Sequence Limitations of Triple Helix Formation by Alternate-Strand Recognition†

Sumedha D. Jayasena*,‡ and Brian H. Johnston*

Cell and Molecular Biology Laboratory, SRI International, 333 Ravenswood Avenue, Menlo Park, California 94025

Received September 3, 1992; Revised Manuscript Received December 15, 1992

ABSTRACT: Until recently, oligonucleotide-directed triplex formation has been limited to oligopurine tracts of target DNA. Triplex formation by alternate-strand recognition relaxes this limitation by allowing triplexes to form at 5′-$(Pu)_m(Py)_n$-3′ and 5′-$(Py)_m(Pu)_n$-3′ sequences, with the third strand pairing first with purines on one strand and then switching to pair with purines on the other strand. In this study, the interaction of several oligonucleotides with the potential to form triplexes by alternate-strand recognition at the sequence 5′-$A_8C_8A_8$-3′ was studied by chemical probing and affinity cleaving. The results show that triplex formation can be readily accomplished at the 5′-A_8C_8-3′ part of the sequence; however, base triplet formation is disrupted on either side of the strand switch and the Watson–Crick helix is distorted in such a way as to expose the N7 positions of purines adjoining the strand switch. Triplex formation is weak or nonexistent at the 3′-most A_8 block, despite the opportunity for recruiting a spacer sequence for the second (C_8-A_8) strand switch by "slippage". This finding indicates that the C_8-A_8 strand switch is energetically unfavorable, although pairing at other 5′-$(Py)_n(Pu)_n$-3′ sequences has been observed, with or without a spacer [Beal, P. A., & Dervan, P. B. (1992) *J. Am. Chem. Soc. 114*, 1470–1478; Jayasena, S. D., & Johnston, B. H. (1992) *Nucleic Acids Res. 20*, 5279–5288]. Thus, alternate-strand recognition may not be feasible for certain sequences of 5′-$(Py)_m(Pu)_n$-3′, at least under the conditions examined.

Site-specific recognition of nucleic acid sequences has become an area of active study, in part because of the therapeutic potential of an ability to control expression or to cleave DNA at specific genes. Site-specific complexation of oligonucleotides with intact double-helical DNA sequences has been accomplished by forming triple helices (Moser & Dervan, 1987; Le Doan et al., 1987). A substantial body of physical and chemical evidence indicates that the nucleotide bases of the third strand occupy the major groove of the target duplex and form specific hydrogen bonds with bases of the Watson–Crick duplex (Felsenfeld et al., 1957; Morgan & Wells, 1968; Lee et al., 1979; Praseuth et al., 1988; Johnston, 1988; Rajagopal & Feigon, 1989; de los Santos et al., 1989). The specificity of these hydrogen-bonding interactions dictates the specific sequence of a duplex with which a third strand can form a triplex.

Depending on the nature of the third-strand oligonucleotide, two principal classes of triplexes can be identified. An R-type triple helix (Figure 1a) stabilized by Pu·PuPy base triplets (G·GC and A·AT) can be formed by an oligopurine strand that binds antiparallel to the purine tract of the Watson–Crick duplex (Beal & Dervan, 1991; Pilch et al., 1991; Radhakrishnan et al., 1991a). An oligopyrimidine strand that binds parallel to the purine tract of the Watson–Crick duplex forms a Y-type triple helix stabilized by Py·PuPy base triplets (T·AT and C^+·GC) (Moser & Dervan, 1987; Le Doan et al., 1987). So far, except for the base triplet G·TA (Griffin & Dervan, 1989; Radhakrishnan et al., 1991b), the nucleotide bases of the third strand have been found to form specific hydrogen bonds only with purines of the target duplex. The purine recognition by the third strand nucleotides substantially limits the type of DNA sequences that can be targeted. As a result, triplex formation has been limited to DNA tracts made up largely or exclusively of purines [see Cooney et al. (1988), Beal and Dervan (1991), Orson et al. (1991), and Blume et al. (1992) for examples of sequences containing a few pyrimidines within a purine block].

To overcome the sequence limitation and to target DNA sequences containing both purines and pyrimidines, Horne and Dervan (1990) proposed an approach involving the recognition of blocks of purines on alternate strands by a homopyrimidine third-strand molecule. This approach, called alternate-strand triple-helix formation, required the synthesis of third-strand oligonucleotides having either 3′–3′ (Horne & Dervan, 1990; Ono et al., 1991; Froehler et al., 1992) or 5′–5′ phosphodiester linkages (Ono et al., 1991) between two pyrimidine blocks that were designed to recognize purine tracts on alternate strands of the target via Py·PuPy base triplets. Recently, we described an approach to overcome the synthetic requirement of unnatural interphosphate linkages in the third strand in which both classes of natural base triplets, Pu·PuPy and Py·PuPy, participate in triplex formation (Jayasena & Johnston, 1992a). In our approach, a third strand consisting of both purine and pyrimidine blocks and having normal phosphodiester linkages throughout pairs with purines in the Watson–Crick duplex, switching strands at the junction between the oligopurine and oligopyrimidine tracts but maintaining the required strand polarity. The success of this approach in targeting both 5′-$(Pu)_m(Py)_n$-3′ and 5′-$(Py)_m$-$(Pu)_n$-3′ target sequences has recently been demonstrated (Beal & Dervan, 1992; Jayasena & Johnston, 1992b). However, both studies show that triplex formation at 5′-$(Py)_m(Pu)_n$-3′ sequences [producing an R triplex followed by a Y triplex, designated an RY triplex (Figure 1a; Jayasena & Johnston, 1992b)] occurs less efficiently than triplex formation at 5′-$(Pu)_m(Py)_n$-3′ sequences (producing a YR triplex). Furthermore, in one study (Beal & Dervan, 1992), efficient RY triplex formation required a 2-nucleotide spacer between purine and pyrimidine blocks, whereas in the other study (Jayasena & Johnston, 1992b) such a spacer was not essential to form an RY triplex at a different sequence, also of the 5′-$(Py)_m(Pu)_n$-

† This work was supported by NIH Grant 1R29GM41423 and the SRI Internal Research and Development Program.

* Correspondence can be addressed to either author.

‡ Present address: NeXagen, Inc., 2860 Wilderness Place, Boulder, CO 80301.

0006-2960/93/0432-2800$04.00/0

3′ type. These results suggest that triplex formation by alternate-strand recognition at 5′-$(Py)_m(Pu)_n$-3′-type sequences may be highly sequence-dependent.

