Nucleic Acids Research, Vol. 20, No. 20 5279–5288

Oligonucleotide-directed triple helix formation at adjacent oligopurine and oligopyrimidine DNA tracts by alternate strand recognition

Sumedha D.Jayasena+ and Brian H.Johnston*
Cell and Molecular Biology Laboratory, SRI International, 333 Ravenswood Avenue, Menlo Park, CA 94025, USA

Received August 12, 1992; Revised and Accepted September 22, 1992

ABSTRACT

A significant limitation to the practical application of triplex DNA is its requirement for oligopurine tracts in target DNA sequences. The repertoire of triplex-forming sequences can potentially be expanded to adjacent blocks of purines and pyrimidines by allowing the third strand to pair with purines on alternate strands, while maintaining the required strand polarities by combining the two major classes of base triplets, Py·PuPy and Pu·PuPy. The formation of triplex DNA in this fashion requires no unusual bases or backbone linkages on the third strand. This approach has previously been demonstrated for target sequences of the type 5′-(Pu)$_n$(Py)$_n$-3′ in intramolecular complexes. Using affinity cleaving and DNase I footprinting, we show here that intermolecular triplexes can also be formed at both 5′-(Pu)$_n$(Py)$_n$-3′ and 5′-(Py)$_n$(Pu)$_n$-3′ target sequences. However, triplex formation at a 5′-(Py)$_n$(Pu)$_n$-3′ sequence occurs with lower yield. Triplex formation is disfavored, even at acid pH, when a number of contiguous C$^+$·GC base triplets are required. These results suggest that triplex formation via alternate strand recognition at sequences made up of blocks of purines and pyrimidines may be generally feasible.

INTRODUCTION

Following the early discovery of the triple helix structure in synthetic polyribonucleic acids (1), the more recent demonstration of oligonucleotide-directed triple helix formation within natural DNA sequences (2, 3) has shown the potential usefulness of this structure in a number of applications, including therapeutics. Oligonucleotide-directed triplex formation has been used for sequence-specific DNA cleavage (2, 4–6), to block sequence-specific DNA binding proteins (7, 8), and to repress specific gene transcription (9–12). The third strand of a triple helix, which is believed to lie in the major groove of the target DNA duplex, makes hydrogen bonds with purines of the duplex to form specific base triplets. Two different classes of triplexes are known; they are distinguished by their types of base triplets and strand polarity. When the third strand consists mainly of pyrimidines, Hoogsteen-type Py·PuPy base triplets (T·AT and C$^+$·GC) are formed and the third strand is parallel to the purine strand of the target duplex (2, 3) (Figure 1A, top). When the third strand is predominantly purines, Pu·PuPy-type base triplets (G·GC and A·AT) are formed and the third strand is antiparallel to the Watson–Crick purine strand (Figure 1A, bottom) (13–15). [In this motif, thymines on the third strand can also recognize adenines in the duplex by forming T·AT triplets (9, 10, 13), but probably in a reverse Hoogsteen rather than a Hoogsteen orientation. If the third strand consists of guanine and thymine residues, the preferred polarity appears to be determined by the number of GpT and TpG steps (17)]. For brevity, Py·PuPy-type triplexes will be referred to here as Y-type and Pu·PuPy-type triplexes as R-type.

Because the third strand generally recognizes and pairs only with purines, triplex formation with a mixed sequence of purines and pyrimidines would require the third strand to alternate between pairing with the 'Watson' and pairing with the 'Crick' strands. However, the relative polarity between the third strand and the purine-containing strand of the duplex would reverse with each switch. As a result, triplex formation has been generally observed only in sequences that are predominantly or exclusively oligopurine-oligopyrimidine tracts, and its practical applications have been considerably constrained by this sequence restriction.

Several approaches have been described to overcome this limitation and allow targeting of sequences made up of both purines and pyrimidines. Some of them require the synthesis of either unnatural interphosphate linkages (18, 19) or novel base analogs (20, 21). We (22) and others (17, 23) have recently explored another approach which has no special synthetic requirements, and we demonstrated its success in forming intramolecular triplexes at short adjacent blocks of purines and pyrimidines (22). The third strand (consisting of both purine and pyrimidine blocks) pairs with purines in the Watson–Crick duplex, switching strands at the junction between the oligopurine and oligopyrimidine blocks. The required polarity is maintained

* To whom correspondence should be addressed

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

by utilizing both Y- and R-type base triplets, which have opposite polarity requirements. This approach is referred to as triplex formation by alternate strand recognition.

Most practical applications require intermolecular triplex formation between an oligonucleotide and a target duplex rather than intramolecular triplex formation within a single DNA strand. Here we demonstrate alternate strand recognition in such intermolecular triplexes. Two different target sequences that belong to the general forms 5′-$(Pu)_n(Py)_n$-3′ and 5′-$(Py)_n(Pu)_n$-3′, differing in the order of their purine and pyrimidine blocks, are examined. Triplex formation at these two types of target sequences by alternate strand recognition is schematically represented in Figure 1B. In both cases, the required polarities of purine and pyrimidine blocks in the third strand are maintained in the triple helix. These two different triplexes can be distinguished by the order in which the two types of base triplets are arranged. The triplex shown in the upper part of Figure 1B has the orientation 5′-(Py·PuPy)-(Pu·PuPy)-3′ with respect to the third strand and is therefore designated a YR triplex, and that shown in the lower part of Figure 1B has the opposite orientation, 5′-(Pu·PuPy)-(Py·PuPy)-3′, and is referred to as an RY triplex.

By using defined target sequences, we present evidence for intermolecular triplex formation by alternate strand recognition at both types of triplexes, although, for the sequences examined, RY-type triplexes are formed less efficiently than YR-type triplexes.

MATERIALS AND METHODS

Materials

Nucleoside phosphoramidites and other general chemicals for DNA synthesis were purchased from Applied Biosystems, C6-Thiolmodifier from Clonetech, and T4 polynucleotide kinase from United States Biochemicals. Restriction enzymes were obtained from New England Biolabs, deoxyribonuclease I (DNase I) from Promega, 5-nitro-1,10-phenanthroline from G.F. Smith (Columbus, Ohio), and γ-^{32}P-ATP from Du Pont NEN Research Products.

Oligonucleotide synthesis

Oligonucleotides were synthesized on an Applied Biosystems Model 380 A automated DNA synthesizer using β-cyanoethyl phosphoramidite chemistry. After deprotection, the oligonucleotides were purified by electrophoresis on 20% denaturing polyacrylamide gels. 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 according to the method of Chen and Sigman (24).

After solid phase synthesis of oligonucleotides containing 5′ thiol groups, the oligonucleotides were deprotected in ammonia for 8 hr at 55°C, then lyophilized and resuspended in 100 mM triethylammonium acetate (TEAA) (pH 7.5) buffer. Deprotection of the 5′ thiol group was performed according to the manufacturer's instructions. Briefly, 10 A_{260} units of an oligonucleotide in 100 μL of 100 mM TEAA was mixed with 15 μL of 1 M $AgNO_3$ and incubated 30 min at room temperature. 20 μL of 1 M dithiothreitol (DTT) was then added, and the mixture was vortexed and incubated for a further 15 min. The mixture was centrifuged for 15 min at 1500 *g* and the supernatant recovered from the yellowish-white precipitate. The precipitate was washed once with 100 μL of TEAA buffer and the supernatants were combined. DTT in the supernatant was then removed by extracting four times with an equal volume of ethyl acetate. The oligonucleotide with the reactive 5′ thiol group (in ~200 μL) was immediately mixed with 10 μL of a saturated solution of 5-iodoacetamido-1,10-phenanthroline in dimethyl formamide. The coupling reaction was conducted at 4°C overnight, and unreacted 5-iodoacetamido-1,10-phenanthroline was removed by gel filtration through a Sephadex G-50 matrix equilibrated in TE buffer. Finally, OP-modified oligonucleotides were purified by electrophoresis on 20% denaturing polyacrylamide gels.

