AD-A251 499

‖‖‖‖‖‖‖‖‖‖‖‖‖

OFFICE OF NAVAL RESEARCH

Grant N00014-92-J1052

R&T Code 4135018

Technical Report #5

"Recognition of Double Helical DNA by Alternate Strand Triple Helix Formation"

by
P. A. Beal and P. B. Dervan

California Institute of Technology
Division of Chemistry and Chemical Engineering
Pasadena, CA

June 1, 1992

92-14891

‖‖‖‖‖‖‖‖‖‖‖‖‖‖

92 6 0 01ö

Recognition of Double Helical DNA by Alternate Strand Triple Helix Formation

Peter A. Beal and Peter B. Dervan*

Arnold and Mabel Beckman Laboratories of Chemical Synthesis
California Institute of Technology
Pasadena, California 91125

Abstract:

The triplet specificities and required strand orientations of two classes of DNA triple helices can be combined to target double helical sequences containing all four base pairs by alternate strand triple helix formation. This allows for the use of oligonucleotides containing only natural 3'-5' phosphodiester linkages to simultaneously bind both strands of double helical DNA in the major groove. The stabilities and structures of these alternate strand triple helices depend on whether the binding site sequence is 5'-(purine)$_m$(pyrimidine)$_n$-3' or 5'-(pyrimidine)$_m$(purine)$_n$-3'. The sequence type 5'-(purine)$_m$(pyrimidine)$_n$-3' was targeted with an oligonucleotide consisting of a pyrimidine domain and a purine domain linked by a 3'-5' phosphodiester. To bind the duplex sequence type 5'-(pyrimidine)$_m$(purine)$_n$-3', the third strand requires at least two nucleotides linking binding domains at the site of crossover in the major groove. Since both of these sequence types have been successfully targeted, a new class of oligonucleotides capable of binding a variety of duplex sequences by multiple crossovers in the major groove may now be possible.

---

*To whom correspondence should be addressed

1

Introduction

Oligonucleotide-directed triple helix formation is one of the most versatile among the methods for the sequence specific recognition of double helical DNA.[1-4] This approach has recently been used to mediate single site specific cleavage of human chromosomal DNA[1k] as well as to block transcription *in vitro*.[1m,3a] The generalizability to target a broad range of DNA sequences and the high stabilities of the resulting local triple helical structures makes this a powerful technique for binding single sites within large segments of double helical DNA.

At least two classes of DNA triple helices exist which differ in the sequence composition of the third strand, the relative orientations and postitions of the three strands and the base triplet interactions.[1-4] Pyrimidine oligonucleotides bind in the major groove specifically to purine tracts of double helical DNA parallel to the Watson-Crick (WC) purine strand. Specificity is derived from thymine (T) recognition of adenine-thymine base pairs (T•AT base triplets) and protonated cytosine (C+) recognition of guanine-cytosine base pairs (C+GC base triplets).[1,2,4a-e] Within this structural motif, guanine (G) has been shown to recognize certain thymine-adenine base pairs depending on the flanking sequence (G•TA base triplets).[1f] An additional family of triple helical structures consists of purine-rich oligonucleotides bound in the major groove to purine tracts of double helical DNA antiparallel to the WC purine strand.[3,4f] Sequence specificity is derived from G recogonition of GC base pairs (G•GC base triplets) and adenine (A) recognition of AT base pairs (A•AT base triplets). Within this structural motif, T has been shown to recognize AT base pairs (reverse Hoogsteen T•AT base triplets).

In these two triple helical structures, sequence specific binding occurs by the formation of hydrogen bonds between bases in the third strand and duplex base pairs, predominantly at the purines of the base pairs. Oligonucleotides designed to bind to purine target sequences allow for hydrogen bonding to the purines in

2
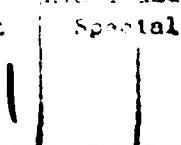
consecutive base pairs without significant distortion of the sugar phosphate backbone, maximizing stability by allowing stacking of bases in the third strand. Therefore, oligonucleotide directed triple helix formation appears to be limited to mostly purine tracts. A formidable challenge in the sequence specific recognition of duplex DNA by triple helix formation is designing oligonucleotides capable of binding mixed purine/pyrimidine sequences. In this work, we show that the specificities of two classes of DNA triple helices can be used in tandem to target mixed sequences of duplex DNA by alternate strand triple helix formation.

Results and Discussion

Alternate Strand Triple Helix Formation Recently, it has been shown that pyrimidine oligonucleotide-directed triple helix formation can be extended to target mixed DNA sequences of the type $5'-(Pu)_m(Py)_n-3'$ and $5'-(Py)_m(Pu)_n-3'$ by simultaneous binding to adjacent purine tracts on alternate strands of the WC duplex.[5] To preserve the required strand orientation, pyrimidine oligonucleotides were linked 3' to 3' or 5' to 5'. However, since in the two triple helix motifs, the third strands bind in opposite orientations with respect to the purine strand of the duplex, it seems plausible that an oligonucleotide consisting of pyrimidines and purines with natural 3'-5' phosphodiesters might bind in the major groove and simultaneously recognize both strands. Base triplet specificities and required strand orientations would be maintained by alternating between structural motifs at junctions between purines and pyrimidines on one strand of the duplex.(Figure 1) The base triplets of such a structure are shown in Figure 2, where thymidine and 5-methyldeoxycytidine are bound to one strand of the duplex and deoxyadenosine, deoxyguanosine and thymidine to the other.

3

The two duplex sequences shown in Figures 3 and 8 were prepared, each containing 18 bp binding sites composed of two adjacent nine bp half sites on alternate strands capable of being bound in a triple helical sense by either pyrimidine or purine oligonucleotides. The sequence content is identical in both binding sites, however they are opposite in polarity such that the two sequence types 5'-$(Pu)_m(Py)_n$-3' and 5'-$(Py)_m(Pu)_n$-3' can be compared. Modelling of these two sequences suggests that the adjacent triple helix binding sites are in different relative positions in space. Therefore, designing triple helix forming oligonucleotides to bind to these two sequence types will likely require different structures at the junction where the transition from one motif to the other takes place.