The current study focuses on triplex formation at a defined target sequence consisting of homopurine and homopyrimidine blocks (5′-$A_8C_8A_8$-3′) by several oligonucleotides having the potential to generate triplexes with either YR or RY or both types of strand switches. The runs of identical nucleotides allow triplex formation with potential recruitment of spacer nucleotides by "slippage". Affinity-cleaving data, together with chemical-probing results, indicate that a YR-type triplex is formed at the A_8C_8 portion of the target. No clear evidence is seen for triplex formation at the 3′ end of the target sequence (C_8A_8), where the RY junction is required. These results indicate that triplex formation by alternate-strand recognition with third strands containing natural phosphodiester bonds may not be generally applicable to all 5′-$(Py)_n(Pu)_n$-3′-type sequences. However, this approach appears to be a general technique for targeting 5′-$(Pu)_n(Py)_n$-3′ sequences.

MATERIALS AND METHODS

Materials. Nucleoside phosphoramidites and other general chemicals for DNA synthesis were purchased from Applied Biosystems, C6-Thiolmodifier was from Clonetech, T4 polynucleotide kinase was from United States Biochemicals, 5-nitro-1,10-phenanthroline was from G. F. Smith (Columbus, OH), and radiolabeled nucleotide triphosphates were from Du Pont NEN Research Products. Dimethyl sulfate (DMS) and diethyl pyrocarbonate (DEPC) were obtained from Aldrich, and restriction enzymes came from New England Biolabs.

Oligonucleotide Synthesis. Oligonucleotides were synthesized on an Applied Biosystems Model 380A automated DNA synthesizer using β-cyanoethyl phosphoramidite chemistry. For 1,10-phenanthroline (OP) coupling, a thiol moiety was linked to the 5′ ends of oligonucleotides by using C6-Thiolmodifier according to the manufacturer's instructions.

Attachment of 1,10-Phenanthroline (OP) to Oligonucleotides. 5-Iodoacetamido-1,10-phenanthroline was synthesized from 5-nitro-1,10-phenanthroline by using the method of Chen and Sigman (1986).

After solid-phase synthesis of oligonucleotides containing protected 5′ thiol groups, the bases were deprotected with ammonia and the deprotection of the 5′ thiol group was performed according to the manufacturer's instructions. The reactive 5′-SH-oligonucleotides (10 A_{260} units in ~200 μL) were mixed with 10 μL of a saturated solution of 5-iodoacetamido-1,10-phenanthroline in dimethylformamide. The mixture was incubated at 4 °C overnight, and unreacted 5-iodoacetamido-1,10-phenanthroline was removed by gel filtration through Sephadex G-50 equilibrated in TE buffer. Finally, the OP-modified oligonucleotides were purified by electrophoresis on 20% denaturing polyacrylamide gels.

Construction of Target DNA. The complementary oligonucleotides 5′-AGCTAAAAAAAACCCCCCCC-AAAAAAA-3′ and 3′-TTTTTTTTGGGGGGGG-TTTTTTTCTAG-5′, which carry the potential triplex-forming motif flanked by *Hin*dIII and *Bam*HI linkers, were synthesized and cloned into pUC18 DNA that had been previously digested with *Hin*dIII and *Bam*HI, using standard techniques (Sambrook et al., 1989). A plasmid carrying the 5′-$A_8C_8A_8$-3′ insert, called pTGT-IV, was purified from a large-scale culture by using a Qiagen plasmid purification kit according to the manufacturer's instructions.

Radiolabeled DNA Substrates. Plasmid pTGT-IV was linearized with *Xmn*I, dephosphorylated with calf intestine phosphatase, and treated with [γ-^{32}P]ATP and T4 polynucleotide kinase to obtain full-length radiolabeled plasmid DNA. For 3′ end labeling, plasmid DNA was initially digested with *Nar*I and then with *Hae*II. The 475-base pair (bp) DNA fragment was gel-purified, and the bottom strand was radiolabeled by filling in with [α-^{32}P]dCTP using DNA polymerase I (Klenow fragment) at the *Nar*I site. For 5′ end-labeling, plasmid DNA was initially digested with *Fsp*I, dephosphorylated, and then precipitated with ethanol. The *Fsp*I-digested DNA was then cleaved with *Hae*II, and the 475-bp restriction fragment was gel-purified and radiolabeled at the dephosphorylated *Fsp*I end with [γ-^{32}P]ATP by using T_4 polynucleotide kinase.

Affinity Cleaving. In a typical experiment, 10 μL of a mixture containing about 10 000 counts/min (cpm) radiolabeled target DNA, 3 μM OP-oligonucleotide (oligonucleotide derivatized with 1,10-phenanthroline at the 5′ end), 100 mM NaCl, 50 mM Tris-HCl (pH 6.8), 10 mM $MgCl_2$, 1 mM spermine, 1 μg of sonicated salmon sperm DNA, and 10% ethylene glycol was incubated at 0 °C for 10 min to allow triplex formation. Then, $CuSO_4$ and mercaptopropionic acid (MPA) were added to final concentrations of 1 μM and 2.5 mM, respectively (Jayasena & Johnston, 1992c). After incubation for 12 h at 20 °C, the cleavage reaction was quenched by adding 2,9-dimethyl-1,10-phenanthroline to 3 mM and the cleavage products were separated by gel electrophoresis. Gels were dried and visualized by autoradiography.

Chemical Sequencing. G and A + G ladders of uniquely end-labeled DNA were prepared according to Maxam and Gilbert (1980) with a slight modification. For the G-specific reaction, 9 μL of end-labeled DNA in TE (10 mM Tris-HCl and 0.1 mM EDTA, pH 7.5) buffer was mixed with 1 μL of a freshly prepared 1:200 dilution of dimethyl sulfate (DMS) in distilled water and the mixture was incubated for 10 min at ambient temperature. For the purine-specific reaction, 9 μL of end-labeled DNA was reacted with 1 μL of 1 M piperidine formate (pH 2.0) in TE buffer for 5 min at 60 °C. The chemically modified DNAs were then treated with 1 M piperidine at 90 °C for 30 min (Maxam & Gilbert, 1980).

Modification of Triplex DNA with DMS. Triplex formation between 3′-labeled DNA and 3 μM OP-oligonucleotide was carried out in a 10-μL volume as described for affinity cleaving, but salmon sperm DNA was omitted. The DMS modification was performed as described for chemical sequencing.