Plasmid construction

The following complementary oligonucleotide pairs carrying potential triple helix forming motifs flanked by Hind III and BamHI linkers were synthesized:

I 5′-AGCTAAAAAAAGAATTCCCCCCCC-3′
3′-TTTTTTTCTTAAGGGGGGGGCTAG-5′

II 5′-GATCCCCCCCCCTTAAGAAAAAAA-3′
3′-GGGGGGGGAATTCTTTTTTTTCGA-5′

III 5′-GATCCCCCCCCCTTTTCTTTTTTT-3′
3′-GGGGGGGGAAAAGAAAAAAATCGA-5′

Equimolar amounts of gel-purified complementary oligonucleotides were annealed together and the resulting duplex (1 μg) was ligated to pUC18 (0.5 μg) DNA previously digested with Hind III and BamHI. *E.coli* DH5 cells were transformed with ligated products and the cells grown on X-gal- and IPTG-coated, ampicillin-containing LB plates. Transformants were identified by blue-white colony screening, and successful transformants were further characterized by restriction and sequence analysis. Recombinant plasmids were purified from large-scale cultures by using Qiagen plasmid purification kits according to the manufacturer's instructions.

Duplex pair I was used to obtain plasmid pTGT-I. As shown in Figure 2, the successful recombinants (pTGT-II) resulting from the cloning of the duplex pair II sequence carried two copies of Target II sequences (separated by a single GC base pair) inserted at the BamHI site of the polylinker. The recombinant plasmid pTGT-III (derived from the cloning of duplex pair III) had the Target III sequence inserted at the BamHI site.

Radiolabeled DNA substrates

Recombinant plasmids carrying triple helix-forming motifs were linearized with XmnI, dephosphorylated with calf intestine phosphatase, and reacted with γ-^{32}P-ATP and T4 polynucleotide kinase to obtain full-length radiolabeled plasmid DNA. To obtain uniquely end-labeled restriction fragments, plasmid DNA was initially digested with FspI, dephosphorylated, and then precipitated with ethanol. The FspI-digested DNA was then cleaved with HaeII. The restriction fragments were resolved on 1% agarose gels and the 465 base pair (bp) fragment was isolated, purified, and radiolabeled at the dephosphorylated FspI end with γ-^{32}P-ATP using T4 polynucleotide kinase.

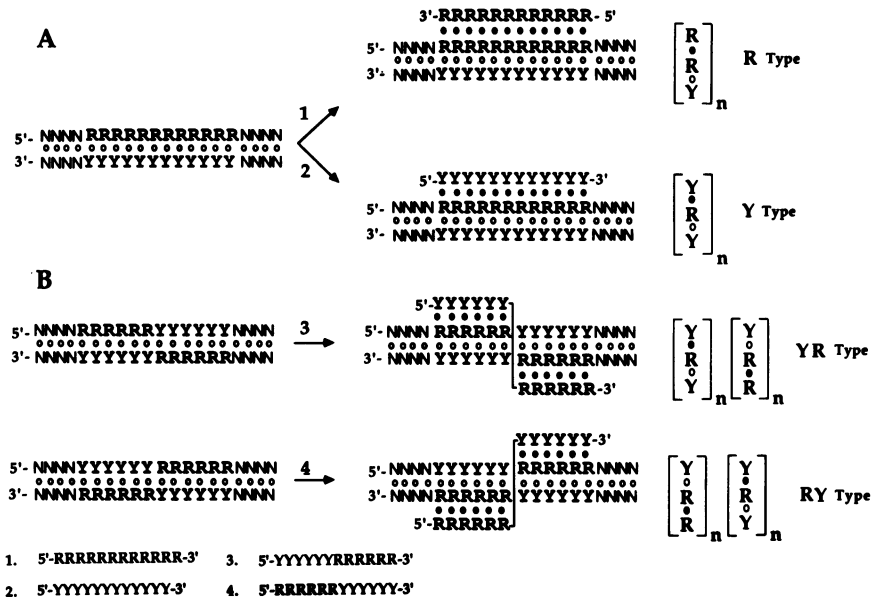

Figure 1. Schematic views of different triple helices. **A.** The top right structure is an R-type triplex formed by an oligopurine strand antiparallel to the purine tract of the target duplex, with the triplex stabilized by Pu·PuPy base triplets. Below it is a Y-type triplex formed by an oligopyrimidine strand parallel to the purine tract of the duplex, with the triplex stabilized by Py·PuPy triplets. **B.** Two types of triplexes formed by alternate strand recognition. The top right structure is a YR-type triplex, formed at a 5′-$(Pu)_n(Py)_n$-3′ target sequence by a third strand oligonucleotide of the form 5′-$(Py)_n(Pu)_n$-3′. Below is an RY-type triplex, formed at a 5′-$(Py)_n(Pu)_n$-3′ target sequence with oligonucleotide of the form 5′-$(Pu)_n(Py)_n$-3′. In both cases, the third strand pairs with the purine tract of one strand of the Watson–Crick duplex via Pu·PuPy base triplets and then crosses over to recognize the purine tract of the alternate strand of the duplex via Py·PuPy base triplets. Open circles represent Watson–Crick hydrogen bonding; closed circles indicate Pu·Pu-type or Hoogsteen hydrogen bonding.

Affinity cleaving

In a typical affinity cleaving experiment, 10 μL of a mixture containing about 10,000 counts/min (cpm) of radiolabeled target DNA (full-length plasmid DNA digested with XmnI and radiolabeled), 3 μM OP-oligonucleotide (oligonucleotide derivatized with 1,10-phenanthroline at the 5′ end), 100 mM NaCl, 50 mM Tris-HCl (pH as specified in the figures), 10 mM $MgCl_2$, 1 mM spermine, 1 μg 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 (25). After further incubation for 12 hr 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 on 1% agarose gels. The gels were dried and visualized by autoradiography.

Efficiency of DNA cleavage

Cleavage efficiencies were quantitated for reactions carried out at pH 6.0, using 3 μM OP-oligonucleotides and radiolabeled full-length plasmid DNA (i.e., XmnI-digested and radiolabeled) and analyzed on 1% agarose gels. The gels were dried and autoradiographed, then all three bands (2.6, 1.8, and 0.79 kilobase DNA fragments) were excised and the radioactivity of each band determined by scintillation counting. The efficiency of cleavage was defined as the sum of the radioactivity of the product bands (1.8 and 0.79 kb) divided by the total radioactivity of all three bands.

DNase I footprinting

DNase I footprinting was performed on end-labeled 465-bp FspI-HaeII DNA fragments in the same buffer used for affinity cleaving (pH 6.0), but without salmon sperm DNA. The mixture of end-labeled DNA (~10,000 cpm) and OP-oligonucleotide (concentration as indicated in figures) in a volume of 10 μL was maintained at 0° for 10 min and then at 20°C for 30 min to allow triple helix formation. Nuclease digestion was initiated by addition of 1 U of DNase I at 20°C. After the indicated time, the reaction was stopped by adding 6 μL of 0.5 M EDTA. The digested DNA was precipitated with ethanol in the presence of 4 μg of carrier tRNA, washed twice with 70% ethanol and resuspended in 80% formamide gel loading buffer. The cleavage products were separated on 5% sequencing gels and the gels visualized by autoradiography.

