)-5'
G
G-3'-3'-AAAAAAGAAAGAAGGA)d-5'

2 A G G T CC TTCT TT CTT TT TT)-3'
G
G-5'-5'-AGGAAGAAAGAAAAAA)d-3'

Complementary Pyrimidine and Purine Sequences

3 5'-d(TTT TTT CTT TCT TCC T)-3'

4 5'-UUU UUU CUU UCU UCC U-3'

5 5'-d(AGGAAGAAAGAAAAAA)-3'

Hoogsteen Pyrimidine Sequence

6 5'-d(TCC TTC TTT CTT TTT T)-3'

Antiparallel-Stranded Hairpin Sequence

G G TT T TT TCT TTC TTCC T)-3'
7 A
G G AAAAAAGAAAGAAGGA)d-5'

Table 2: Thermal Melting Data of Duplexes

oligo number	oligonucleotide sequences/complexes[a]	T_m, °C pH 5.0	T_m, °C pH 7.6[b]
1	5'-TCCTTCTTTCTTTTTT-G(GA)G 5'-AGGAAGAAAGAAAAAA-3'-3'-G	47.8	48.6
2	3'-TTTTTTCTTTCTTCCT-G(GA)G 3'-AAAAAAGAAAGAAGGA-5'-5'-G	45.3	46.9
5+6	5'-TCCTTCTTTCTTTTTT-3' 5'-AGGAAGAAAGAAAAAA-3'	41.0	45.3
7	3'-TCCTTCTTTCTTTTTT-GGA 5'-AGGAAGAAAGAAAAAA-GG	67.4	69.2
3+5	3'-TCCTTCTTTCTTTTTT-5' 5'-AGGAAGAAAGAAAAAA-3'	48.6	52.0
4+5	3'-UCCUUCUUUCUUUUUU-5' 5'-AGGAAGAAAGAAAAAA-3'	32.7	32.1

[a] Arrows indicate polarity of the strands (5′ → 3′) and thereby the parallel or antiparallel nature of the duplexes. [b] In the presence of 1 mM spermine.

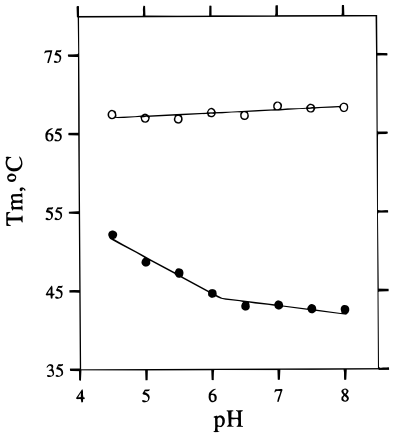

FIGURE 2: pH dependence of the Watson−Crick (oligonucleotide **7**; ○) and the Hoogsteen (oligonucleotide **1**; ●) hairpin duplex T_ms in 100 mM sodium acetate, and 10 mM magnesium chloride buffers.

also the modeling section). The T_ms of oligonucleotides **1** and **2** in pH 7.6 buffer in the presence of 1 mM spermine are comparable to those obtained in pH 5.0 buffer (Table 2). These results suggest that oligonucleotides **1** and **2** form stable Hoogsteen duplex structures (Kandimalla et al., 1995a) that have thermal stabilities lower than that of the WC hairpin duplex **7**.

pH Dependence of the Hoogsteen Duplex. Hoogsteen base pairing between G and C requires protonation of cytosine (C^+) which could occur below pH 6.0. We measured the T_ms of oligonucleotides **1** and **7** as a function of pH. Oligonucleotide **7** which forms WC hairpin duplex shows a pH-independent T_m over the entire range of pHs studied (4.5−8.0) (Figure 2). The duplex of oligonucleotide **1** shows decreasing thermal stability with increasing pH in the pH 4.5−6.0 range (Figure 2). Above pH 6.0, the structure does not show pH dependence. These results suggest that in the lower pH range (<6.0) the structure is Hoogsteen base-paired and at higher pH conditions it forms a pH-independent structure as reported earlier for linear Hoogsteen duplexes (Liu et al., 1992). These results suggest that (i) oligonucleotides **1** and **2** form pH-dependent Hoogsteen duplex structures and (ii) the thermal stability of the WC hairpin duplex is higher than that of the PS Hoogsteen duplex (**1** or **2**) over the entire range of pHs studied.