Modification of Triplex DNA with DEPC. After allowing triplex formation between 5′-labeled DNA and 3 μM OP-oligonucleotide in a 10-μL volume of the same buffer used for affinity cleaving, 1 μL of DEPC was added and the samples were mixed and incubated at 20 °C for 15 min. The modified DNA was then precipitated with ethanol, washed twice with 70% ethanol, dried, and treated with 1 M piperidine at 90 °C for 30 min.

RESULTS AND DISCUSSION

The target sequence used, 5′-$A_8C_8A_8$-3′ from plasmid pTGT-IV, has the general form 5′-$(Pu)_n(Py)_n(Pu)_n$-3′. Triple helices can potentially be formed at this target sequence by the five third-strand oligonucleotides equipped with a 1,10-orthophenanthroline (OP) cleaving moiety at their 5′ ends, OP-1 through OP-5, as illustrated in Figure 1b. All except OP-5 can potentially pair by alternate-strand recognition. OP-1 (OP-T_8G_8) can potentially form a triplex of YR-type at the

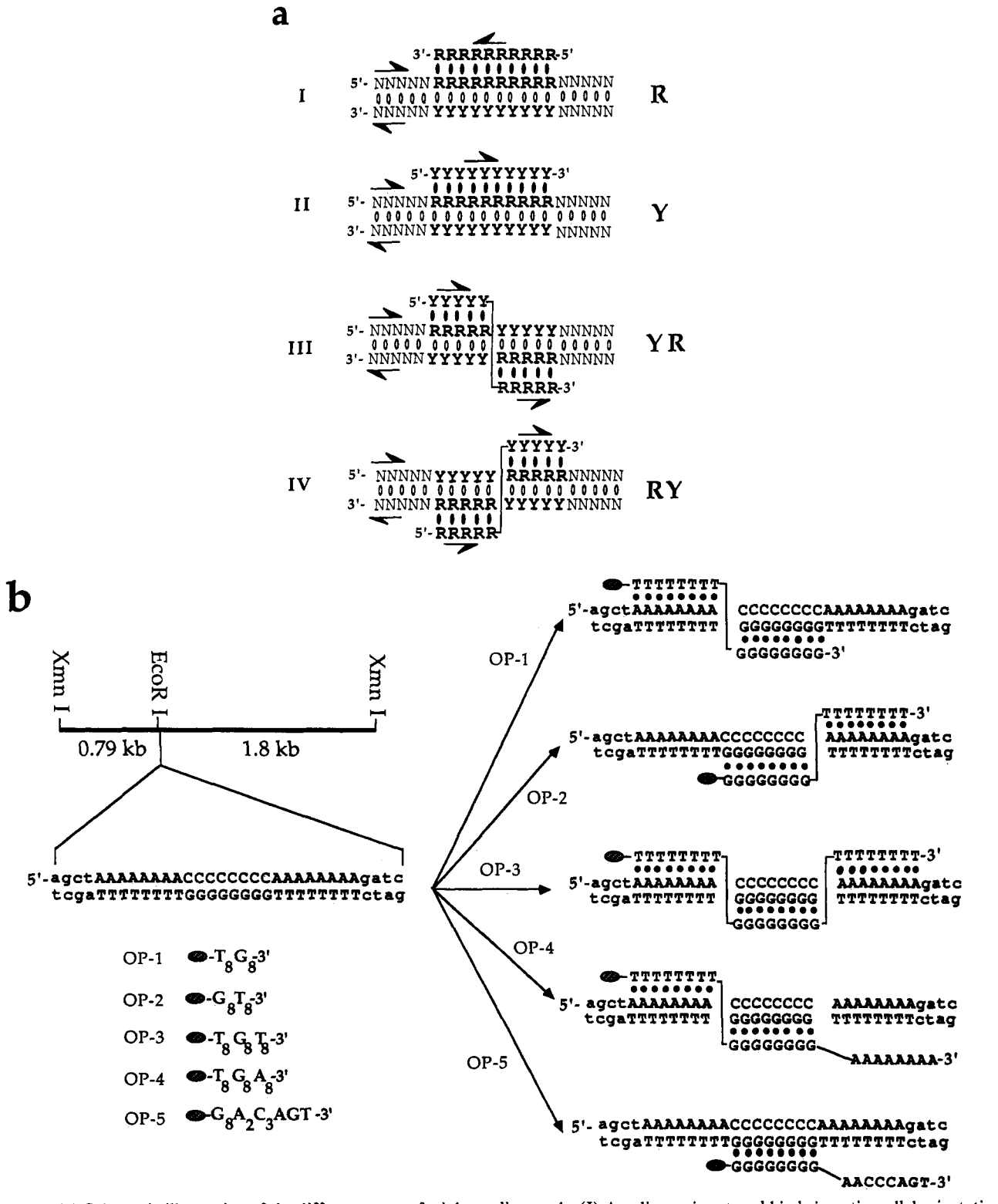

FIGURE 1: (a) Schematic illustration of the different types of triplexes discussed. (I) An oligopurine strand binds in antiparallel orientation to the purine tract of a target duplex to form an R-type triplex stabilized by Pu·PuPy base triplets. (II) An oligopyrimidine strand binds parallel to the purine tract of a target duplex to form a Y-type triplex stabilized by Py·PuPy base triplets. (III) A third strand of the form 5′-$(Py)_n(Pu)_n$-3′ binds to the purine tracts of alternate strands within a 5′-$(Pu)_n(Py)_n$-3′-type target sequence to form a YR-type triplex [$(Py{\cdot}PuPy)_n(Pu{\cdot}PuPy)_n$ viewed 5′ to 3′ along the third strand]. (IV) A 5′-$(Pu)_n(Py)_n$-3′ third strand binds to a 5′-$(Py)_n(Pu)_n$-3′ target sequence to form an RY-type triplex [$(Pu{\cdot}PuPy)_n(Py{\cdot}PuPy)_n$]. Open ovals represent Watson–Crick hydrogen bonding, and filled ovals represent Hoogsteen or Pu·Pu-type hydrogen bonding. The polarity of each strand is shown by a half-arrow. (b) Location and sequence of the target sequence within pTGT-IV plasmid are shown on the left. On the right are the structures of potential triplexes that can be formed by oligonucleotides OP-1 through OP-5. The ovals at the 5′ ends of the oligonucleotides represent 1,10-phenanthroline (OP) moieties. Filled circles (●) represent Hoogsteen or Pu·Pu-type hydrogen bonding.