Chemical sequencing

Maxam–Gilbert chemical sequencing reactions (26) were used to generate G and A+G ladders of end-labeled DNA. For the G-specific reaction, 9 μL of uniquely end-labeled DNA in TE buffer (10 mM Tris-HCl and 0.1 mM EDTA, pH 7.5) were 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 were 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.

RESULTS

Triplex formation monitored by affinity cleaving

Within its polylinker, plasmid pTGT-III carries the Target III sequence consisting of a 20-nucleotide-long oligopurine tract (Figure 2). The triple helix generated at Target III has the same

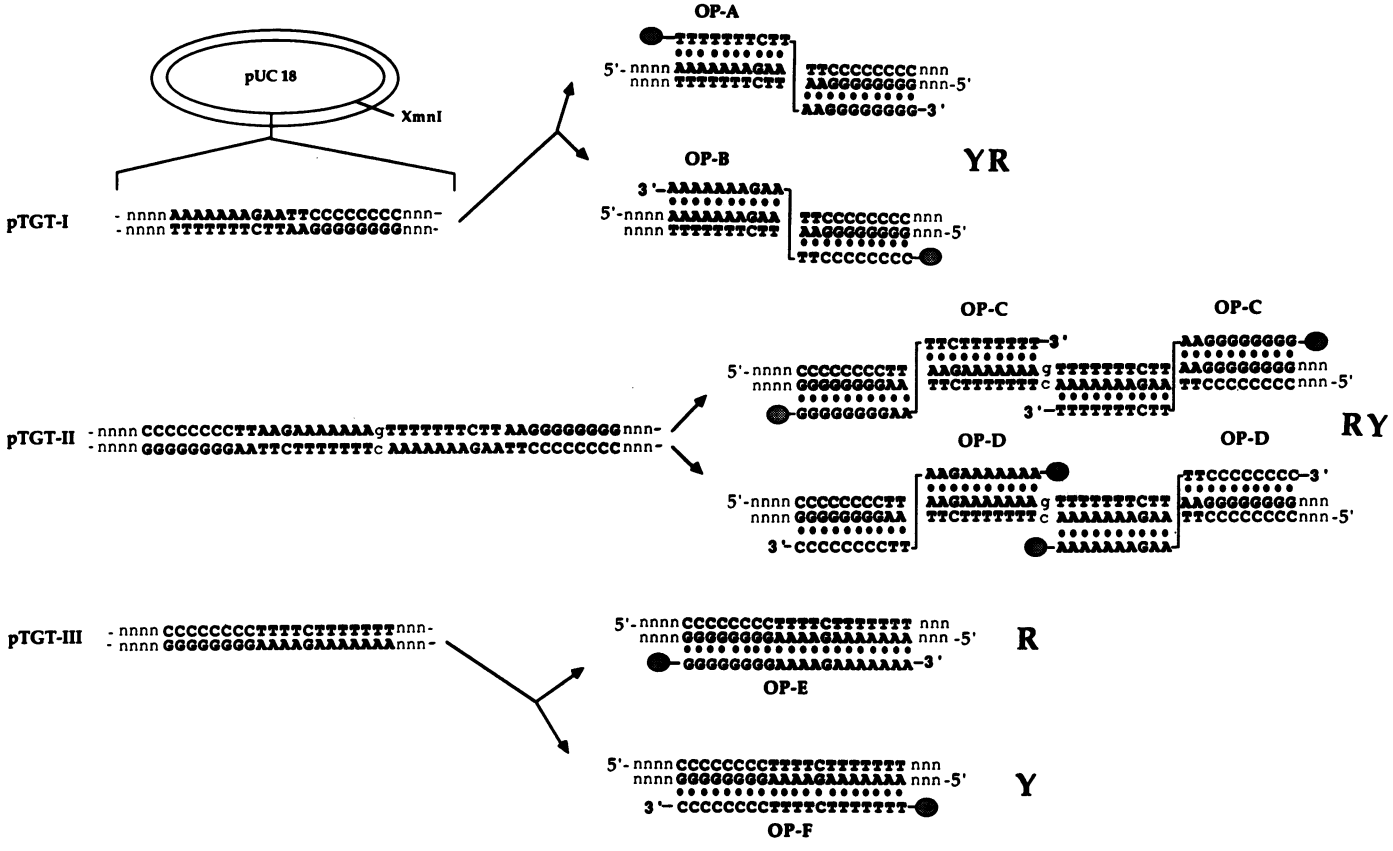

Figure 2. Plasmid constructs and structures of potential triple helices used in this study. pTGT-I, pTGT-II, and pTGT-III plasmids were constructed by inserting Targets I, II and III, respectively, into the polylinker of pUC18. The potential triple helices generated by oligonucleotides OP-A and OP-B at Target I (of the general form 5′-$(Pu)_n(Py)_n$-3′) are of the YR-type (top right). Triplexes generated by oligonucleotides OP-C and OP-D at Target II (of the form 5′-$(Py)_n(Pu)_n$-3′) are of the RY-type (middle right). At the bottom are simple Pu·PuPy (R) and Py·PuPy (Y) triplexes, formed respectively by interaction of OP-E (oligopurine) and OP-F (oligopyrimidine) oligonucleotides with oligopurine Target III. The stippled oval at the 5′ end of each oligonucleotide represents a 1,10-phenanthroline moiety.

sequence composition as those for Targets I and II but does not require a strand switch; therefore, the Target III triplex was used as a control in this study.

Triple helices at Target III can be generated by either an oligopurine strand (OP-E, where OP stands for the 1,10-phenanthroline moiety at the 5′ end of the oligonucleotide) or an oligopyrimidine strand (OP-F) under suitable conditions. As shown in Figure 2 (lower right), both oligonucleotides form hydrogen bonds to the purine strand of the target. Triple helix formation at Target III by OP-E and OP-F was monitored by affinity cleaving (Figure 3A), in which site-specific cleavage of a target DNA occurs when an oligonucleotide equipped with a nonspecific nuclease forms a stable and sequence-specific complex with its target (Figure 3A). In the present case, the nuclease activity is due to the generation of cuprous-oxo species (27) by oligonucleotide-bound 1,10-phenanthroline under reducing conditions.

Lanes 1–3 and 10–12 of Figure 3B show the results of the cleavage of XmnI-cut, radiolabeled pTGT-III DNA by oligonucleotides OP-E and OP-F, respectively, at different pH values. OP-E causes strong cleavage of the target DNA, yielding two fragments—one with 1.8 kilobases (kb) and the other with 0.79 kb—at pH 5.5, 6.0, and 6.8. Comparing these results with the marker lane that contains the same DNA digested with EcoRI, we see that this is the cleavage expected for triplex formation at Target III within the polylinker. Thus, oligonucleotide OP-E forms a stable triplex with Target III at all three pHs. In contrast, OP-F shows a relatively low cleavage rate, and the efficiency of cleavage decreases with increasing pH (data for pH >6.8 not shown). Oligonucleotide OP-E forms Pu·PyPu base triplets whereas oligonucleotide OP-F forms Py·PuPy base triplets, including C^+·GC base triplets, which require acid pH for protonation of cytosines on the third strand (28). Thus, the pH independence with oligonucleotide OP-E and the acid pH dependence with OP-F are as expected. The relatively low rate of cleavage induced by oligonucleotide OP-F at the pH values examined may be due to the presence of a contiguous string of cytosines at the 3′ end of the third strand; the expected increase in triplex formation resulting from protonation of cytosine may be partially canceled by the repulsive effect of the adjacent positive charges (21).