The 5'-$(Pu)_m(Py)_n$-3 sequence The potential binding site, 5'-AAGAAAAAGCTC CTCCCT-3', shown in Figure 3, consists of two adjacent nine bp purine tracts on alternate strands with a 5'-GC-3' junction. Oligonucleotides **1-5** were designed to bind this sequence by the formation of a pyr•pur•pyr triple helix at the 5' half site and/or a pur•pur•pyr triple helix at the 3' half site. Oligonucleotides **1-4** were modified with thymidine-EDTA so that the affinity cleaving method could be used to monitor their binding properties.[6]

Under conditions in which very little binding is detected for the nine-mer controls, **3** and **4**, alone or in combination, both termini of the 18-mer oligonucleotide are bound to the duplex as indicated by cleavage generated by 5' thymidine-EDTA in **1** and 3' thymidine-EDTA in **2**.(Figure 4) In data not shown, a 16 mer oligonucleotide lacking the center two nucleotides cleaves the target equally as efficiently as **1**.

DNase I footprinting was used to confirm the binding site size for **5**, identical in sequence to **1** and **2**, lacking thymidine-EDTA. (Figure 5) As expected from the affinity cleaving results, **5** protects both strands of the duplex from cleavage by

4

DNase I over a region of approximately 25 base pairs centered around the proposed binding site. DNase I footprint titrations with **5** yielded an association constant K_a = 5.7± 1.4 x 10^6 M^{-1} (pH=7.0, 1.5 mM spermine, 5 mM NaCl, 10 mM $MgCl_2$, 10 mM $CaCl_2$) corresponding to a binding free energy ΔG_b = -9.2± 0.1 kcal/mol.[7]

The reactivity of dimethylsulfate (DMS) can be used to determine the accessibility of the N-7 position of guanine,[8] which is involved in both the MeC+GC and CG•G base triplets.(See Figure 2) Formation of these triplets should protect the guanines in the duplex from methylation by DMS. In the presence of 10 µM **5**, all the guanines in the binding site are fully protected from methylation except the two at the junction, which are only marginally protected. (Figure 6)

The results of these experiments are summarized in Figure 7. These data show that this duplex sequence can be targeted for alternate strand triple helix formation by linking binding domains with a 3'-5' phosphodiester. However, the two bases in the center of the third strand apparently do not stabilize the structure to a great extent and DMS footprinting indicates that the two base pairs at the junction are likely not involved in stable base triplets.

The 5'-(Py)$_m$(Pu)$_n$-3' sequence The target site, 5'-TCCCTCCTCGAAAAAGAA-3', shown in Figure 8, is identical in sequence content to that shown in Figure 3, but opposite in polarity. Therefore, the two adjacent nine-mer purine tracts on alternate strands are joined at a 5'-CG-3' junction. Oligonucleotides **6-12** were designed to bind the target by the formation of a pur•pur•pyr triple helix at the 5' half site and/or a pyr•pur•pyr triple helix at the 3' half site.

In this case, when the two nine-mers are linked by a phosphodiester, there is no significant increase in binding affinity, as evidenced by comparison of the cleavage efficiency for **6** with that for **10** plus **11**.(Figure 9, lanes 5 and 6) However, when the oligonucleotide contains two thymidines between the purine and

pyrimidine domains as in **8** and **9**, both termini are bound to the duplex.[9](Figure 9,

lanes 8 and 9) One thymidine is not sufficient, as shown by the low cleavage

efficiency of **7**.(Figure 9, lane 7) Oligonucleotide **12**, which contains the two

thymidine linker but lacks thymidine-EDTA was used for footprinting studies. In

the presence of 10 μM **12**, both strands of the duplex are protected from DNase I

degradation in the region corresponding to the putative binding site. (Figure 10)

DNase I footprint titrations yielded an association constant $K_a = 1.5 \pm 0.4 \times 10^6 \, M^{-1}$

(pH=7.0, 1.5 mM spermine, 5 mM NaCl, 10 mM $MgCl_2$, 10 mM $CaCl_2$) corresponding

to a binding free energy $\Delta G_b = -8.4 \pm 0.2$ kcal/mol. DMS footprinting studies on this

structure show that all the guanines in the binding site are protected equally from

methylation. (data not shown)

The footprinting reagent MPE•Fe was used to investigate the triplex formed

by **12** at the 5'-$(Py)_m(Pu)_n$-3' target site.[10](Figure 11) In addition to protection from

cleavage within the binding site, a region of increased reactivity at the center of the

5'-$(Py)_m(Pu)_n$-3' site was detected. This single locus of hyperreactivity suggests a

single high affinity intercalation site at the junction.[11]

A summary of the studies for this duplex target is given in Figure 12. These

results show that targeting the 5'-$(Py)_m(Pu)_n$-3' sequence requires at least a two

nucleotide linker between the two binding domains. This is in contrast to the 5'-

$(Pu)_m(Py)_n$-3' sequence, for which no linker nucleotides were necessary and in fact

the two nucleotides at the center likely do not form base triplets. The triplex formed

at the 5'-$(Py)_m(Pu)_n$-3' site is less stable, with an association constant lower by a

factor of four. Also, the MPE•Fe reactivity pattern for this structure indicates that

the base pairs near the site of crossover between strands provide an unusually stable

binding site for an intercalator. Interestingly, Ono *et al.* have proposed an ethidium

bromide binding site at the center of the sequence 5'-$(Py)_8(Pu)_8$-3' when bound on

alternate strands by pyrimidine oligonucleotides linked at their 5'ends.[5c]

6

It is clear that the stucture and stability of the alternate strand triple helix depends on whether the duplex target is $5'\text{-}(\text{Pu})_m(\text{Py})_n\text{-}3'$ or $5'\text{-}(\text{Py})_m(\text{Pu})_n\text{-}3'$. Analysis of models of these sequences with oligonucleotides bound to adjacent binding sites may explain some of these observations. For the sequence $5'\text{-}(\text{Pu})_m(\text{Py})_n\text{-}3'$, the site of crossover between strands is located where adjacent binding sites overlap.(Figure 13, left) Therefore, no linker between binding domains besides a phosphodiester is necessary for a successful crossover. In fact, it is likely that the two base pairs at the junction are not specifically recognized in base triplets, allowing the third strand to maintain helical continuity as it winds along the major groove.