5′-A_8C_8-3′ region of the target, OP-2 (OP-G_8T_8) has the potential to form an RY-type triplex at the 5′-C_8A_8-3′ region, and OP-3 (OP-$T_8G_8T_8$) can potentially use the entire sequence of the target to form a triplex having two strand switches. In contrast to the triple helix formed by OP-3 in which DNA strands have the right polarities in all three segments, a triplex

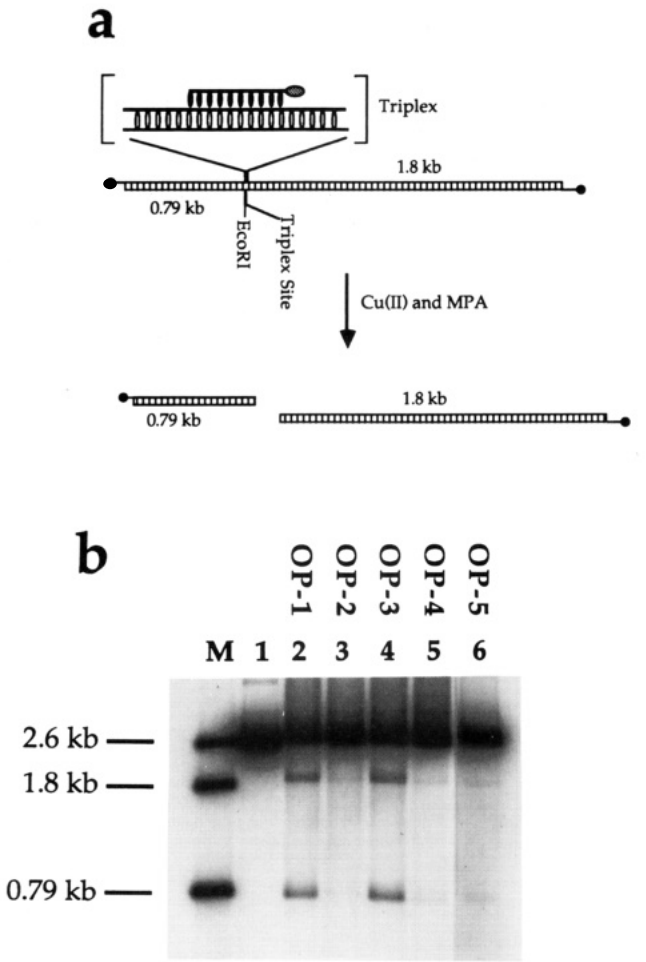

FIGURE 2: (a) Schematic representation of affinity cleavage of pTGT-IV plasmid DNA. Site-specific cleavage of plasmid DNA labeled at both ends is induced by the formation of a triple helix with an appropriate oligonucleotide equipped with a 1,10-phenanthroline moiety (stippled oval) in the presence of Cu(II) ions and mercaptopropionic acid (MPA), cleaving the plasmid into two DNA fragments of 1.8 and 0.79 kb. Open ovals represent Watson–Crick base pairing, and filled ovals represent Hoogsteen or Pu·Pu-type pairing. (b) Autoradiogram of a 1% agarose gel used to resolve products of affinity cleaving. The cleavage reactions were performed as described under Materials and Methods at pH 6.8. Lane 1 is unreacted *Xmn*I-cut, end-labeled pTGT-IV DNA, and lanes 2–6 show affinity cleavage induced by different oligonucleotides as indicated. Lane M contains markers derived from the digestion of radiolabeled pTGT-IV DNA with *Eco*RI; the size of each fragment is shown on the left.

generated by OP-4 (OP-$T_8G_8A_8$) would have a polarity conflict at the third (3′-most) segment of the target and therefore would be expected to pair only at the two 5′-most segments. OP-5 (OP-G_8AACCCAGT) can pair only at the middle segment, with no strand switch possible.

Triplex formation can be monitored by affinity cleavage, in which an OP-equipped third strand induces site-specific cleavage of the target DNA in the presence of Cu(II) ions under reducing conditions upon triplex formation (Figure 2a) (Sigman & Chen, 1990). Figure 2b shows the results of affinity cleavage induced by oligonucleotides OP-1 through OP-5 on radiolabeled pTGT-IV plasmid DNA in pH 6.8 buffer. OP-1 produces site-specific cleavage of plasmid DNA (lane 2), yielding fragments that comigrate on agarose gels with *Eco*RI fragments of the same DNA (lane M), indicating that triplex formation occurs at or near the 5′-$A_8C_8A_8$-3′ target sequence inserted in the polylinker. The cleavage induced by OP-2 on the same DNA substrate is substantially less than the cleavage induced by OP-1 (compare lanes 2 and 3) and similar to that caused by OP-5 (compare lanes 3 and 6). The triplexes formed by OP-1 and OP-2 by alternate-strand recognition would have identical numbers and types of base triplets (eight T·AT and eight C·GC), differing only in the order in which these triplets are arranged. As mentioned above, previous studies have also reported that the efficiency of RY triplex formation is lower than that of YR triplex (Beal & Dervan, 1992; Jayasena & Johnston, 1992b). The cleavage induced by OP-3 (lane 4) is as effective as that caused by OP-1, whereas the cleavage induced by OP-4 is substantially poorer (lane 5).

To determine whether the cleavage of plasmid DNA occurs precisely at the triplex-forming target sequence as expected, affinity cleavage was carried out on the 5′-end-labeled top strand of a 475-bp *Fsp*I–*Hae*II fragment derived from pTGT-IV and the cleavage products were resolved on a 5% polyacrylamide sequencing gel. As shown in Figure 3a, cleavage occurs at the target site for all oligonucleotides, indicating that it results from triplex formation. Cleavage by OP-1 occurs at the 5′ end of the 5′-A_8C_8-3′ sequence (lane 3), confirming that the orientation of the third strand is as illustrated in Figure 1b. If the cleavage were due to the formation of triplexes without a strand switch, involving just one half-site, two cleavage sites would be expected for OP-1, corresponding to the T_8 block of OP-1 interacting with each of the two adenine tracts of the target to form triplexes stabilized by T·AT triplets. However, the cleavage site expected from the binding of OP-1 at the 5′ A_8 block is far more intense than cleavage at the 3′ A_8 block, which is barely visible, suggesting that binding in this case involves both the 5′ A_8 and the C_8 half-sites.