Circular Dichroism of the Hoogsteen Hairpin Duplex. Both oligonucleotides **1** and **2** have similar CD spectra. The CD spectra of these oligonucleotides are very different from the CD spectrum of the WC hairpin duplex of oligonucleotide **7** in pH 5.0 buffer (Figure 3). Oligonucleotide **7** exhibits a canonical B-type CD spectrum (Cantor & Schimmel, 1980; Gray et al., 1992, 1995) with positive maxima at 282 and 219 nm and a negative minimum at 247 nm (Figure 3B). The CD spectrum of the linear WC duplex of oligonucleotides **3**+**5** under the same (pH 5.0) experimental conditions is shown in Figure 3B for comparison. Oligonucleotides **1** and **2** show lower-intensity positive maxima (277 and 224 nm) with a negative minimum at 247 nm almost equal to that of the WC duplex CD spectrum (Figure 3A). In addition, the CD spectra of Hoogsteen duplexes have a new characteristic negative minimum around 213−215 nm. These CD spectral characteristics differentiate the secondary structures formed by oligonucleotides **1** and **2** from those of WC duplexes.

The CD spectra of oligonucleotides **1** and **2** exhibit a temperature dependence (Figure 4). The T_ms measured from thermally dependent CD curves correlate with the T_ms measured by UV thermal melting experiments within ±3.0 °C (data not shown). This result further supports the fact that oligonucleotides **1** and **2** form PS secondary structures.

pH-dependent secondary-structure formation by oligonucleotides **1** and **2** is also studied by CD spectroscopic measurements and compared with the formation of the WC hairpin duplex oligonucleotide (**7**). Oligonucleotide **7** shows similar spectra at both of the pHs (5.0 and 7.6) studied (data not shown), suggesting that the secondary structure formed by this oligonucleotide is not pH-dependent. By contrast, the CD spectra of oligonucleotides **1** and **2** show different spectral characteristics depending on the pH of the solution (Figure 3A). At pH 7.6, the CD spectrum looks like a B-DNA spectrum as in the case of oligonucleotide **7**.

We examined the reversible nature of the pH-dependent secondary-structure formation by oligonucleotides **1** and **2**. When the pH of a neutral solution containing oligonucleotide **1** or **2** is decreased to pH 5.0 with dilute HCl, a Hoogsteen duplex is formed readily, as determined by the CD spectrum (Figure 3A). Similarly, when the low-pH structure (Hoogsteen duplex) of these oligonucleotides is titrated with dilute alkali to raise the pH to neutrality (around 7.5), a structure with a CD spectrum similar to the one observed at pH 7.6 is formed (data not shown). Oligo[d(A)] and oligo[d(T)] sequences attached through 3′−3′ and 5′−5′ linkages form pH-independent PS duplex structures involving rWC hydrogen bonding (van de Sande, 1988; Rippe & Jovine, 1989). The presence of G·C base pairs in such sequences destabilizes rWC parallel duplex formation (Rentzeperis et al., 1992). The CD spectra of oligonucleotides **1** and **2** suggest that they do not form PS structures with rWC base pairing.

Effect of the Nucleotide Loop Size Joining 3′−3′ or 5′−5′ Ends. Studies with WC hairpin duplexes of RNA (Groebe & Uhlenbeck, 1988) and DNA (Haasnoot et al., 1986) showed that duplexes with four- or five-base loops are the most stable. In the case of foldback (Prakash & Kool, 1992; Kandimalla et al., 1996) and circular (Prakash & Kool, 1992) oligonucleotides that form triplexes with single-stranded purine target sequences, a loop of five or six bases is optimum for joining the two pyrimidine strands. This is also the case with H-DNA (Shimizu et al., 1989).

The nucleotide loop that bridges the two strands should have an appropriate length for optimum binding of the two binding domains. In addition, introduction of a 3′−3′ or 5′−5′ linkage results in an intranucleotide distance that is different from the normal 3′−5′ intranucleotide distance. Therefore, the length of a linker should compensate for these differences in intranucleotide and interstrand distances for optimum binding.

We have synthesized 3′−3′- and 5′−5′-linked oligonucleotides with loops of zero to three, five, and seven purine or pyrimidine nucleotides (see Table 3 for the loop sequences). All the oligonucleotides contain the same complementary binding sequences as oligonucleotides **1** and **2**. Measurement of the melting temperatures for these Hoogsteen duplexes (Table 3) with different lengths of pyrimidine loops reveals that 3′−3′-linked sequences prefer a three-base loop (Table 3), and a two-base loop is optimum for 5′−5′-linked oligonucleotides (Table 3). The results with different pyrimidine sequence loops suggest that parallel duplex