For the sequence $5'\text{-}(\text{Py})_m(\text{Pu})_n\text{-}3'$, the adjacent binding sites do not overlap.(Figure 13, right) A linker domain is necessary for successfully crossing the major groove in this fashion, with a minimum length determined here to be two nucleotide residues. Although thymidines were used for this purpose, further optimization of the linker structure may still be possible. Also, the unusual MPE•Fe reactivity at the center of this triple helix may indicate that the duplex undergoes conformational reorganization upon binding of the third strand.   This might involve a local unwinding or bending at the center of this alternate strand triple helical structure, creating a site for intercalation.

Conclusion   By combining the base triplet specificities and required strand orientations of two distinct classes of DNA triple helices, we have designed oligonucleotides containing natural $3'\text{-}5'$ phosphodiester linkages which bind mixed sequences of duplex DNA by alternate strand triple helix formation. The duplex sequence type $5'\text{-}(\text{Pu})_m(\text{Py})_n\text{-}3'$ was targeted by linking two binding domains with a phosphodiester. However to bind the duplex sequence type $5'\text{-}(\text{Py})_m(\text{Pu})_n\text{-}3'$, the third strand requires at least two linker nucleotides at the site of crossover in the major groove. The design of oligonucleotides containing both pyrimidines and

7

purines and all natural phosphodiester linkages capable of binding a variety of sequence types by multiple crossovers in the major groove may now be possible.

Experimental

General. Distilled, deionized water was used for all aqueous reactions and dilutions. Enzymes were purchased from Stratagene, Boehringer-Mannheim or New England Biolabs. Enzyme reactions were performed using the manufacturer's recommended protocol in the activity buffer provided. pUC19 plasmid DNA was purchased from Sigma. DNase I was obtained from Stratagene and dilutions were made in the activity buffer recommended by the manufacturer. Deoxynucleoside triphosphates were purchased from Pharmacia as 100 mM solutions. $5'$-(α-^{32}P) dGTP (>3000 Ci/mmol) and $5'$-(γ-^{32}P)ATP (>5000 Ci/mmol) were obtained from Amersham. Competent cells were purchased from Stratagene. Transformation, selection and maintanence of cell lines were accomplished using standard procedures.[12] Calf thymus DNA was purchased from Pharmacia. Polyacrylamide gel electrophoresis was performed in 1XTBE buffer.[12] $5'$ and $3'$ end labelling was accomplished using standard procedures.[12] Autoradiography was carried out using Kodak X-Omat film.

Synthesis and Purification of Oligonucleotides. Oligonucleotides were prepared on an Applied Biosystems Model 380B DNA synthesizer with β-cyanoethyl phosphoramidites. Thymidine-EDTA was prepared as described[6] and incorporated at the $3'$ end of oligonucleotides via the $5'$-O-DMT-thymidine-EDTA-triethylester $3'$-succinyl controlled pore glass or at the $5'$ end via the $5'$-O-DMT-thymidine-EDTA-triethylester $3'$-O-(2-cyanoethyl,N,N-diisopropyl) phosphoramidite. Deprotection was carried out in 0.1 N NaOH at 55° C for 24 hours. Oligonucleotides were purified

8

by reverse phase chromatography on a Pharmacia FPLC system using a ProRPC 10/10 (C2-C8) column with a 0-40% CH_3CN gradient in 100 mM triethylammonium acetate, pH=7. The concentration of single-stranded oligonucleotides was determined at 260 nm, using the following molar extinction coefficients for each base: 15400 (A), 11700 (G), 7300 (C), 5700 (Me^5C) 8800 (T and T*) $cm^{-1}M^{-1}$.

Construction of Plasmid DNA. Using T4 DNA ligase, the plasmid pPBCRI was prepared by ligation of the duplex formed between oligonucleotides of the sequence 5'-AATTCTCTCTAAGAAAAAGCTCCTCCCTCTCTCT-3' and 5'-CTAGAGAGAG AGGGAGGAGCTTTTTTCTTAGAGAG-3' into pUC19, which had been previously digested with Eco RI and Xba I. The plasmid pPBCRII was analogously prepared from insertion of the duplex formed between 5'-AATTCTCTCTTCCCTCCTCG AAAAAGAACTCTCT-3' and 5'-CTAGAGAGAGTTCTTTTTCGAGGAGGGA AGAGAG-3'. Ligation products were used to transform Epicurian™ Coli XL 1 Blue competent cells. Colonies were selected for α-complementation on 25 mL Luria-Bertani medium agar plates containing 50 µg/mL ampicillin and treated with XGAL and IPTG solutions.[12] Large scale plasmid purification was performed using Qiagen purification kits according to the manufacturer's protocol. Plasmid DNA concentration was determined at 260 nm using the relation 1 OD unit = 50 µg/mL duplex DNA.

Generation of Labelled Duplex DNA. Labelled restriction fragments from pPBCRI and pPBCRII were generated as follows. Plasmid DNA (20 µg) was linearized with Hind III followed either by calf alkaline phosphatase treatment and 5' end labelling by T4 polynucleotide kinase with γ-^{32}P ATP or 3' end labelling by filling in with the Klenow fragment of DNA polymerase I and dATP, dCTP, dTTP and α-^{32}P dGTP. The labelled plasmid DNA was digested with Ssp I and the 639 bp HindIII/SspI

restriction fragment was isolated by nondenaturing 5% PAGE. Gel bands were visualized by autoradiography, cut out and eluted with 200 mM NaCl. The suspension was filtered to remove polyacrylamide and the DNA was purified by ethanol precipitation.

Affinity Cleaving. In a typical cleavage experiment, the labelled DNA was mixed with salts, buffer, spermine, and the oligonucleotide EDTA•Fe(II). This solution was incubated for one hour at 37° C before the cleavage reaction was initiated by the addition of DTT giving a final volume of 20 µL. Cleavage reactions were allowed to proceed for 18 hours at 37° C. The reactions were stopped by precipitation with ethanol. Electrophoretic separation of the cleavage products was achieved on a denaturing 8% polyacrylamide gel. The gel was transferred to filter paper and dried for one hour at 80° C. Cleavage products were visualized by autoradiography and the relative amounts of cleavage products were analyzed by storage phosphor autoradiography. Specific reaction conditions are stated in the figure legends.

Storage Phosphor Autoradiography. Storage phospor imaging plates (Kodak Storage Phosphor Screen S0230 obtained from Molecular Dynamics) were pressed flat against gel samples and exposed in the dark for 12-17 hours.[13] A Molecular Dynamics 400S PhosphorImager was used to obtain all data from the storage screens. The data were analyzed by performing volume integrations of the regions corresponding to intact DNA, the target site and reference sites using the ImageQuant v. 3.0 software running on an AST Premium 386/33 computer.