OP-2 cleaves DNA mainly at the A–C junction of the A_8C_8 sequence, with relatively low efficiency (lane 4). This is the expected cleavage site whether OP-2 utilizes one half-site or both. A secondary cleavage occurs at the other end of the C_8 tract, suggesting that some binding occurs with the G_8 block of OP-2 aligned parallel to the G_8 block of the duplex (and possibly a similar reverse alignment at the 5′ A_8 tract after a YR-type strand crossing). The cleavage induced by OP-5, an oligonucleotide that presumably can bind only at the middle segment via G·GC triplets, is shown in lane 8. The site of cleavage indicates that the interaction occurs with an orientation antiparallel to that of the G_8 tract, as previously seen (Kohwi & Kohwi-Shigematsu, 1988; Pilch et al., 1991; Beal & Dervan, 1991). The efficiencies of cleavage induced by OP-2 and OP-5 are similar, suggesting that OP-2 also binds for the most part only to the middle segment.

Triplex formation at a 5′-$(Py)_n(Pu)_n$-3′ sequence by alternate-strand recognition with a third-strand oligonucleotide containing a special 5′–5′ linkage has been reported to show enhanced stability in the presence of the intercalator ethidium bromide (Ono et al., 1991). There was no apparent effect on the efficiency of cleavage when the affinity cleaving experiment was performed in the presence of ethidium bromide (Figure 3b), suggesting that the enhanced stability in the system studied by Ono et al. could be due to the interaction of ethidium bromide with the special linker or with some feature of the particular sequence used in that study.

The cleavage induced by OP-3 and OP-4 occurs at the 5′ end of the $A_8C_8A_8$ target sequence, as expected (Figure 3a, lanes 5 and 6, respectively). Although both OP-1 and OP-4 have the same sequence available for triplex formation at the target, the efficiency of OP-4-induced cleavage is considerably lower than that of OP-1. The A_8 and T_8 blocks of OP-4 permit it to form a hairpin, thereby potentially reducing the

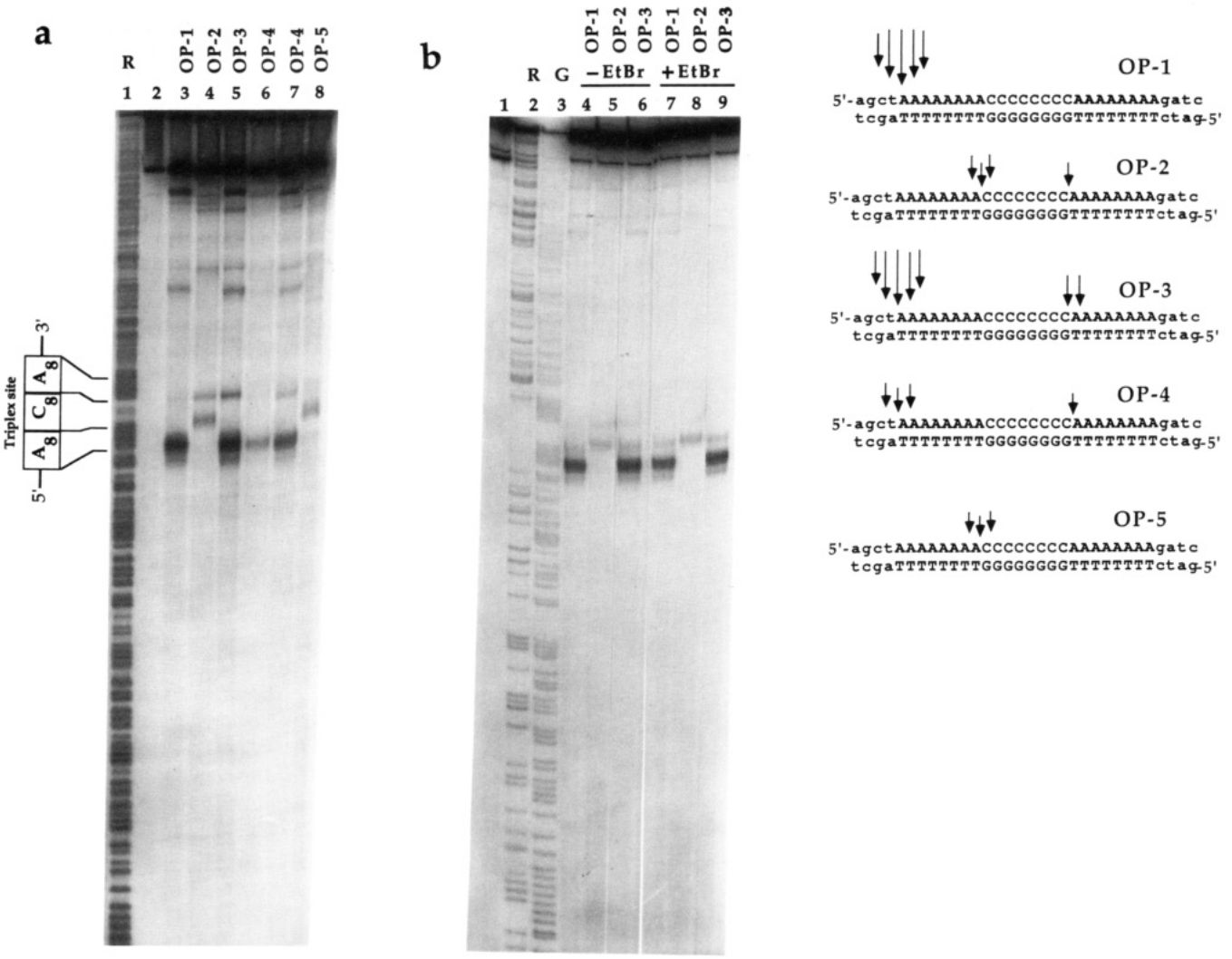

FIGURE 3: Affinity cleavage products of the 5′-labeled top strand of the 475-bp *Fsp*I–*Hae*II fragment of pTGT-IV caused by triplex formation with OP-oligonucleotides, resolved on 5% polyacrylamide sequencing gels. (a) Lane 1, A + G sequencing reaction; lane 2, unreacted DNA; lanes 3–6, cleavage reactions with OP-1 through OP-4 as indicated; lane 7, cleavage by OP-4 that was heated to 90 °C, rapidly chilled, and then complexed with the end-labeled substrate; lane 8, cleavage induced by OP-5. The position of the triplex site is shown to the left of the autoradiogram. (The lighter bands appearing above the main bands in lanes 3–8 appear to be due to low-level labeling of the opposite strand, generating a second set of cleavage sites that parallel the main set.) (b) Affinity cleavage in the presence and absence of 20 μM ethidium bromide. Lane 1, unreacted end-labeled DNA; lanes 2 and 3, A + G and G sequencing reactions. Lanes 4–6 and 7–9 show the cleavage induced by different OP-oligonucleotides as indicated in the absence and presence of ethidium bromide, respectively. (c) Summary of cleavages caused by different OP-oligonucleotides at the target. Cleavage sites are indicated by arrows, and the length of the arrow roughly correlates with the extent of cleavage.