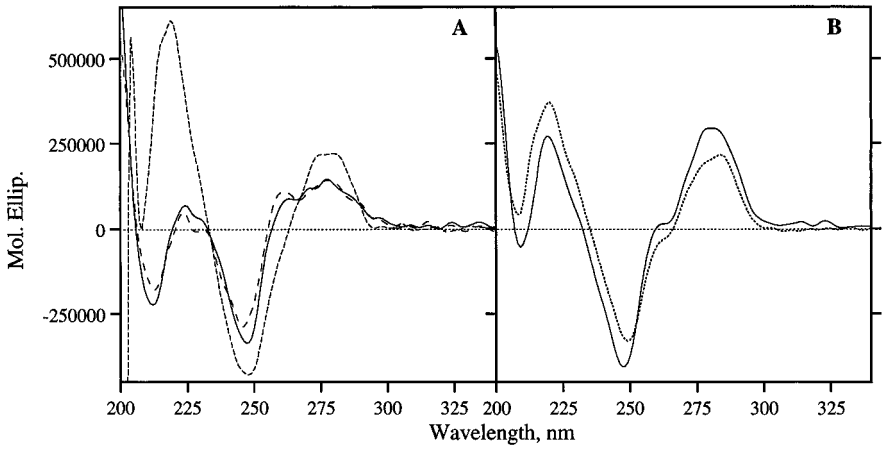

FIGURE 3: (A) CD spectrum of the Hoogsteen duplex of oligonucleotide **1** in 10 mM sodium cacodylate (pH 5.0) (—), 100 mM sodium chloride, and 10 mM magnesium chloride buffer. CD spectra of oligonucleotide **1** in pH 7.6 buffer (-----) and after adjusting the pH of the same solution (pH 7.6) to pH 5.0 with HCl and incubating for 15 min at 15 °C (- - -). (B) CD spectra of WC hairpin (**7**) (—) and linear (**3**+**5**) (···) antiparallel duplexes in 10 mM sodium cacodylate (pH 5.0), 100 mM sodium chloride, and 10 mM magnesium chloride buffer.

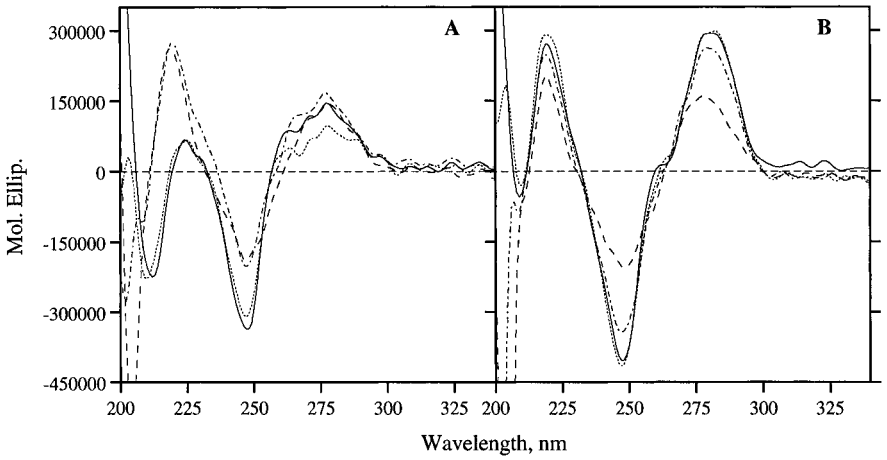

FIGURE 4: Temperature-dependent CD spectra of oligonucleotides (A) **1** and (B) **7** at 15 (—), 30 (···), 50 (-·-), and 70 °C (- - -). Experimental conditions were as in the legend for Figure 3.

Table 3: Effect of Loop Length and Sequence on the Hoogsteen Duplex Formation

loop size	sequence	T_m, °C[a] 3′−3′	5′−5′	sequence	T_m, °C[a] 3′−3′	5′−5′
0	−	43.8	44.5	−	43.2	44.0
1	G	45.3	46.0	C	45.4	45.8
2	AG	47.0	47.9	TC	46.6	47.7
3	GAG	47.8	46.4	CTC	47.5	46.0
5	GGAGG	47.8	45.3	CTCTC	45.0	43.3
7	GAGAGAG	45.6	43.8	CTCTCTC	43.0	41.6

[a] T_ms were measured in 100 mM sodium acetate (pH 5.0) and 10 mM magnesium chloride at a strand concentration of 1.2 μM at 260 nm as described in Experimental Procedures.

formation is not pyrimidine loop sequence-dependent. These results are supported by molecular modeling studies (see Molecular Modeling). When purine nucleotides are present in the loop, however, a different order of loop size preference and influence of loop bases on the Hoogsteen duplex stability are observed (Table 3). In the case of 3′−3′-linked oligonucleotides, two- to five-base loops appear to have a similar effect on duplex stability. Whereas in the case of 5′−5′-linked oligonucleotides, a two-base long purine loop is better. The higher duplex stabilizing effect of the GGAGG loop in 3′−3′-linked oligonucleotides may result from interaction of guanines in the loop. It is not clear why the GGAGG loop has a less stabilizing effect on a 5′−5′-linked duplex than on a 3′−3′-linked duplex. The interaction of guanines in the loop is suggested by molecular modeling. We used a five-base GGAGG loop in the studies reported here.