DNase I Footprinting. In a typical DNase I footprinting experiment, the labelled DNA was mixed with salts, buffer, spermine, and the oligonucleotide. This solution was incubated for two hours at 24° C before the addition of nonspecific

oligonucleotide and DNase I giving a final volume of 20 μL. The nonspecific single stranded oligonucleotide 5'-ATCTGATGAGGTGCTGAATAGGACC -3' was included in all DNase I footprinting reactions to maintain uniform DNase I activity in the presence and absence of triple helix forming oligonucleotides. For footprint titrations, the incubation period was extended to 24 hours to insure samples containing low concentrations of triple helix forming oligonucleotides had reached equilibrium. The DNase I reactions were allowed to proceed for 10 minutes at 24°C and stopped by the addition of calf thymus DNA and EDTA, followed by precipitation with ethanol. Electrophoretic separation of the cleavage products was achieved on an 8% polyacrylamide gel. The gel was transferred to filter paper and dried for one hour at 80° C. DNase I cleavage products were visualized by autoradiography and analyzed by storage phosphor autoradiography . Specific reaction conditions are stated in the figure legends. Data analysis for the footprint titrations was performed according to previously described methods.[7] Curves were fitted to the experimental data using KaleidaGraph v. 2.1 (Abelbeck Software) running on a Macintosh IIfx personal computer, with the upper and lower end points of the curve and K_a as adjustable parameters. The association constants are reported in the text as the mean ± one standard deviation for three experiments.

Dimethylsulfate Footprinting. In a typical DMS footprinting experiment, the labelled DNA was mixed with salts, buffer, spermine, and the oligonucleotide. This solution was incubated for two hours at 24° C before the addition dimethylsulfate giving a final volume of 20 μL. The reactions were allowed to proceed for 90 seconds at 24°C, stopped as described by Maxam and Gilbert, and the DNA was precipitated with ethanol.[13] The reaction products were redissolved in 10% aqueous piperidine and heated at 90°C for 30 minutes. The solutions were lyophilized to dryness. Electrophoretic separation of the cleavage products was achieved on an 8%

polyacrylamide gel. The gel was transferred to filter paper and dried for one hour at 80° C. Cleavage products were visualized by autoradiography and analyzed by storage phosphor autoradiography. Specific reaction conditions are stated in the figure legends.

MPE•Fe(II) Footprinting. In a typical MPE•Fe(II) footprinting experiment, the labelled DNA was mixed with salts, buffer, spermine, and the oligonucleotide. This solution was incubated for two hours at 24°C before the addition of MPE•Fe(II) and DTT to a final volume of 20 μL. The reactions were allowed to proceed for five minutes at 24° C and stopped by precipitation with ethanol. Electrophoretic separation of the cleavage products was achieved on an 8% polyacrylamide gel. The gel was transferred to filter paper and dried for one hour at 80° C. Cleavage products were visualized by autoradiography and analyzed by storage phosphor autoradiography. Specific reaction conditions are stated in the figure legends.

References

1. (a) Moser, H. E.; Dervan, P.B. *Science* **1987**, 238, 645. (b) Strobel, S. A.;
    Moser, H.E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, 110, 7927. (c) Povsic,
    T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, 111, 3059. (d) Strobel, S. A.;
    Dervan, P. B. *J. Am. Chem. Soc.* **1989**, 111, 7286. (e) Maher, L. J.; Wold, B.;
    Dervan, P. B. *Science* **1989**, 245, 725. (f) Griffin, L. C.; Dervan, P. B.
    Science, **1989**, 245. 967. (g) Strobel, S. A.; Dervan, P. B. *Science* **1990**, 249, 73. (h)
    Plum, G.E.; Park, Y.W.; Singleton, S.F.; Dervan, P.B.; Breslauer, K.T.; *Proc.*
    Natl. Acad. Sci. USA **1990**, 87, 9436. (i) Maher, L.J.; Dervan, P.B.; Wold, B.J.;
    Biochemistry, **1990**, 29, 8820. (j) Strobel, S.A.; Dervan, P.B.; *Nature* **1991**,
    350, 172. (k) Strobel, S.A.; Doucette-Stamm, L.A.; Riba, L.; Housman, D.E.;
    Dervan, P.B.; *Science* **1991**, 254, 1639. (l) Distefano, M.D.; Shin, J.A.; Dervan,
    P.B.; *J. Am. Chem. Soc.* **1991**, 113, 5901. (m) Maher, L.J.; Dervan, P.B.; Wold, B.;
    Biochemistry **1992**, 31, 70.

2. (a) Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.L.;
    Thoung, N.T.; Lhomme, J.; Helene, C. *Nucleic Acids Res.* **1987**, 15, 7749. (b)
    Praseuth, D.; Perrouault, L.; Le Doan, T.; Chassignol, M.; Thuong, N. T.;
    Lhomme, J.; Helene, C. *Proc. Natl. Acad. Sci. USA* **1988**, 85, 1349. (c) Francois,
    J. C.; Saison-Behmoaras, T.; Chassignol, M.; Thuong, N. T.; Helene, C. *J.*
    Biol. Chem. **1989**, 264, 5891. (d) Lyamichev, V. I.; Mirkin, S. M.; Frank-
    Kamenetskii, M. D.; Cantor, C. R. *Nucleic Acids Res.* **1988**, 16, 2165. (e)
    Francois, J. C.; Saison-Behmoaras, T.; Thuong, N. T.; Helene, C.
    Biochemistry **1989**, 28, 9617. (f) Sun, J. S.; Francois, J. C.; Montenay-Garestier,
    T.; Saison-Behmoaras, T.; Roig, V.; Thuong , N. T.; Helene, C. *Proc. Natl.*
    Acad. Sci. USA **1989**, 86, 9198. (g) Sun, J.S.; Giovannangeli, C.; Francois, J.C.;

Kurfurst, R.; Montenay-Garestier, T.; Asseline, U.; Saison-Behmoaras, T.; Thuong, N.T.; Helene, C.; *Proc. Natl. Acad. Sci. USA* **1991**, 88, 6023.