availability of the single-stranded form for triplex formation at the target and possibly accounting for the lower level of OP-4 cleavage compared with that of OP-1. Although hairpin formation is expected to be rapid, heating OP-4 to 90 °C before mixing it with the target DNA resulted in an enhancement of cleavage under these conditions (lane 7).

To identify the triplexes formed at the target sequence in more detail, we used the chemical probes dimethyl sulfate (DMS) and diethyl pyrocarbonate (DEPC). DMS reacts predominantly at the N-7 position of guanines (Singer, 1975). In a triple helix stabilized by G·GC and C^+·GC base triplets, the N-7 positions of guanines in the Watson–Crick duplex are involved in hydrogen bonding with guanine and cytosine residues on the third strand. Consequently, the guanines of the Watson–Crick duplex within a triple helix show resistance to DMS modification (Jayasena & Johnston, 1992a; Johnston, 1988; Hanvey et al., 1988).

The results of DMS modification of the 3′-end-labeled bottom strand of the *Nar*I–*Hae*II DNA fragment of pTGT-IV in the presence of oligonucleotides are shown in Figure 4a. All eight guanines are equally reactive toward DMS in the absence of an oligonucleotide (lane 4). However, when the modification was carried out in the presence of OP-1, except for the single guanine at the 3′ end of the G_8 tract, the remaining seven guanines became resistant to DMS modification (lane 5). The same pattern of DMS reactivity is seen in the presence of OP-3 and (to a lesser extent) OP-4 (lanes 7 and 8). The guanine at the junction is hypermodified [compare with the control lane (lane 4)], suggesting that that base does not participate in a G·GC triplet and, furthermore, that the Watson–Crick duplex is distorted at that position. The DMS reactivity of the same guanine tract in the presence of OP-2 is different; all eight guanines are nearly equally protected, suggesting that all guanines are involved in pairing with the third strand and that the hyperreactive guanine seen in the presence of OP-1, OP-3, and OP-4 results from the YR strand switch.

To investigate the interaction of third-strand oligonucleotides with the two adenine tracts of the target, end-labeled target DNA was probed with DEPC, which carboethoxylates