Triplex formation at a 5′-$(Pu)_n(Py)_n$-3′ target site: the YR triplex

To examine triple helix formation by alternate strand recognition, two different target motifs were inserted into pUC18 DNA. Target I in plasmid pTGT-I has the general form 5′-$(Pu)_n(Py)_n$-3′. To maintain correct strand polarities in the complex, stable triple helix formation at 5′-$(Pu)_n(Py)_n$-3′-type targets requires a third-strand oligonucleotide of the form

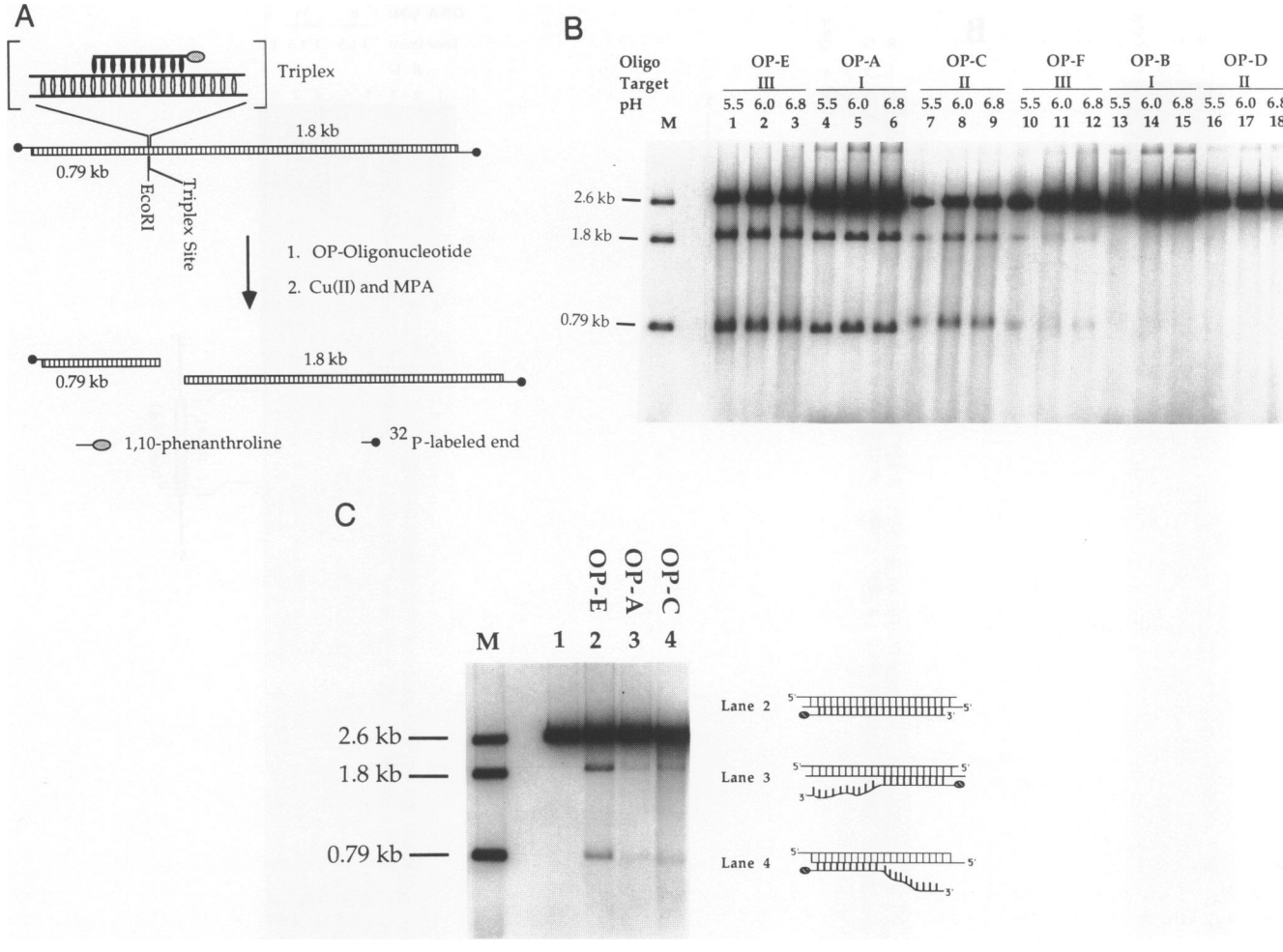

Figure 3. A. Schematic representation of affinity cleavage of pTGT plasmid DNA. Site-specific cleavage of plasmid DNA labeled at both ends is induced by the formation of a triple helix with an appropriate oligonucleotide equipped with a 1,10-phenanthroline moiety in the presence of Cu(II) ions and mercaptopropionic acid (MPA), cleaving the plasmid into two DNA fragments 1.8 and 0.79 kb long. Open ovals represent Watson–Crick base pairs, and filled ovals represent Pu·Pu-type or Hoogsteen pairing. **B.** Autoradiogram of a 1% agarose gel used to resolve products of affinity cleaving. The cleavage reactions were carried out as described in 'Materials and Methods' at the pH values indicated. Lanes 1–3 and 10–12 show the cleavage produced by OP-E and OP-F, respectively, on pTGT-III DNA; lanes 4–6 and 13–15, cleavage of pTGT-I DNA by OP-A and OP-B, respectively; lanes 7–9 and 16–18, cleavage of pTGT-II DNA by OP-C and OP-D, respectively. Lane M contains markers derived from the digestion of pTGT-III DNA with EcoRI; the size of each DNA fragment is shown on the left. **C.** Effect of length of triplex on cleavage efficiency. Target III was reacted with oligonucleotides that can pair with either the entire motif (OP-E; lane 2) or half of the motif (OP-A or OP-C; lanes 3 and 4), and cleavage products were resolved on a 1% agarose gel and autoradiographed. Marker lane M is the same as in panel B.

5′-$(Py)_n(Pu)_n$-3′, which would then switch from pairing with one Watson–Crick strand to the other at the Py-Pu junction. For a given target, two such oligonucleotides are possible. Figure 2 (upper right) shows the two potential triplexes that can be generated by oligonucleotides OP-A and OP-B at Target I. The order with respect to the third strand of base triplets within these YR-type triplexes is 5′-(Y·RY)(R·RY)-3′.

Oligonucleotide OP-A produces site-specific cleavage on linearized pTGT-I DNA at all three pHs (5.5, 6.0, and 6.8) examined (lanes 4–6 in Figure 3B), indicating that a stable triple helix is formed. Because the cleaving moiety is attached to the pyrimidine half of OP-A, cleavage could be accomplished without a strand switch if triplex formation at only this half-site were stable. However, cleavage of Target III by OP-A, which can only pair with the pyrimidine half, is much less efficient than cleavage of Target I by OP-A or cleavage of Target III by OP-E (Figure 3C, lanes 2 and 3), indicating that pairing at both half-sites is necessary for efficient cleavage. These results, together with footprinting data described below, demonstrate that alternate strand triplex formation at 5′-$(Pu)_n(Py)_n$-3′ sequences can be successfully applied to intermolecular complexes. However, unlike oligonucleotide OP-A, OP-B does not produce detectable cleavage at Target I at pH 5.5–6.8 (Figure 3B, lanes 13–15). The absence of cleavage at pH 5.5 is surprising because oligonucleotide OP-F, with the same number of contiguous cytosines as OP-B but without the strand switch, was able to cleave the control plasmid (Target III). Evidently the repulsive effect of several adjacent protonated cytosines combined with the energy cost of a strand switch presents a sufficient energy barrier to preclude triplex formation.