3.   (a) Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, J.; Hogan, M. E. *Science* **1988**, 241, 456. (b) Beal, P.A.; Dervan, P.B.; *Science* **1991**, 251, 1360. (c) Pilch, D.S.; Levenson, C.; Shafer, R.H.; *Biochemistry* **1991**, 30, 6081. (d) Durland, R.H.; Kessler, D.J.; Gunnell, S.; Duvic, M.; Pettitt, B.M.; Hogan, M.E., *Biochemistry,* **1991**, 30, 9246.

4.   (a) Rajogopal, P.; Feigon, J. *Nature* **1989**, 339, 637. (b) Rajagopal, P.; Feigon, J.; *Biochemistry* **1989**, 28, 7859. (c) Sklenar, V.; Feigon, J.; *Nature* **1990**, 345, 836. (d) Macaya, R.F.; Schultze, P.; Feigon, J.; *J. Am. Chem. Soc.* **1992**, 114, 781. (e) de los Santos, C.; Rosen, M.; Patel, D.; *Biochemistry* **1989**, 28, 7282. (f) Radhakrishnan, I., de los Santos, C., Patel, D.J., *J. Mol. Biol.,* **1991**, 221, 1403.

5.   (a) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, 112, 2435. (b) McCurdy, S., Moulds, C., Froehler, B., *Nucleosides & Nucleotides,* **1991**, 10, 287-290. (c) Ono, A., Chen, C., Kan, L., *Biochemistry,* **1991**, 30, 9914-9921.

6.   Dreyer, G. B.; Dervan, P. B.; *Proc. Natl. Acad. Sci. USA* **1985**, 82, 968.

7.   Brenowitz, M., Senear, D., Shea, M.A., Ackers, G.K., *Proc. Natl. Acad. Sci. USA,* **1986**, 83, 8462.

8.   Williamson, J.R., Raghuraman, M.K., Cech, T.R., *Cell,* **1989**, 59, 871.

9.   The 11 mer 5'-T*GGGAGGAGTT-3' binds the target site less efficiently than 10, indicating that the linker thymidines alone do not stabilize the triple helix.

10.   Hertzberg, R.P., Dervan, P.B., *J. Am. Chem. Soc.,* **1982**, 104, 313.

11.   Collier, D.A., Mergny, J., Thuong, N.T., Helene, C.;*Nucleic Acids Res.,* **1991**, 19, 4219.

12.   Sambrook, J., Fritsch, E.F., Maniatis, T., *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989).

13. Johnston, R.F., Pickett, S.C., Barker, D.L., *Electrophoresis,* **1990,** 11, 355.

14. Maxam, A., Gilbert, W., *Methods in Enzymology,* **1980,** 65, 499.

15. Iverson, B. L.; Dervan, P. B. *Nucleic Acids Res.* **1987,** 15, 7823.

Figure Legends

Figure 1 Models representing oligonucleotides binding in the major groove and simultaneously recognizing both strands of the duplex by forming hydrogen bonds to the purine in a WC base pair. Base triplet specificities and required strand orientations are maintained by alternating between triple helix motifs at junctions between purines and pyrimidines on one strand of the duplex. (Above) Triple helix formed at a $5'-(Pu)_m(Py)_n-3'$ target site. (Below) Triple helix formed at a $5'-(Py)_m(Pu)_n-3'$ target site.

Figure 2 Models for base triplets formed when an oligonucleotide binds in the major groove and simultaneously recognizes both strands. All bases are in the anti conformation.

Figure 3 The duplex target site present on the 639 bp Hind III/Ssp I restriction fragment from plasmid pPBCRI containing two nine-mer purine tracts on alternate strands joined at a 5'-PuPy-3' junction. The double stranded region bound in the triple helix is boxed. The sequences of oligonucleotide-EDTA's 1-5 are shown, where T* indicates the position of thymidine-EDTA and bold type C indicates 5-methyldeoxycytidine.

Figure 4 (Left) Autoradiogram of an 8% denaturing polyacrylamide gel used to separate affinity cleavage products. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (200 nM) and $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ (500 nM) with the 3' end ^{32}P labelled Hind III/Ssp I restriction fragment from plasmid pPBCRI [~15,000 cpm] in a solution of tris acetate, pH=7.0 (50 mM), spermine (1 mM), NaCl (10 mM), and calf thymus DNA (0.1 mM bp), and then incubating for one hour at 37° C. Concentrations listed are final concentrations. The reactions were initiated by the addition of dithiothreitol (DTT) (4 mM) and allowed to proceed for 18 hours at 37° C. The DNA was precipitated with ethanol and the cleavage products were analyzed by gel electrophoresis. (Lane 1) Products of an adenine-specific sequencing reaction.[15] (Lane 2) Intact 3' labelled fragment obtained after incubation under the conditions of the cleavage reactions in the absence of oligonucleotide-EDTA•Fe(II). (Lanes 3-7) DNA cleavage products produced by oligonucleotide-EDTA•Fe(II) 1-4; 3 (lane 3); 4 (lane 4) 3+4(200 nM each) (lane 5); 1 (lane 6); 2 (lane 7). **(Right)** Sequences of triple helical complexes formed between oligonucleotides 1 and 2 and the 5'-(Pu)9(Py)9-3' target site.

Figure 5 Autoradiogram of an 8% denaturing polyacrylamide gel used to separate DNase I footprinting products. The cleavage reactions were carried out by incubating oligonucleotide 5 with the 3' end or 5' end ^{32}P labelled Hind III/Ssp I restriction fragment from plasmid pPBCRI [~20,000 cpm] in a solution of trisHCl, pH=7.0 (40 mM), spermine (1.5 mM), NaCl (5 mM), MgCl2 (10 mM), CaCl2 (10 mM) for two hours at 24° C. The reactions were initiated by the addition of nonspecific single stranded oligonucleotide (1 μM) and DNase I (0.25 units/μL) and allowed to proceed for 10 minutes at 24° C. The reactions were stopped by the addition of calf thymus DNA (0.1 μM bp) and

16

EDTA (50 mM). Concentrations listed are final concentrations. The DNA was precipitated with ethanol and the cleavage products were analyzed by gel electrophoresis. Odd numbered lanes contain 5' end labelled fragment, even numbered lanes contain 3' end labelled fragment. (Lanes 1 and 2) Intact labelled fragment in the absence of oligonucleotide 5 and DNase I. (Lanes 3 and 4) Products of adenine-specific sequencing reactions.[15] (Lanes 5 and 6) DNase I cleavage products in the absence of oligonucleotide 5. (Lanes 7 and 8) DNase I cleavage products in the presence of 10 μM oligonucleotide 5. (Lanes 9 and 10) DNase I cleavage products in the presence of 1 μM oligonucleotide 5.