Triplex Formation with DNA/RNA Pyrimidine Complementary Strands. In the presence of a complementary pyrimidine strand (antiparallel to the purine strand, **3** or **4**), oligonucleotides **1** and **2** form triplexes. These triplexes with both DNA and RNA pyrimidine single strands are more stable than the triplex formed by adding a pyrimidine Hoogsteen third strand (**6**) to oligonucleotide **7** (WC hairpin duplex). The T_m data are shown in Table 4. The triplexes of oligonucleotides **1** and **2** with the DNA (**3**) and RNA (**4**) pyrimidine single strands showed single, cooperative melting transitions (not shown) with T_m values comparable to those obtained with foldback triplexes (Kandimalla & Agrawal, 1994; Kandimalla et al., 1995a,b).

Hoogsteen duplexes of oligonucleotides **1** and **2** formed stable triplexes (Kandimalla et al., 1995a), with the complementary pyrimidine DNA strand **3**, that have T_ms of 63.1 and 63.4 °C, respectively, at pH 5.0 (Table 4). These T_m values are about 14 and more than 15 °C higher than those of the corresponding linear WC (**3**+**5**) and the Hoogsteen hairpin duplexes (**1** or **2**), respectively, under the same experimental conditions. The conventional triplex of the WC hairpin duplex **7** and oligonucleotide **6** showed two distinct melting transitions at 24.6 and 67.2 °C. The lower-

Table 4: Thermal Melting Data of Triplexes

oligo number	oligonucleotide sequences/complexes[a]	T_m, °C pH 5.0	T_m, °C pH 7.6[b]
1 + 3	5'-TCCTTCTTTCTTTTTT-G G A G 5'-AGGAAGAAAGAAAAAA-3'-3'-G 3'-TCCTTCTTTCTTTTTT-5'	63.1	54.6
2 + 3	3'-TTTTTTCTTTCTTCCT-G G A G 3'-AAAAAAGAAAGAAGGA-5'-5'-G 5'-TTTTTTCTTTCTTCCT-3'	63.4	53.3
7 + 6	3'-TCCTTCTTTCTTTTTT-G G A 5'-AGGAAGAAAGAAAAAA-G G 5'-TCCTTCTTTCTTTTTT-3'	24.6[c], 67.2	18.2[c], 69.3
1 + 4	5'-TCCTTCTTTCTTTTTT-G G A G 5'-AGGAAGAAAGAAAAAA-3'-3'-G 3'-UCCUUCUUUCUUUUUU-5'	53.3	41.1
2 + 4	3'-TTTTTTCTTTCTTCCT-G G A G 3'-AAAAAAGAAAGAAGGA-5'-5'-G 5'-UUUUUUCUUUCUUCCU	53.7	43.1

[a] Arrows indicate polarity of each strand (5′ → 3′). [b] In the presence of 1 mM spermine. [c] Lower melting transition that corresponds to third-strand (triplex) dissociation.

temperature transition is a result of the dissociation of the third strand (**6**) from the duplex, and the higher-temperature transition is the result of denaturation of the hairpin duplex itself (Roberts & Crothers, 1992). The triplexes involving oligonucleotides **1** and **2** with the RNA pyrimidine stand **4** also showed cooperative single melting transitions (figure not shown) at 53.3 and 53.7 °C, respectively, at pH 5.0. These are about 20 and 6 °C higher than the corresponding linear WC heteroduplex (**4**+**5**) and the Hoogsteen duplexes (**1** or **2**), respectively, under the same experimental conditions.

In order to study triplex formation at physiological pH conditions, we have carried out experiments at pH 7.6 in the presence of 1 mM spermine (Moser & Dervan, 1987; Hanvey et al., 1991; Han & Dervan, 1993) with both the DNA and RNA complementary strands. The T_m data are shown in Table 4. In both cases, we observed a single melting transition, indicating that the Hoogsteen and WC hydrogen-bonded strands dissociate simultaneously (data not shown). The increase in the T_m that results from triplex formation, compared to the T_ms of corresponding duplexes, is lower than that observed at pH 5.0. Comparison of the T_m values suggests that the RNA complementary strand forms a more stable triplex at physiological pH conditions than at lower pH. The T_m data at pH 7.6 in Tables 2 and 4 suggest that spermine has a higher stabilizing effect on PS duplexes than the triplexes. Polyamines such as spermine and spermidine stabilize DNA duplex and triplex structures by interacting with phosphate backbone charges (Hampel et al., 1991; Thomas & Thomas, 1993). However, the precise mechanism of promotion of C^+•G:C triplexes by polyamines at neutral pH is not well understood. The basis for the different stabilizing effects of spermine on PS duplexes and their triplexes is unknown.