Triplex formation at a 5′-$(Py)_n(Pu)_n$-3′ target site: RY triplex

The second category of target sequences that are candidates for triplex formation by alternate strand recognition is 5′-$(Py)_n(Pu)_n$-3′. To test for triplex formation at a target of this type, plasmid pTGT-II was constructed by inserting the

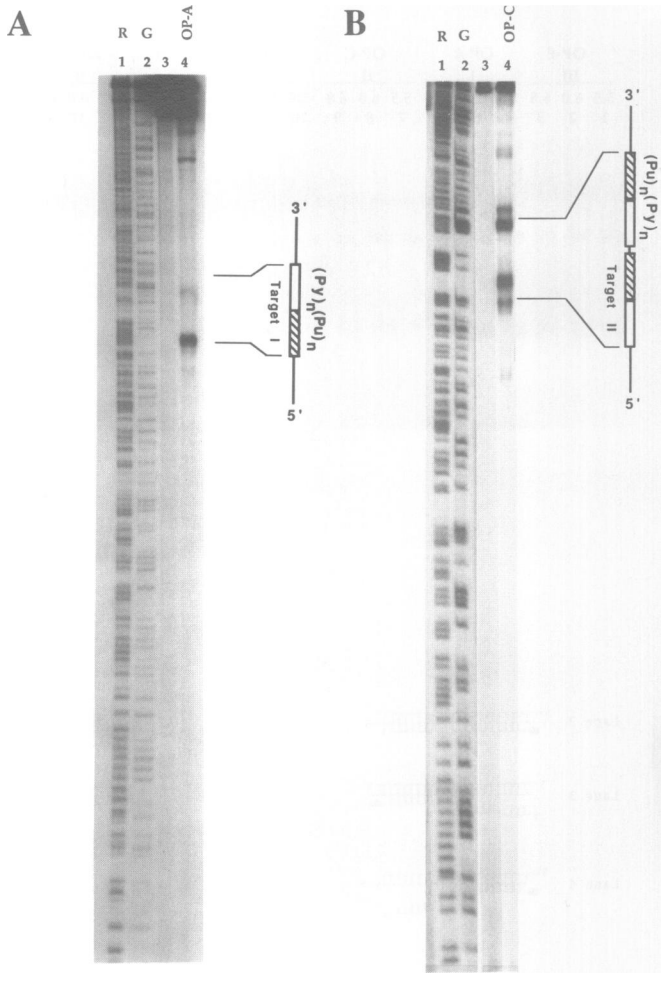

Figure 4. Autoradiograms of 5% polyacrylamide sequencing gels used to resolve products of affinity cleavage of end-labeled Fsp I-Hae II 465-bp fragments of pTGT-I with OP-A (panel A) and pTGT-II with OP-C (panel B). Cleavage was carried out with 3 μM OP-oligonucleotide and ~10,000 cpm end-labeled DNA at 20°C for 12 hr. Lanes 1 and 2 are the products of A+G and G chemical sequencing reactions. Lane 3 shows unreacted DNA. Lane 4 indicates the cleavage induced by OP-A (panel A) and OP-C (panel B). The position and orientation of the purine and pyrimidine blocks of the triple helix site are shown on the right of each panel.

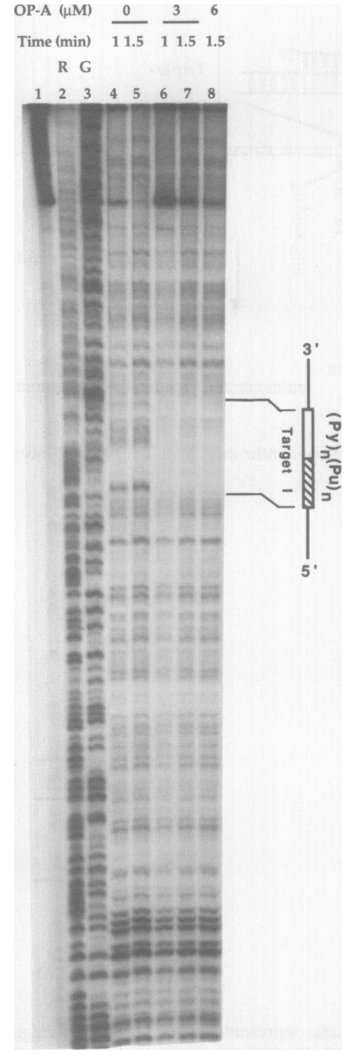

Figure 5. DNase I footprinting of a 465-bp FspI-HaeII fragment of plasmid pTGT-I; autoradiogram of a 5% sequencing gel. Lane 1, untreated DNA; lanes 2 and 3, A+G and G chemical sequencing reactions; lanes 4 and 5, DNase I digestion in the absence of an oligonucleotide; lanes 6–8, DNase I digestion in the presence of OP-A for the indicated concentrations and reaction times.

5′-C_8TTAAGA$_7$-3′ (Target II) sequence. Again, as shown in Figure 2 (middle right), 2 third-strand oligonucleotide sequences can be designed to form triple helices having the required strand polarities. In this case (an RY triplex), the order (with respect to the third strand) of base triplets within the triple helix is 5′-(R·RY)-(Y·RY)-3′.

As shown in lanes 7–9 of Figure 3B, oligonucleotide OP-C efficiently cleaves end-labeled pTGT-II DNA at Target II at all three pH values examined (pH 5.5, 6.0, and 6.8), indicating stable triplex formation at Target II. The inability of OP-C to efficiently cleave Target III, with which it shares the purine sequence to which the cleaving moiety is attached (Figure 3C, lane 4), again indicates that both halves of the sequence must be involved in triplex formation for significant cleavage to be observed. These results, together with footprinting data described below, indicate that the formation of RY as well as YR triplexes can be achieved. OP-D, the other oligonucleotide designed to form a triplex at the same target site, does not show any detectable cleavage at Target II (lanes 16–18 in Figure 3B), a result analogous to the inability of OP-B to cleave Target I. Again, this lack of cleavage could be due to repulsion by the string of eight cytosines that would need to be protonated for triplex formation to occur, combined with the energy cost of a strand switch.

Fine mapping of cleavage sites

The exact sites of oligonucleotide-induced cleavages at Targets I and II were characterized by using uniquely 5′-end-labeled Fsp-I-HaeII DNA fragments of pTGT-I and pTGT-II DNA, with the cleavage products resolved on sequencing gels.