Figure 6 Autoradiogram of an 8% denaturing polyacrylamide gel used to separate DMS footprinting products. The reactions were carried out by incubating oligonucleotide 5 with the 3' end or 5' end ^{32}P labelled Hind III/Ssp I restriction fragment from plasmid pPBCRI [~20,000 cpm] in a solution of trisHCl, pH=7.0 (40 mM), spermine (1.5 mM), NaCl (5 mM), MgCl$_2$ (10 mM), CaCl$_2$ (10 mM) for two hours at 24°C. The reactions were initiated by the addition of dimethylsulfate (DMS) (0.2% v/v) and allowed to proceed for 90 seconds at 24°C. The reactions were stopped as described by Maxam and Gilbert.[14] The DNA was precipitated with ethanol, redissolved in 50 μL 10% aqueous piperidine and heated at 90° C for 30 min. Concentrations listed are final concentrations. The solutions were lyophilized and the cleavage products were analyzed by gel electrophoresis. Odd numbered lanes contain 5' end labelled fragment, even numbered lanes contain 3' end labelled fragment. (Lanes 1 and 2) Intact labelled fragment in the absence of oligonucleotide 5 and DMS. (Lanes 3 and 4) Products of adenine-specific sequencing reactions.[15] (Lanes 5 and 6) Cleavage products generated in the

17

presence of DMS in the absence of oligonucleotide 5. (Lanes 7 and 8) Cleavage products generated in the presence of DMS and 10 μM oligonucleotide 5. Arrows indicate guanines in the binding site not protected by oligonuceotide 5.

Figure 7 (A) Cleavage pattern generated by oligonucleotide 1. The box indicates the double stranded sequence bound. Positions of the arrows show the sites of cleavage and heights indicate extent of cleavage at that site. (B) Cleavage pattern generated by oligonucleotide 2. (C) DNase I footprint generated by oligonucleotide 5. Bases within the brackets are protected from cleavage by DNase I. (D) Dimethylsulfate (DMS) reactivity within the target site bound by oligonucleotide 5. Asterisks (*) indicate bases within the binding site protected from methylation by DMS. Arrows indicate bases within the binding site which remain reactive to DMS upon binding.

Figure 8 The duplex target site present on the 639 bp Hind III/Ssp I restriction fragment from plasmid pPBCRII containing two nine-mer purine tracts on alternate strands joined at a 5'-PyPu-3' junction. The double stranded region bound in the triple helix is boxed. The sequences of oligonucleotideEDTA's 6-12 are shown, where T* indicates the position of thymidine-EDTA, bold type C indicates 5-methyldeoxycytidine and the underlined thymidines indicate the linker domain.

Figure 9 (Left) Autoradiogram of an 8% denaturing polyacrylamide gel used to separate affinity cleavage products. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (200 nM) and $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ (500 nM) with the 5' end ^{32}P labelled Hind III/Ssp I

18

restriction fragment from plasmid pPBCRII [~15,000 cpm] in a solution of tris acetate, pH=7.0 (50 mM), spermine (1 mM), NaCl (10 mM), and calf thymus DNA (0.1 mM bp) and then incubating for one hour at 37° C. The reactions were initiated by the addition of dithiothreitol (DTT) (4 mM) and allowed to proceed for 18 hours at 37° C. Concentrations listed are final concentrations. The DNA was precipitated with ethanol and the cleavage products were analyzed by gel electrophoresis. (Lane 1) Products of an adenine-specific sequencing reaction.[15] (Lane 2) Intact 5' labelled fragment obtained after incubation under the conditions of the cleavage reactions in the absence of oligonucleotide-EDTA•Fe(II). (Lanes 3-9) DNA cleavage products produced by oligonucleotide-EDTA•Fe(II) 6-11; 10 (lane 3); 11 (lane 4) 10+11 (200 nM each) (lane 5); 6 (lane 6); 7 (lane 7); 8 (lane 8); 9 (lane 9). (Right) Sequences of triple helical complexes formed between oligonucleotides 8 and 9 and the 5'-(Py)$_9$(Pu)$_9$-3' target site.

Figure 10 Autoradiogram of an 8% denaturing polyacrylamide gel used to separate DNase I footprinting products. The cleavage reactions were carried out by incubating oligonucleotide 12 with the 3' end or 5' end ^{32}P labelled Hind III/Ssp I restriction fragment from plasmid pPBCRII [~20,000 cpm] in a solution of trisHCl, pH=7.0 (40 mM), spermine (1.5 mM), NaCl (5 mM), MgCl$_2$ (10 mM), CaCl$_2$ (10 mM) for two hours at 24°C. The reactions were initiated by the addition of nonspecific single stranded oligonucleotide (1 µM) and DNase I (0.25 units/µL) and allowed to proceed for 10 minutes at 24°C. The reactions were stopped by the addition of calf thymus DNA (0.1 µM bp) and EDTA (50 mM). Concentrations listed are final concentrations. The DNA was precipitated with ethanol and the cleavage products were analyzed by gel electrophoresis. Odd numbered lanes contain 5' end labelled fragment, even

19

numbered lanes contain 3′ end labelled fragment. (Lanes 1 and 2) Intact labelled fragment in the absence of oligonucleotide **12** and DNase I. (Lanes 3 and 4) Products of adenine-specific sequencing reactions.[15] (Lanes 5 and 6) DNase I cleavage products in the absence of oligonucleotide **12**. (Lanes 7 and 8) DNase I cleavage products in the presence of 10 µM oligonucleotide **12**. (Lanes 9 and 10) DNase I cleavage products in the presence of 1 µM oligonucleotide **12**.