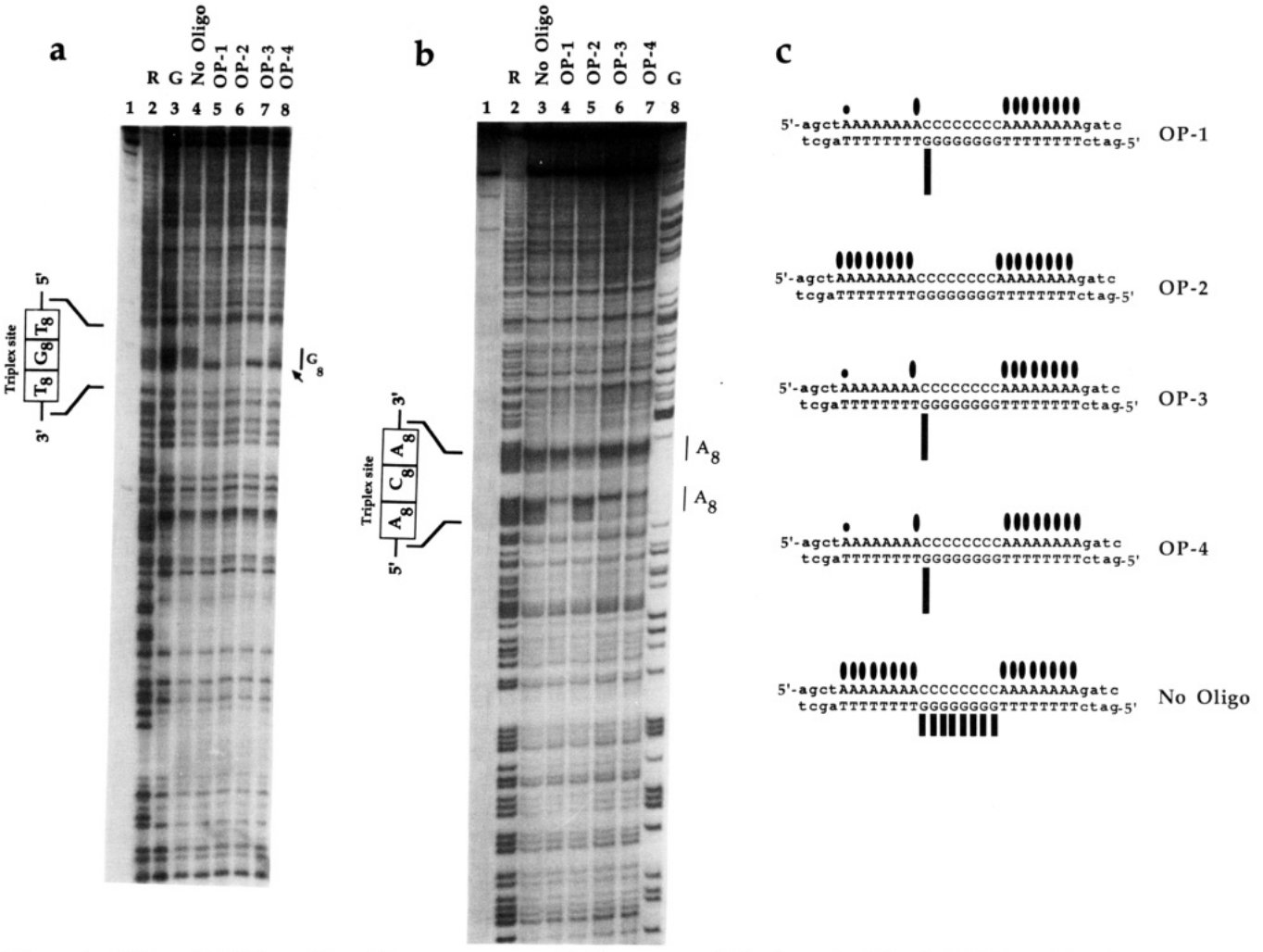

FIGURE 4: DMS and DEPC reactivity of the target sequence in the presence of OP-oligonucleotides. (a) DMS reactivity of the 3′-labeled bottom strand of the 475-bp *Nar*I–*Hae*II fragment of pTGT-IV. Lane 1, unreacted DNA; lanes 2 and 3, A + G and G sequencing reactions; land 4, reactivity toward DMS in the absence of an oligonucleotide but in the triplex-forming buffer; lanes 5–8, DMS modification in the presence of different OP-oligonucleotides as indicated. The arrow indicates the 5′-most guanine within the G8 tract that remains reactive. (b) DEPC reactivity of a 5′-labeled 475-bp *Fsp*I–*Hae*II fragment of pTGT-IV. Lane 1, unreacted DNA; lane 3, DEPC reactivity in the absence of any oligonucleotide; lanes 4–7, DEPC reactivity in the presence of different OP-oligonucleotides as indicated. Lanes 2 and 8 show the products of A + G and G sequencing reactions. On the left of each autoradiogram is the location and the sequence of the triplex-forming site. (c) Summary of DMS and DEPC modification data on the sequence $A_8C_8A_8$. Rectangles show guanines reactive toward DMS; cigar-shaped symbols show adenines reactive toward DEPC.

purines (A > G) primarily at the N-7 position and is a versatile structural probe for DNA (Johnston & Rich, 1985; Herr, 1985). The N-7 positions of Watson–Crick adenines are involved in hydrogen bonding in both T·AT and A·AT base triplets and thus are expected to be resistant to modification with DEPC. Therefore, DEPC is a useful probe for detecting triplex formation at AT tracts (Collier et al., 1991; Jayasena & Johnston, 1992a). The results of DEPC modification of the 5′-end-labeled top strand of the *Fsp*I–*Hae*II DNA fragment of plasmid pTGT-IV are shown in Figure 4b. In the absence of an oligonucleotide, both adenine tracts of the target are approximately equally reactive toward DEPC (lane 3). In the presence of OP-1, the adenines of the 5′ tract are protected but those in the 3′ tract remain reactive except for slight protection of those nearest the C_8 tract (lane 4). The protection of adenines in the 5′-adenine tract from DEPC indicates the formation of T·AT base triplets using the T_8 block of OP-1. This result, together with the DMS modification and affinity cleavage results, confirms that OP-1 forms a YR triplex by alternate-strand recognition as depicted in Figure 1b. Furthermore, the first seven adenines in the 5′ A_8 tract are resistant toward DEPC but the eighth adenine is reactive, suggesting that that base, the last before the strand switch, is not involved in a T·AT base triplet, just as was the case with the adjacent guanine on the other strand as shown by DMS reactivity. Thus, in a YR triplex, the two base pairs surrounding the strand switch appear not to be involved in base triplets and the Watson–Crick duplex appears to be distorted at that position. A similar observation of reactivity at a YR junction was recently reported (Beal & Dervan, 1992).

Potassium permanganate, a chemical that reacts with the 5,6 double bond of unpaired thymines, showed no detectable modification within the target sequence in the presence of oligonucleotide, indicating that the distortion of the helix at the YR junction implied by the DMS and DEPC hyperreactivity was insufficient to render the bottom-strand thymine at the junction reactive (data not shown).

The 5′ A_8 tract is not protected from DEPC by OP-2, as expected (lane 5). In fact, the 3′-most adenine is hyperreactive, consistent with other observations that the 5′-duplex–triplex-3′ boundary is hyperreactive (Collier et al., 1991). However, OP-2 only slightly protects the adenines of the 3′ A_8 tract, suggesting that the OP-2 induced triplex is mainly limited to the G_8 tract, without formation of the entire RY triplex as shown in Figure 1b. The formation of an RY triplex within a mixed sequence studied by Beal and Dervan (1992) required

a 2-nt spacer between the purine and pyrimidine blocks of the third strand. The system studied here was designed so that if a spacer between the guanine tract and the thymine tract of OP-2 were required, it could be recruited by shortening the length of the triplex at either half-site (i.e., by "slippage"). In such a case, we would expect a shorter region of adenines or guanines (or both) to be protected from modifying agents. Slippage at the G_8 tract appears not to occur because the same number of guanines are protected by OP-1, which cannot make an RY strand crossing, as by OP-3, which can (Figure 4a). The slight protection of the 5′ part of the 3′ A_8 block is consistent with slippage at the 3′ A_8 tract because the 3′ part of that tract remains reactive. However, that pattern is also seen for OP-1, which cannot bind at the 3′ A_8 tract. Shortening the 3′ block from 8 to 7 or 6 nucleotides to accommodate a spacer would by itself seem unlikely to preclude triplex formation because YR triplexes consisting of four nucleotides in each half-site (Jayasena & Johnston, 1993), or 3 nucleotides at one half-site and 15 nucleotides at the other have previously been shown to be stable (Froehler et al., 1992).

The DEPC modification patterns in the presence of OP-3 and OP-4 (lanes 6 and 7, respectively) are identical to those in the presence of OP-1, showing protection of the 5′ A_8 tract but no protection of the 3′ A_8 tract (lane 3). This is the expected reactivity pattern for OP-4, because of the unfavorable polarity for A·AT triplet formation in the 3′ A_8 tract with this oligonucleotide. However, OP-3, which has the potential to incorporate all three purine tracts of the target sequence, does not protect the 3′ A_8 tract any more than does OP-1, indicating that OP-3 mainly forms a triplex with only a YR junction and without the additional RY junction as depicted in Figure 1b. The poor ability of OP-3 to extend the triplex to incorporate a YR junction is analogous to the results with OP-2.

Deoxyribonuclease I (DNase I) digestion in the absence of an oligonucleotide reveals only two DNase I-sensitive sites within the entire target sequence, consistent with previous findings for oligo(dA) tracts (Drew & Travers, 1984). Consequently, DNase I protection could not be applied successfully to probe the entire length of the target for triplex formation. However, the DNase I protection seen at the two DNase I-sensitive regions (not shown) is in agreement with the affinity cleaving and chemical probing data.