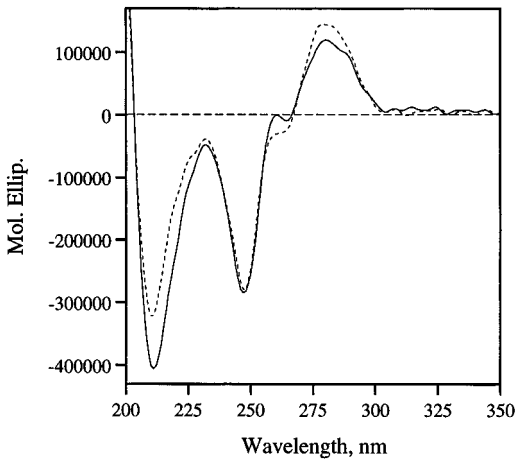

FIGURE 5: CD spectra of the triplexes of oligonucleotides **1**+**3** (—) and **7**+**6** (- - -) in 10 mM sodium cacodylate, (pH 5.0), 100 mM sodium chloride, and 10 mM magnesium chloride buffer.

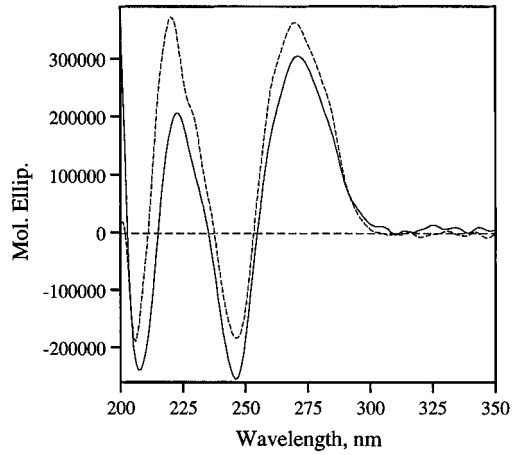

FIGURE 6: CD spectra of the triplex of oligonucleotides **1**+**4** (—) and the WC heteroduplex of oligonucleotides **4**+**5** (- - -).

Circular Dichroism of Triplexes. The triplexes of oligonucleotides **1** and **2** with the DNA complementary pyrimidine strand show a positive maximum around 282 nm and two negative minima around 248 and 215 nm (Figure 5). The CD spectrum of the triplex of oligonucleotide **7** with the pyrimidine DNA third strand (**6**) is similar to the CD spectra of triplexes of oligonucleotides **1** and **2** with **3**. This result confirms that the global structure of the triplex of oligonucleotide **1** or **2** (with **3**) is similar to the triplex of oligonucleotides **7**+**6**. The triplex of an FTFO (Kandimalla et al., 1996), in which the Watson−Crick and Hoogsteen pyrimidine strands are connected by a nucleotide loop (Figure 1B), and a complementary purine strand also has a similar CD spectrum shown in Figure 5 (data not shown). These CD results support the finding that the triple-helix structure is not affected by the way in which the two strands of a triplex are attached (Figure 1). The thermal stability of the triplex is dramatically affected; however, the conventional triplex (Figure 1A) is much less stable than the other two types of triplexes (Figure 1B,C).

The triplexes of oligonucleotides **1** and **2** with the RNA pyrimidine complementary strand (**4**) exhibited different CD spectral characteristics than the triplexes formed with the DNA pyrimidine complementary strand (Figure 6). The CD spectrum of the triplex with the RNA pyrimidine strand has two positive maxima at 271 and 223 nm and two negative minima at 247 and 211 nm. The CD spectrum of a linear

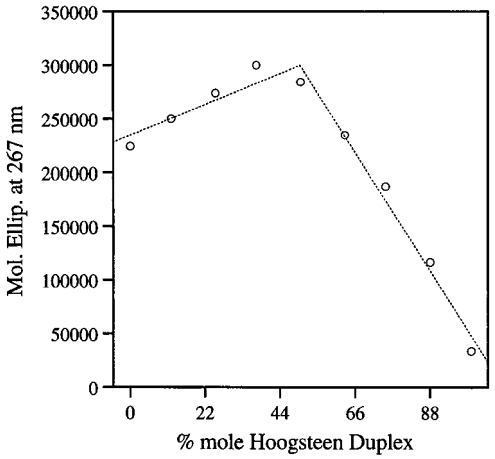

FIGURE 7: CD mixing curve for the interaction of oligonucleotides **1** and **4**.