Figure 11 Autoradiogram of an 8% denaturing polyacrylamide gel used to separate MPE•Fe footprinting products. The cleavage reactions were carried out by incubating oligonucleotide **12** with the 3′ end or 5′ end ^{32}P labelled Hind III/Ssp I restriction fragment from plasmid pPBCRII [~20,000 cpm] in a solution of trisHCl, pH=7.0 (40 mM), spermine (1.5 mM), NaCl (5 mM), MgCl₂ (10 mM), CaCl₂ (10 mM) for two hours at 24°C. The reactions were initiated by the addition of MPE•Fe (5 µM) and DTT (4 mM) and allowed to proceed for five minutes at 24°C. Concentrations listed are final concentrations. The reactions were stopped by precipitation with ethanol and the cleavage products were analyzed by gel electrophoresis. Odd numbered lanes contain 5′ end labelled fragment, even numbered lanes contain 3′ end labelled fragment. (Lanes 1 and 2) Intact labelled fragment in the absence of oligonucleotide **12** and MPE•Fe. (Lanes 3 and 4) Products of adenine-specific sequencing reactions.[15] (Lanes 5 and 6) MPE•Fe cleavage products in the absence of oligonucleotide **12**. (Lanes 7 and 8) MPE•Fe cleavage products in the presence of 10 µM oligonucleotide **12**. Bracket indicates bases in the binding site hyperreactive to MPE•Fe.

20

Figure 12 (A) Cleavage pattern generated by oligonucleotide **8**. The box indicates the double stranded sequence bound. Positions of the arrows show the sites of cleavage and heights indicate extent of cleavage at that site. **(B)** Cleavage pattern generated by oligonucleotide **9**. **(C)** DNase I footprint generated by oligonucleotide **12**. Bases within the brackets are protected from cleavage by DNase I. **(D)** MPE•Fe(II) reactivity within the target site bound by oligonucleotide **12**. Arrows indicate positions of increased cleavage upon binding. Arrow heights indicate extent of cleavage at that site.

Figure 13 **(Left)** Ribbon model depicting nine-mer oligonucleotides binding to adjacent triple helix binding sites for the sequence 5'-(Pu)$_9$(Py)$_9$-3'. The 3' end of one binding site appears to overlap with the 5' end of the adjacent site in the major groove. **(Right)** Ribbon model depicting nine mer oligonucleotides binding to adjacent triple helix binding sites for the sequence 5'-(Py)$_9$(Pu)$_9$-3'. No binding site overlap is apparent in the major groove.

5'-(Pu)$_m$(Py)$_n$-3'

junction

5'- T———MeC⁺———G———A -3'   - Third Strand

5'- A———G———C———T -3' ⎫ Watson-Crick
3'- T———C———G———A -5' ⎬ Duplex

5'-(Py)$_m$(Pu)$_n$-3'