SUMMARY AND CONCLUSIONS

Data from affinity cleaving (summarized in Figure 2c) and chemical probing experiments (Figure 4c) strongly support the formation of a YR triplex by OP-1 at the 5′-A_8C_8-3′ sequence of the target. At the target examined, OP-1 could potentially form two types of triplexes, one having a strand switch at the 5′-A_8C_8-3′ sequence and the other lacking a strand switch. However, cleavage results largely exclude the latter possibility, indicating that the binding of the YR triplex at th A_8C_8 sequence is significantly stronger than binding at the 3′ A_8 tract and may block the latter binding, despite possible energy costs associated with the strand switch.

The results of DEPC modification (Figure 4c) do not support significant formation of an RY triplex within the target by either OP-2 or OP-3. The data suggest that the low level of cleavage induced by OP-2 is largely due to the formation of an R-type triplex at the G_8 half-site, as was seen with OP-5. Interestingly, a secondary cleavage site suggests that OP-2 may also bind in the reverse orientation to that expected. The cleavage caused by OP-3 is somewhat greater than that caused by OP-1, suggesting that the 3′ T_8 block of OP-3 is transiently interacting with the 3′ adenine tract of the target, but the interaction is not strong enough to significantly protect those adenines from DEPC.

These results contrast with observations of significant RY triplex formation at other sequences (Beal & Dervan, 1992; Jayasena & Johnston, 1992b). Thus, the poor ability of OP-2 and OP-3 to form RY triplexes at the sequence 5′-C_8A_8-3′ appears to be due to a sequence effect. Whether this effect is caused by the C_8 block or the A_n block, or both, would be interesting to discover. Because adenine tracts are known to possess an unusual structure, with bifurcated hydrogen bonds and a high propeller twist (Coll et al., 1987; Nelson et al., 1987; Yoon et al., 1988), the sequence effect might be due to the A_n block. However, the formation of the YR triplex is apparently not affected by any special structure at the A_8 tract.

In conclusion, the use of alternate-strand recognition allows the expansion of potential triple-forming targets to include 5′-$(Pu)_n(Py)_n$-3′ and 5′-$(Py)_n(Pu)_n$-3′ sequences, but certain 5′-$(Py)_n(Pu)_n$-3′ sequences may not be successfully targeted by this approach. An understanding of the structural basis for the sequence dependence of RY triplex stability awaits further study.

REFERENCES

Beal, P. A., & Dervan, P. B. (1991) *Science 251*, 1360–1363.
Beal, P. A., & Dervan, P. B. (1992) *J. Am. Chem. Soc. 114*, 1470–1478.
Blume, S. W., Gee, J. E., Shrestha, K., & Miller, D. M. (1992) *Nucleic Acids Res. 20*, 1777–1784.
Chen, C. B., & Sigman, D. S. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7147–7151.
Coll, M., Federick, C. A., Wang, A. H.-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–8389.
Collier, D. A., Mergny, J.-L., Thuong, N. T., & Hélène, C. (1991) *Nucleic Acids Res. 9*, 4291–4224.
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.
Drew, H. R., & Travers, A. A. (1984) *Cell 37*, 491–502.
Felsenfeld, G., Davies, D. R., & Rich, A. (1957) *J. Am. Chem. Soc. 79*, 2023–2024.
Froehler, B. C., Terhorst, T., Shaw, J.-P., & McCurdy, S. N. (1992) *Biochemistry 31*, 1603–1609.
Griffin, L. C., & Dervan, P. B. (1989) *Science 245*, 967–971.
Hanvey, J. C., Shimizu, M., & Wells, R. D. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6292–6296.
Herr, W. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 8009–8013.
Horne, D. A., & Dervan, P. B. (1990) *J. Am. Chem. Soc. 112*, 2435–2437.
Jayasena, S. D., & Johnston, B. H. (1992a) *Biochemistry 31*, 320–327.
Jayasena, S. D., & Johnston, B. H. (1992b) *Nucleic Acids Res. 20*, 5279–5288.
Jayasena, S. D., & Johnston, B. H. (1992c) *Proc. Natl. Acad. Sci. U.S.A. 89*, 3526–3530..
Johnston, B. H. (1988) *Science 241*, 1800–1804.
Johnston, B. H., & Rich, A. (1985) *Cell 42*, 713–724.
Kohwi, Y., Kohwi-Shigematsu, T. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 3781–3785.
Le Doan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J. L., Thuong, H. T., Lhomme, J., & Hélène, C. (1987) *Nucleic Acids Res. 15*, 7749–7760.
Lee, J. S., Johnson, D. A., & Morgan, A. R. (1979) *Nucleic Acids Res. 6*, 3073–3085.
Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.
Morgan, A. R., & Wells, R. D. (1968) *J. Mol. Biol. 37*, 63–80.

Moser, H., & Dervan, P. B. (1987) *Science 238*, 645–650.
Nelson, H. C. M., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature 330*, 221–226.
Ono, A., Chen, C., & Kan, L. (1991) *Biochemistry 30*, 9914–9921.
Orson, F. M., Thomas, D. W., McShan, W. M., Kessler, D. J., & Hogan, M. E. (1991) *Nucleic Acids Res. 19*, 3435–3441.
Pilch, D. S., Levenson, C., & Shafer, R. H. (1991) *Biochemistry 30*, 6081–6087.
Praseuth, D., Perroualt, L., Le Doan, T., Chassignol, M., Thuong, N., & Hélène, C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1349–1353.
Radhakrishnan, I., de los Santos, C., & Patel, D. J. (1991a) *J. Mol. Biol. 221*, 1403–1418.
Radhakrishnan, I., Gao, X., de los Santos, C., Live, D., & Patel, D. J. (1991b) *Biochemistry 30*, 9022–9030.
Rajagopal, P., & Feigon, J. (1989) *Nature 239*, 637–640.
Sambrook, J., Fritsch, E. F., & Maniatis, T., Eds. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sigman, D. S., & Chen, C. B. (1990) *Annu. Rev. Biochem. 59*, 207–236.
Singer, B. (1975) *Prog. Nucleic Acids Res. Mol. Biol. 15*, 219–284.
Yoon, C., Privé, G. G., Goodsell, D. S., & Dickerson, R. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6332–6336.