WC heteroduplex of oligonucleotide **4** and its complementary DNA purine strand **5** is shown in Figure 6 for comparison. The CD spectrum of this heteroduplex indicates an A-type structure (Gray et al., 1995; Kandimalla et al., 1996). Compared with the heteroduplex, the triplex of oligonucleotide **1** (or **2**) with the RNA pyrimidine strand **4** appears to be in the A-type conformation. It is difficult, however, to interpret the sugar geometry and the strand conformation solely on the basis of the CD data in the absence of IR or other spectroscopic data for such complexes. The CD spectra of the triplexes studied exhibit a temperature dependence (data not shown). The T_ms measured from thermally dependent CD curves correlate with the T_ms measured by UV thermal melting experiments (±3.0 °C) (data not shown).

Stoichiometry of the Interaction of the Hoogsteen Duplex with Pyrimidine Single Strands. The stoichiometry of the interaction of oligonucleotide **1** with both DNA (**3**) and RNA (**4**) complementary pyrimidine strands is studied by CD mixing experiments. The mixing curves demonstrate that oligonucleotide **1** interacts with **4** (Figure 7) and **3** with 1:1 stoichiometry. The 1:1 stoichiometric interaction between **1** and **3** or **4** may not, however, indicate triplex formation. That is, the 3′−3′ (**1**) or 5′−5′ (**2**) oligonucleotide might only be capable of forming a WC base-paired duplex involving its purine domain with the complementary RNA/DNA pyrimidine target strand in 1:1 ratio, while the Hoogsteen pyrimidine domain remains as an overhanging sequence. The sharp, cooperative UV melting profiles (vs broad melting profile for a duplex with a long overhanging Hoogsteen sequence) (Kandimalla & Agrawal, 1994; Kandimalla et al., 1995b), the CD spectral characteristics, and the difference CD spectra do suggest triplex formation.

Molecular Modeling. We explored whether stereochemically feasible loop structures can be generated without and with hydrogen bonding between the guanines in the loop in order to rationalize the higher T_ms observed with hairpin duplexes containing guanines in the loop (Table 3). The models suggest that the five-base loops in both 3′−3′- and 5′−5′-attached oligonucleotides are capable of adopting a range of stereochemically feasible conformations. In one of these conformations, the bases in the loop are oriented radially away from the loop backbone, but in random directions (Figure 8A). There is no observable stacking or hydrogen bonding between the bases in the loop region or between the loop and the purine or pyrimidine strands. The

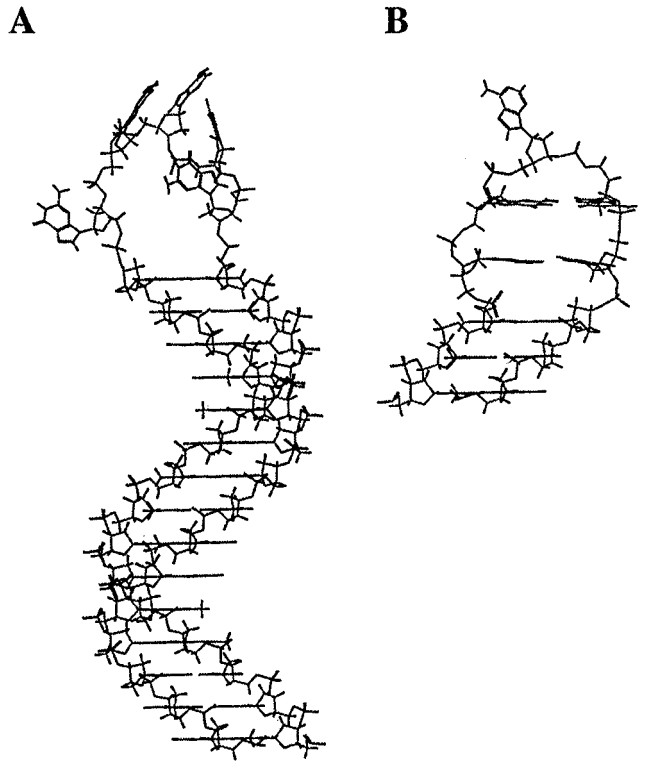

FIGURE 8: Loop of the 3′−3′-attached Hoogsteen PS duplex (A) in the absence of any hydrogen-bonding and stacking constraints and (B) with hydrogen-bonding constraints (only first four bases of the duplex adjacent to the loop are shown for clarity).

conformation of the Hoogsteen duplex is almost identical to the conformation reported earlier (Raghunathan et al., 1994). Using the loop framework shown in Figure 8A, we imposed hydrogen-bonding restraints on the guanines in the loop (Mohanty & Bansal, 1993, 1994). When only one pair of guanines is restrained, the backbone of the loop converges readily to a viable structure. Interestingly, stereochemically feasible structures are obtained even if both guanine pairs in the loop are hydrogen bonded, and the adenine forms a hairpin turn linking the guanine pairs (Figure 8B). In this situation as well, the Hoogsteen duplex conformation is not significantly altered.