Panel A of Figure 4 shows the results when cleavage products of an end-labeled top strand of the Fsp-I-HaeII fragment of pTGT-I generated by OP-A were analyzed on a 5% polyacrylamide sequencing gel. As seen in Lane 4, oligonucleotide OP-A cleaves the restriction fragment at the site of Target I. The major cleavage site lies at the 5′ end of the purine tract, indicating that the correct orientation of OP-A is indeed as illustrated in Figure 2. The

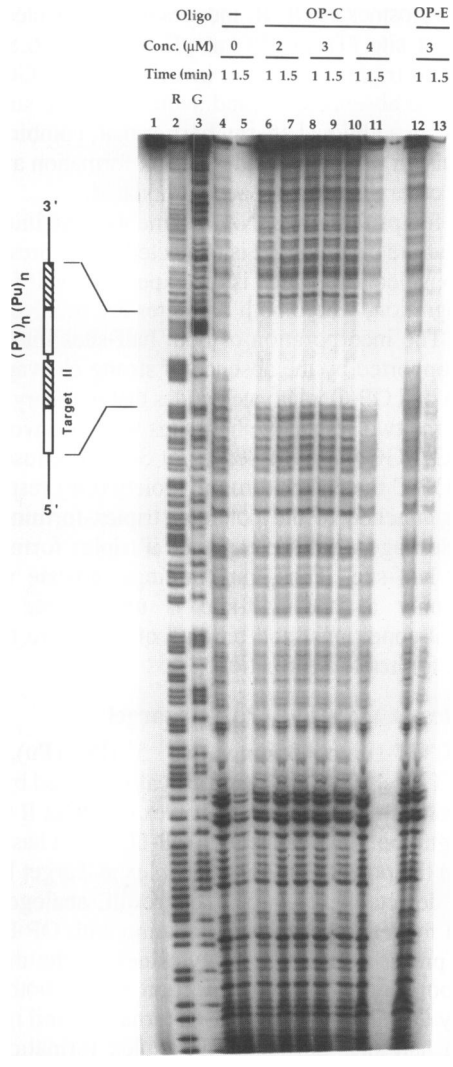

Figure 6. DNase I footprinting of a 465-bp FspI-HaeII fragment of plasmid pTGT-II; autoradiogram of a 5% sequencing gel. Lane 1, untreated DNA; lanes 2 and 3, A+G and G chemical sequencing reactions; lanes 4 and 5, DNase I digestion in the absence of an oligonucleotide; lanes 6–11, DNase I digestion in the presence of OP-C for the indicated concentrations and reaction times; lanes 12 and 13, DNase I digestion in the presence of OP-E.

linkage between the OP moiety and the 5′ end of the third strand includes six C-C bonds and thus may allow cleavage over several adjacent sites.

Panel B of Figure 4 shows cleavage sites induced by oligonucleotide OP-C on the end-labeled top strand of the FspI-HaeII fragment of pTGT-II carrying two Target II sequences. OP-C produces two cleavage sites at each Target II site; for each target site the 5′-most cleavage sites lie at the triplex-duplex junctions, as expected. The other two cleavage sites lie within the 5′ copy of Target II and 3′ to the 3′ copy, respectively, suggesting that additional structures may be present (see Discussion).

Efficiency of oligonucleotide-induced DNA cleavage

The RY triplex formed at Target I by OP-A and the YR triplex formed by OP-C at Target II have the same number of each type of base triplet; the only difference is in the order of the blocks of Pu·PuPy and Py·PuPy base triplets. Therefore, pTGT-I and pTGT-II DNA were used as substrates to assess the cleavage efficiency due to the generation of triplexes at Targets I and II, respectively. The efficiency of cleavage was calculated by scintillation counting of excised agarose gel bands as described in Materials and Methods, above.

The overall efficiency of DNA cleavage induced by triplex formation on these two DNA substrates is the same within experimental error (26% for OP-A + Target I; 24% for OP-C at Target II). Since pTGT-II has two Target II sequence sites, the efficiency of cleavage at each site is ~12%. The efficiency of oligonucleotide-induced cleavage is assumed to be directly proportional to the occupancy of the triplex state, which reflects either the thermodynamic stability or the kinetic availability of the triplex state, or both. Assuming that occupancy of one Target II site by one molecule of OP-C does not interfere with the occupancy of the adjacent Target II site by a second molecule (since there is a 1-base pair spacer between the two sites), the above results indicate that RY triplex formation at a 5′-$(Py)_n(Pu)_n$-3′ sequence is less favorable than YR triplex formation at a 5′-$(Pu)_n(Py)_n$-3′ sequence, even though the base triplet compositions of the two triplexes are identical. However, caution must be exercised in generalizing this conclusion because of possible sequence effects (29) with a given target sequence.

Footprinting

Evidence presented above from affinity cleavage experiments indicates that pairing at both half-sites is necessary for cleavage by OP-A and OP-C at Targets I and II, respectively. To further confirm this requirement, DNase I footprinting was carried out in the presence and absence of the same oligonucleotides. Figure 5 shows the results of DNase I footprinting of DNA carrying Target I. At 3 μM concentration, oligonucleotide OP-A protects the entire length of Target I, indicating that OP-A utilizes both half-sites in triplex formation.

The results of DNase I footprinting of Target II (Figure 6) are analogous to those for Target I. In this case, the end-labeled FspI-HaeII restriction fragment carries two copies of Target II and therefore a larger footprint is expected. Oligonucleotide OP-C (lanes 6–11) protects both copies of Target II over the concentration range (2–4 μM) tried. No protection was seen with OP-E, which in theory can utilize only one half-site of Target II (compare lanes 4 and 12). OP-E does, however, protect Target III along its entire length, as expected (data not shown). These results indicate that YR-type triplex formation utilizes both half-sites of Target II.

Effect of pH

Triplex formation by alternate strand recognition occurs readily at low pH, as shown in Figure 3B. To investigate further the effect of pH, especially in the physiological pH range, affinity cleaving was carried out at pH 7.0–8.0. Triplex formation at Target III, which does not require alternate strand recognition, was used as a control. As shown in Figure 7 (lanes 1–4), the cleavage induced by OP-E at Target III decreases sharply when the pH is above 7.2. This is a surprising result, especially for a triple helix stabilized entirely by Pu·PuPy base triplets. In contrast, the cleavage induced by OP-A at Target I (lanes 5–8) showed a gradual decrease as the pH of the medium was increased from 7.0 to 8.0. Taking the single C^+·GC base triplet within the triplex into consideration, this result is expected. As the pH increases, the protonation of the cytosine on the third strand

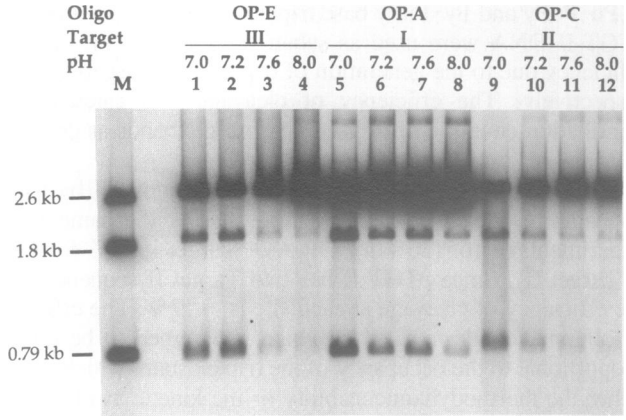

Figure 7. pH dependence of affinity cleavage of pTGT-I, pTGT-II, and pTGT-III DNA over the pH range 7.0−8.0. In this autoradiogram of a 1% agarose gel, the target DNAs, third-strand oligonucleotides, and pH values are as indicated. Lane M is a marker DNA derived from digestion of pTGT-III DNA with EcoRI; the size of each DNA fragment is shown on the left.

becomes less effective and thus may decrease the stability of the triplex. Similarly, cleavage induced by OP-C at Target II, which also depends on formation of a $C^+ \cdot GC$ triplet, shows a decrease in efficiency with increasing pH (lanes 9−12).