The distance between the C3′ of the purine strand and the C3′ of the Hoogsteen strand in the PS duplex is about 10.8 Å. Using an average distance of about 6.5 Å for each nucleotide (Saenger, 1984), plus an additional 2 Å for the additional phosphate group, the spacer should be at least two nucleotides long in order to traverse this distance in the case of a 3′−3′ linkage. A one nucleotide long spacer will distort the sugar geometry of the end bases along the duplex, causing a destabilization of the Hoogsteen-bonded duplex. For very long loops, however, additional destabilization resulting in slippage of the duplex structure could occur. In the case of a 5′−5′ linkage, the distance between the C4′ atoms of the Hoogsteen strand and the purine strand is about 10.8 Å. In this case, the modeling studies indicate that a single nucleotide can traverse this distance as the C4′−C5′ bond of the end bases in the two strands can adjust to form a loop. These findings are in support of the experimental results observed with different loop size PS duplexes (Table 3).

Introduction of the WC pyrimidine strand for triplex formation does not alter the conformation or relative position of the purine and the Hoogsteen-paired pyrimidine strands.

Both the loop structures modeled earlier in the absence of the WC-paired pyrimidine strand are still possible. Since the loop is well removed from the pyrimidine strand, there are no additional interactions between the residues in the loop and the pyrimidine strand. The loop orientation is also such that there would be no hindrance for the third WC-paired pyrimidine strand even it were to continue in 5′-direction. Stereochemically feasible models for hetero-triplex structures, where oligonucleotide **1** and **2** interact with **4**, were also modeled (Kandimalla et al. unpublished results).

Conclusions. We have shown that complementary oligopurine-oligopyrimidine sequences attached through 3′−3′ and 5′−5′ linkages from stable Hoogsteen hairpin duplexes. The studies with oligonucleotides of different loop sizes reveal that three- and two-base loops are optimum for the formation of stable Hoogsteen duplexes with 3′−3′ and 5′−5′ linkages, respectively. The guanine residues in the loop stabilize the duplex as a result of G−G interactions. A Hoogsteen duplex has CD characteristics different from those of a WC duplex but similar to those of a triplex. The parallel-stranded Hoogsteen duplexes readily interact with complementary pyrimidine RNA and DNA single strands, forming stable triplexes. This allows for targeting of pyrimidine single strands through triplex formation in the Py·Pu:Py motif for molecular biological and gene expression control studies.

ACKNOWLEDGMENT

We thank Ms. Adrienne Manning for excellent technical assistance and Prof. V. Sasisekharan and Dr. G. Venkataraman for help with molecular modeling and suggestions.