DISCUSSION

Alternate strand recognition in triple helical DNA

Triple helical DNA is stabilized by two major classes of natural base triplets, Hoogsteen-type Py·PuPy, and non-Hoogsteen triplets including Pu·PuPy. Hoogsteen-type Py·PuPy base triplets are formed by an oligopyrimidine strand parallel to the purine tract of the target, and Pu·PuPy base triplets are formed by an oligopurine strand antiparallel to the purine tract of the target (Figure 1A). As illustrated in Figure 1B, by maintaining the required strand polarities, these two types of base triplets can be combined to target alternate oligopurine and oligopyrimidine tracts in duplex DNA. We earlier demonstrated the success of this approach in intramolecular triple helix formation (22), and have here extended the application of alternate strand recognition to intermolecular triplexes.

YR triplexes at a 5′-$(Pu)_n(Py)_n$-3′ target

Affinity cleaving of plasmid pTGT-I shows that oligonucleotide OP-A, which has the general form 5′-$(Py)_n(Pu)_n$-3′, induces site-specific cleavage at Target I. This result indicates the formation of a stable triple helix at Target I. However, oligonucleotide OP-B, the second oligonucleotide that has the potential to form a triplex at the same target site, does not elicit cleavage, at least for pH ≥ 5.5. The difference between the two triplexes generated by these two oligonucleotides is the type of base triplets involved. At Target I, OP-A forms eight G·GC, one $C^+ \cdot GC$, two A·AT, and nine T·AT base triplets, whereas OP-B must form one G·GC, eight $C^+ \cdot GC$, ten A·AT, and two T·AT base triplets. Of these triplets, $C^+ \cdot GC$ is known to have pH-dependent stability (28, 30). Thus it seems likely that the inability of OP-B to generate a stable triplex at Target I under the conditions tried is due to the need to form several $C^+ \cdot GC$ triplets that would result in unfavorable charge repulsion (20, 21). However, oligonucleotide OP-F, with the same number of contiguous cytosines as OP-B, induces detectable cleavage at the control target site (Target III) at pH values of 6.8 or lower, indicating that triplex formation involving 8 $C^+ \cdot GC$ triplets is possible in the absence of strand switching. The strand switch may present an additional energy barrier that, combined with the charge repulsion factor, precludes triplex formation at oligo (dC) blocks but otherwise can be accommodated.

DNase I footprinting of DNA fragments containing Target I indicates that the entire target is protected in the presence of the OP-A oligonucleotide, which is the expected result for alternate strand recognition, where both half-sites are involved in triplex formation. The incorporation of both half-sites into the triplex is further supported by the absence of strong cleavage of target III by OP-A and OP-C, oligonucleotides that in theory could form triplexes at the two respective half-sites without involving more than one $C^+ \cdot GC$ base triplet (Fig. 3C). Because for both OP-A and OP-C the phenanthroline moiety (OP) responsible for cleavage is attached to the potential triplex-forming half-site, favorable cleavage would be expected if triplex formation could occur at a half-site. Thus, our findings indicate that triplex formation only at one half-site is unfavorable under the experimental conditions, and binding of the entire third strand is required for triplex formation.

RY triplexes at a 5′-$(Py)_n(Pu)_n$-3′ target

Plasmid pTGT-II, containing two 5′-$(Py)_n(Pu)_n$-3′ target sequences (Target II), was site-specifically cleaved by OP-C, an oligonucleotide designed to form a triplex at Target II by alternate strand recognition. Oligonucleotide OP-D, which has the correct polarity and the potential to form a triplex at Target II, does not induce any detectable cleavage. This result, analogous to what is observed for YR-type triplex formation with OP-B, could be due to the presence of a run of cytosines on the third strand. DNase I footprinting shows protection within both halves of Target II by OP-C, consistent with alternate strand recognition, where both half-sites participate in triplex formation. Control oligonucleotides that can pair only at one half-site confer no DNase I protection on Target II. The results of fine mapping indicate two cleavage sites, one site occurring at the triplex-duplex junction and the other lying within the triplex. The origin of this latter cleavage site is unclear. A structure can be drawn in which the 3′-most 12 bases of OP-C are paired in such a way as to place the phenanthroline group in the region where cleavage is seen, near the middle of the 5′ copy of Target II, but this structure involves 9 reverse-Hoogsteen T·AT triplets (see ref. 17). We have observed another case of secondary binding in apparently the reverse orientation from that expected (29), and affinity cleaving experiments by others (23) using Fe(II)-EDTA attached to the 5′ end of a third strand showed minor cleavage sites within the triplex region for a YR triplex.

Efficiency of triplex formation

The YR triplex formed by OP-A at Target I and the RY triplex generated by OP-C at Target II have identical compositions with respect to the types of nucleotides and base triplets. Taking the efficiency of cleavage as a direct measure of the efficiency of triplex formation at a given target, it is calculated that the YR-type triplexes are formed about twice as efficiently as the RY-type triplexes.

After these experiments were completed, Beal and Dervan, in a similar study of alternate strand recognition (23), reported that they were unable to efficiently make RY triplexes without

a 2-nucleotide spacer between the purine and pyrimidine blocks of the third strand. They reasoned that the spacer was required to bridge the distance between two purine blocks in a 5′-$(Py)_n(Pu)_n$-3′ target but none was needed to bridge the shorter distance within a 5′-$(Pu)_n(Py)_n$-3′ target sequence. In contrast, a YR triplex can be formed at Target II without any spacer nucleotides between the purine and pyrimidine blocks of oligonucleotide OP-C. One possible explanation for the discrepancy between our findings and theirs is slippage of the base pairing at Target II. However, slippage of two nucleotides at either half-site would shorten the triplex region at that half-site from 10 to 5 or 6 nucleotides (for slippage at the AT-rich or GC-rich half-sites respectively) due to mismatches at the junction. Slippage could not account for the cleavage site within the 5′-most copy of Target II (Figure 4B, lane 4) because it lies too far from the triplex-duplex junction.

Alternatively, sequence effects could account for the apparently different requirements for a spacer in this study and in the Beal and Dervan study. The conformation at the junction may be heavily influenced by the sequence immediately surrounding the junction. Moreover, because homopurine strings often have distinctive conformations, including A-DNA (31) and the so-called P-DNA formed by oligo dA strings (32), the conformation at the junction may be influenced by the entire oligopurine sequence comprising each half of the triplex. Some conformations may permit RY-type alternate-strand triplex formation without any spacer, while others may require one. The YR triplex studied by Beal and Dervan had one $m^5C^+ \cdot GC$ and one G·GC base triplet at the strand switch, whereas the strand switch of the YR triplex in our study includes one T·AT and one A·AT base triplet. Model building studies (A.Liu and B.H.J., unpublished) suggest that the distance the third-strand backbone must traverse at the RY junction formed by Target II is greater than that for the YR junction at Target I, but both can be accommodated without a spacer. However, this distance (from C4′ to C4′ on the third strand) across the RY junction appears to be somewhat greater for the Beal and Dervan sequence than for the Target II sequence, thus possibly explaining the difference in requirement for a spacer. Since RY triplex formation at Target II is apparently less efficient than YR triplex formation at Target I, it is possible that, even for this sequence, the addition of a spacer could improve the efficiency of triplex formation.

Effect of pH

As this study reveals, the presence of a strand switch within a triplex does not seem to exert any pH-dependent destabilizing effect on the triplex. The stability of such a triplex seems to be governed predominantly by the type of base triplets within the triplex, as is the case for a triplex without a strand switch.