REFERENCES

Blackburn, E. H. (1991) *Nature 350*, 569−573.
Cantor, C. R., & Schimmel, P. R. (1980) in *Biophysical Chemistry*, W. H. Freeman, San Francisco, CA.
Cheng, Y.-K., & Pettitt, B. M. (1992) *Prog. Biophys. Mol. Biol. 58*, 225−257.
D'Souza, D. J., & Kool, E. T. (1992) *J. Biomol. Struct. Dyn. 10*, 141−152.
Escude, C., Francois, J.-C., Sun, J.-S., Ott, G., Sprinzl, M., Garestier, T., & Helene, C. (1993) *Nucleic Acids Res. 21*, 5547−5553.
Escude, C., Mohammadi, S., Sun, J.-S., Nguyen, C.-H., Bisagni, E., Liquier, J., Taillandier, E., Garestier, T., & Helene, C. (1996) *Chem. Biol. 3*, 57−65.
Giovannangeli, C., Montenay-Garestier, T., Rougee, M., Chassignol, M., Thoung, N. T., & Helene, C. (1991) *J. Am. Chem. Soc. 113*, 7775−7776.
Gray, D. M., Ratliff, R. L., & Vaughan, M. R. (1992) *Methods Enzymol. 211*, 389−406.
Gray, D. M., Hung, S.-H., & Johnson, K. H. (1995) *Methods Enzymol. 246*, 19−35.
Groebe, D. R., & Uhlenbeck, O. C. (1988) *Nucleic Acids Res. 16*, 11725−11735.
Haasnoot, C. A. G., Hilbers, C. W., van der Marel, G. A., van Boom, J. H., Singh, U. C., Pattabiraman, N., & Kollman, P. A. (1986) *J. Biomol. Struct. Dyn. 3*, 843−857.
Hakoshima, T., Fukui, T., Ikehara, M., & Tomita, K.-I. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 7309−7313.
Hampel, K. J., Crosson, P., & Lee, J. S. (1991) *Biochemistry 30*, 4455−4459.
Han, H., & Dervan, P. B. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3806−3810.
Hanvey, J. C., Williams, E. M., & Besterman, J. M. (1991) *Antisense Res. Dev. 1*, 307−317.
Hirao, I., Nishimura, Y., Tagawa, Y., Watanabe, K., & Miura, K. (1992) *Nucleic Acids Res. 20*, 3891−3896.
Hoogsteen, K. (1963) *Acta Crystallogr. 16*, 907−916.
Ikehara, M., Hattori, M., & Fukui, T. (1972) *Eur. J. Biochem. 31*, 329−334.
Ishikawa, F., Frazier, J., Howard, F. B., & Miles, H. T. (1972) *J. Mol. Biol. 70*, 475−490.
Kandimalla, E. R., & Agrawal, S. (1993) *J. Biomol. Struct. Dyn. 10*, a086.
Kandimalla, E. R., & Agrawal, S. (1994) *Gene 149*, 115−121.
Kandimalla, E. R., Agrawal, S., Venkataraman, G., & Sasisekharan, V. (1995a) *J. Am. Chem. Soc. 117*, 6416−6417.
Kandimalla, E. R., Manning, A. N., & Agrawal, S. (1995b) *J. Biomol. Struct. Dyn. 13*, 483−491.
Kandimalla, E. R., Manning, A. N., Venkataraman, G., Sasisekharan, V., & Agrawal, S. (1995c) *Nucleic Acids Res. 21*, 4510−4517.
Kandimalla, E. R., Manning, A. N., & Agrawal, S. (1996) *J. Biomol. Struct. Dyn. 14*, 79−90.
Kool, E. T. (1991) *J. Am. Chem. Soc. 113*, 6265−6266.
Lavelle, L., & Fresco, J. R. (1995) *Nucleic Acids Res. 23*, 2692−2705.
Liu, K., Miles, H. T., Frazier, J., & Sasisekharan, V. (1993) *Biochemistry 32*, 11802−11809.
Maher, L. J., III (1992) *BioEssays 14*, 807−815.
Mohanty, D., & Bansal, M. (1993) *Nucleic Acids Res. 21*, 1767−1774.
Mohanty, D., & Bansal, M. (1994) *Biopolymers 34*, 1187−1211.
Moser, H., & Dervan, P. B. (1987) *Science 238*, 645−650.
Noll, D. M., O'Rear, J. L., Cushman, C. D., & Miller, P. S. (1994) *Nucleosides Nucleotides 13*, 997−1005.
Prakash, G., & Kool, E. T. (1992) *J. Am. Chem. Soc. 114*, 3523−3528.
Puglisi, J. D., & Tinoco, I., Jr. (1989) *Methods Enzymol. 180*, 304−325.
Raghunathan, G., Miles, H. T., & Sasisekharan, V. (1994) *Biopolymers 34*, 1573−1581.
Rentzeperis, D., Rippe, K., Jovin, T. M., & Marky, L. A. (1992) *J. Am. Chem. Soc. 114*, 5926−5928.
Rippe, K., & Jovine, T. M. (1989) *Biochemistry 28*, 9542−9549.
Roberts, R. W., & Crothers, D. M. (1992) *Science 258*, 1463−1466.
Saenger, W. (1984) in *Principles of Nucleic Acid Structure*, Springer-Verlag, Berlin.
Shimizu, M., Hanvey, J. C., & Wells, R. D. (1989) *J. Biol. Chem. 264*, 5944−5949.
Smith, P. J. C., & Arnott, S. (1978) *Acta Crystallogr. A34*, 3−11.
Thomas, T., & Thomas, T. J. (1993) *Biochemistry 32*, 14068−14074.
Thoung, N. T., & Helene, C. (1993) *Angew. Chem., Int. Ed. Engl. 32*, 666−690.
van de Sande, J. H., Ramsing, N. B., Germann, M. W., Elhorst, W., Kalisch, B. W., Kitzing, E., Pon, R. T., Clegg, R. M., & Jovin, T. M. (1988) *Science 241*, 551−557.
Wang, S., & Kool, E. T. (1994) *Nucleic Acids Res. 22*, 2326−2333.
Wilkins, M. H. F., Arnott, S., Mavin, D. A., & Hamilton, L. D. (1970) *Science 167*, 1693−1694.
Xodo, L. E., Manzini, G., & Quadrifoglio, F. (1990) *Nucleic Acids Res. 18*, 3557−3564.

BI961505Y