As is evident from affinity cleaving at different pH values, the triple helix formed at Target III seems to be unstable for pH > 7.2, an unexpected observation for a triple helix consisting exclusively of Pu·PuPy base triplets that do not possess clear-cut pH-dependent elements involved in hydrogen bonding (as is the case for the $C^+ \cdot GC$ triplet). In contrast, RY and YR triplexes formed at Targets I and II by OP-A and OP-C, respectively, which both have one C+·GC base triplet and are therefore expected to be less stable at higher pH, show a more modest pH dependence than does the Target III-OP-E triplex. The Target III-OP-E triplex consists of 9 A·AT and 11 G·GC base triplets, whereas the other two consist primarily of T·AT and G·GC triplets. Triple helices entirely stabilized by G·GC base triplets have been detected at higher pHs (pH 8.0) (33, 34). In a separate study, we have observed that both YR and RY triplexes consisting solely of T·AT and G·GC base triplets are stable even at pH 8.0 (unpublished data). Thus the high-pH instability apprears to be associated with the presence of A·AT triplets, although it is not clear why this should be the case. If the observed pH-dependent stability of the A·AT base triplet holds true for other sequences, it may be wise to favor T·AT and G·GC base triplets wherever possible in designing third-strand sequences.

CONCLUSIONS

Alternate strand recognition opens up a new class of target sequences not previously available for triplex formation with 'natural' oligonucleotides. This category includes target sequences of the general form 5′-$(Pu)_m(Py)_n$-3′ and 5′-(Py)m$(Pu)_n$-3′ (m not necessarily equal to n). YR triplexes generated at 5′-$(Pu)_m(Py)_n$-3′ sequences appear to be more stable than RY triplexes formed at 5′-$(Py)_m(Pu)_n$-3′ sequences; further studies are needed to determine whether conditions other than insertion of a spacer can help stabilize RY triplexes. Contiguous strings of $C^+ \cdot GC$ triplets are unfavorable in triplex formation, especially for triplexes generated by alternate strand recognition. Since for each target sequence there are two potential third-strand sequence choices (e.g., OP-A and OP-B), the one resulting in the minimum number of $C^+ \cdot GC$ base triplets can be selected.

Thus far, strong interactions of nucleotide bases (both purines and pyrimidines) on the third strand of a triplex have been detected only with the purine base of the Watson–Crick base pair (A·AT, G·GC, T·AT, and $C^+ \cdot GC$), with the exception of the G·TA base triplet (35). Therefore, unless unnatural base analogs are synthesized which recognize specific pyrimidine bases of the Watson–Crick duplex in the major groove, or reach across the groove to pair with specific purine bases without requiring the backbone to cross the groove (20), alternate strand recognition may be the best candidate for expanding the repertoire of triplex-forming sequences.

ACKNOWLEDGMENTS

We thank Amy Liu for computer modeling studies and help in preparing figures. This work was supported by NIH grant 1R29GM41423 and the SRI Internal Research and Development Program (Project No. 391D32SEO).

REFERENCES

1. Felsenfeld, G., Davies, D. R. and Rich, A. (1957) *J. Am. Chem. Soc.* 79, 2023–2024.
2. Moser, H. and Dervan, P. B. (1987) *Science* 238, 645–650.
3. Le Doan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J. L., Thuong, N. T. , Lhomme, J. and Hélène, C. (1987) *Nucleic Acids Res.* 15, 7749–7760.
4. Perrouault, L., Asseline, U., Rivalle, C., Thuong, N. T., Bisagni, E., Giovannangeli, C. and Le Doan, T. (1990) *Nature* 344, 358–360.
5. Strobel, S. A. and Dervan, P. B. (1990) *Science* 249, 73–75.
6. Strobel, S. A., Doucette-Stamm, L. A, Riba, L., Housman, D. E. and Dervan, P. B. (1991) *Science* **254**, 1639–1642.
7. Maher, L. J. III, Wold, B. and Dervan, P. B. (1989) *Science* 245, 725–730.
8. Maher, L. J. III, Dervan, P. B. and Wold, B. (1992) *Biochemistry* **31**, 70–81.
9. Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J. and Hogan, M. E. (1988) *Science* 241, 456–459.
10. Orson, F. M., Thomas, D. W., McShan, W. M., Kessler, D. J. and Hogan, M. E. (1991) *Nucleic Acids Res.* 19, 3435–3441.

11. Duval-Valentine, G., Thuong, N. T. and Hélène, C. (1992) *Proc. Natl. Acad. Sci. USA* 89, 504–508.
12. Young, S. L., Krawczyk, S. H., Matteucci, M. D. and Tool, J. J. (1991) *Proc. Natl. Acad. Sci. USA* 88, 10023–10026.
13. Beal, P. A. and Dervan, P. B. (1991) *Science* 251, 1360–1363.
14. Pilch, D. S., Levenson, C. and Shafer, R. H. (1991) *Biochemistry* 30, 6081–6087.
15. Radhakrishnan, I, de los Santos, C. and Patel, D. J. (1991) *J. Mol. Biol.* **221**, 1403–1418.
16. Ono, A., Chen, C. and Kan, L. (1991) *Biochemistry* **30**, 9914–9921.
17. Sun, J.-S., De Bizemont, T., Duval-Valentin, T., Montenay-Garestier, T., and Hélène, C. (1991) Comptes Rend. Acad. Sci. Paris, t. **313**, Série III, p. 585–590.
18. Horne, D. A. and Dervan, P. B. (1990) *J. Am. Chem. Soc.* 112, 2435–2437.
19. Froehler, B. C., Terhorst, T., Shaw, J. -P. and McCurdy, S. N. (1992) *Biochemistry* **31**, 1603–1609.
20. Kiessling, L. J., Griffin, L. C. and Dervan, P. B. (1992) *Biochemistry* **31**, 2829–2834.
21. Koh, J. S. and Dervan, P. B. (1992) *J. Am. Chem. Soc.* 114, 1470–1478.
22. Jayasena, S. D. and Johnston, B. J. (1992) *Biochemistry* **31**, 320–327.
23. Beal, P. A. and Dervan, P. B. (1992) *J. Am. Chem. Soc.* **114**, 4976–4982.
24. Chen, C. B.and Sigman, D. S. (1986) *Proc. Natl. Acad. Sci. USA* 83, 7147–7151.
25. Jayasena, S. D. and Johnston, B. J. (1992) *Proc. Natl. Acad. Sci. USA* 89, 3526–3530.
26. Maxam, A. M.and Gilbert, W. (1980) *Methods Enzymol.* 65, 499–560.
27. Sigman, D. S. and Chen, C. B. (1990) *Annu. Rev. Biochem.* **59**, 207–236.
28. Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A. and Frank-Kamenetskii, M. D. (1987) *Nature* 330, 495–497.
29. Jayasena, S. D., and Johnston, B. H., submitted.
30. Povsic, T. J. and Dervan, P. B. (1989) *J. Am. Chem. Soc.* **111**, 3059–3061.
31. Peticolas, W. L., Wang, T. and Thomas, G. A. (1988) Proc. Natl. Acad. Sci. USA **85**, 2579–2583.
32. Coll, M., Frederick, C.A., Wang, A.H.-J., and Rich, A. (1987) Proc. Natl. Acad. Sci. USA **84**, 8385–8389.
33. Kohwi, Y. and Kohwi-Shigematsu, T. (1988) *Proc. Natl. Acad. Sci. USA* 85, 3781–3785.
34. Lyamichev, V., Voloshin, O. N., Frank-Kamenetskii, M. D. and Soyfer, V. N. (1991) *Nucleic Acids Res.* **19**, 1633–1638.
35. Griffin, L. C. and Dervan, P. B. (1989) *Science* 245, 967–971.