junction

5'- A———G———MeC⁺———T -3'   - Third Strand

5'- T———C———G———A -3' ⎫ Watson-Crick
3'- A———G———C———T -5' ⎬ Duplex

Fig 1

T·AT

MeC+GC

TA·A

CG·G

TA·T

Fig 2

639 bp

^{32}P

5'- CTCTCT AAGAAAAAGCTCCTCCCT CTCTCT -3'
3'- GAGAGA TTCTTTTTCGAGGAGGGA GAGAGA -5'
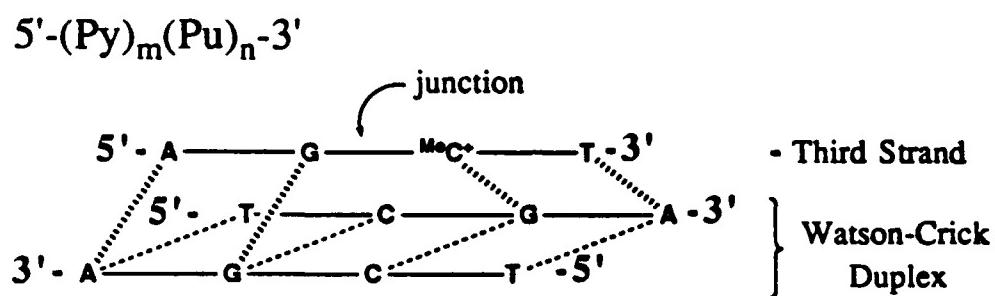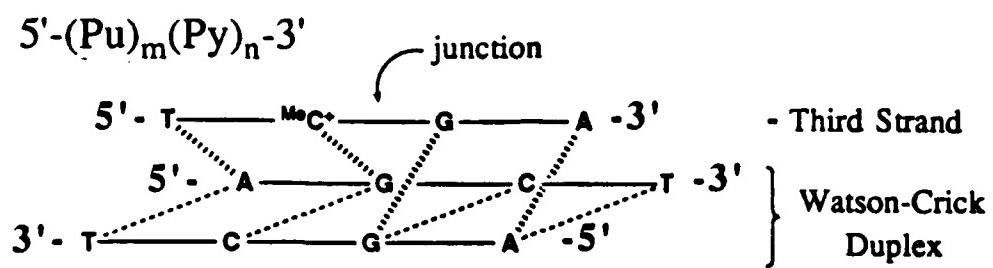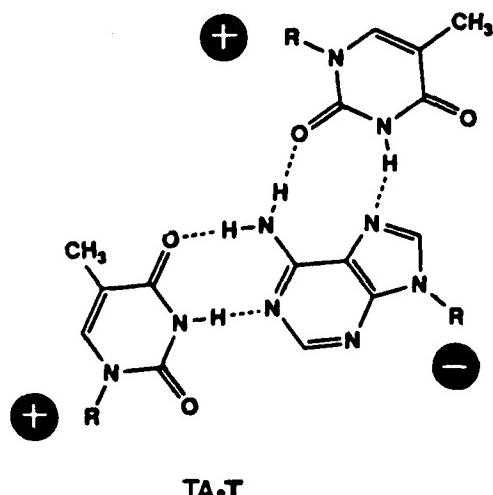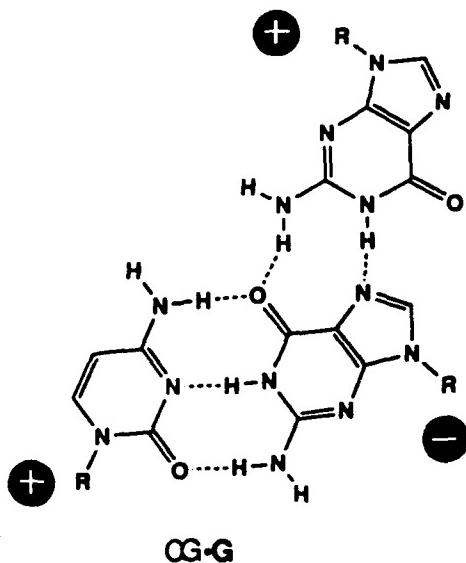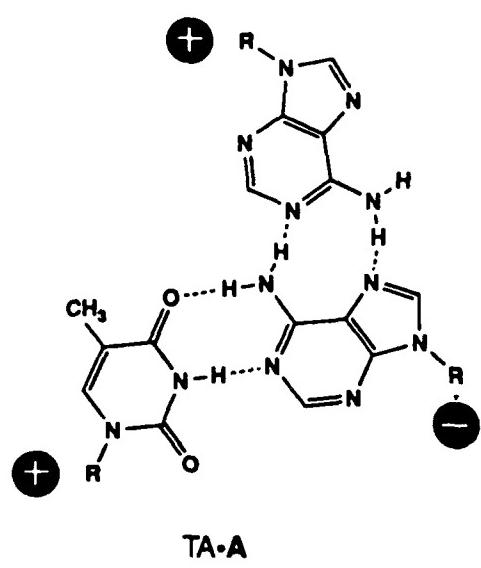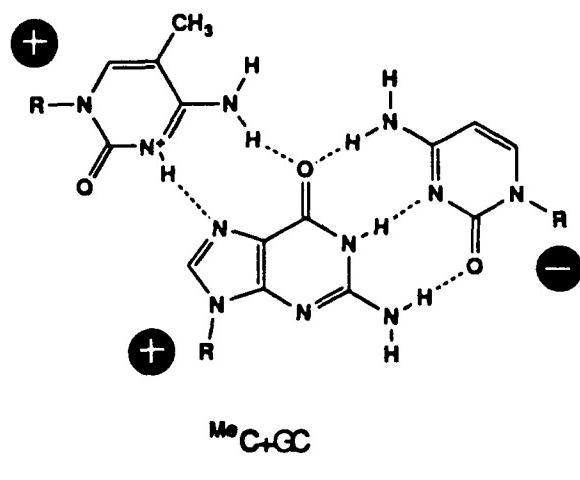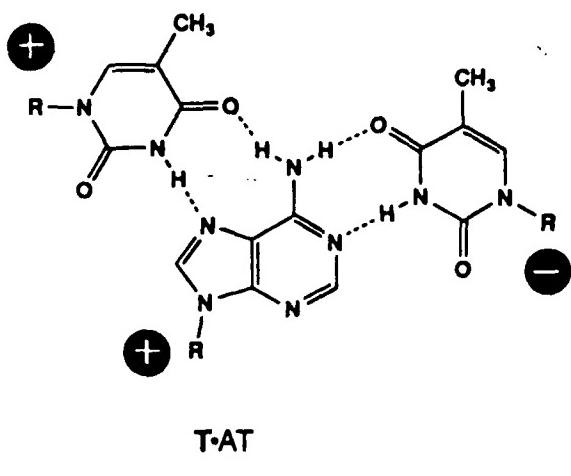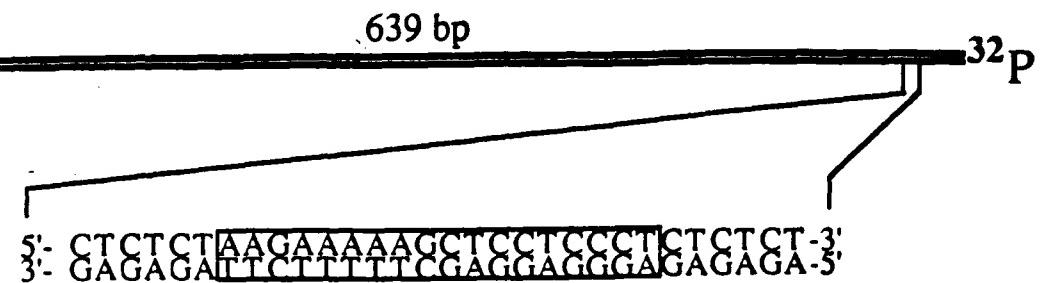
1   5'-*TTCTTTTTCGAGGAGGGT-3'

2   5'- TTCTTTTTCGAGGAGGGT*-3'

3   5'-*TTCTTTTTC -3'

4   5'- GAGGAGGGT*-3'

5   5'- TTCTTTTTCGAGGAGGGT-3'

Fig3

Oligo

	1	2	3	4	5	6	7
1	−	−	−	−	−	+	−
2	−	−	−	−	−	−	+
3	−	−	+	−	+	−	−
4	−	−	−	+	+	−	−

639 bp

Triple Helix Site

1

```
5' 3'
 |  |
T-A
T-A
A-T
A-T
C-G
T-A
C-G
T-A
C-G
T-A
* T•A-T
  T•A-T
  C•G-C
  T•A-T
  T•A-T
  T•A-T
  T•A-T
  T•A-T
  C•G-C
  C-G•G
  T-A•A
  C-G•G
  C-G•G
  T-A•A
  C-G•G
  C-G•G
  C-G•G
  T-A•T
  C-G
  T-A
  C-G
  T-A
  C-G
  T-A
  A-T
  G-C
  A-T
```

2

```
5' 3'
 |  |
T-A
T-A
A-T
A-T
C-G
T-A
C-G
T-A
C-G
T-A
T•A-T
T•A-T
C•G-C
T•A-T
T•A-T
T•A-T
T•A-T
T•A-T
C•G-C
C-G•G
T-A•A
C-G•G
C-G•G
T-A•A
C-G•G
C-G•G
C-G•G
T-A•T *
C-G
T-A
C-G
T-A
C-G
T-A
A-T
G-C
A-T
```

Fig 4

DNase I + + + + + +

10 μM Oligo **5** · · + + · ·

1.0 μM Oligo **5** · · · · + +

1 2 3 4 5 6 7 8 9 10

639 bp

Triple Helix Site

Fig5

Fig 6

A)

Oligo 1

```
5'-CTCTCT AAGAAAAAGCTCCTCCCT CTCTCT
3'-GAGAGA TTCTTTTTTCGAGGAGGGA GAGAGA
```

B)

Oligo 2

```
5'-CTCTCT AAGAAAAAGCTCCTCCCT CTCTCT
3'-GAGAGA TTCTTTTTTCGAGGAGGGA GAGAGA
```

C)

Oligo 5 (DNase I)

```
5'-CTCTCT AAGAAAAAGCTCCTCCCT CTCTCT
3'-GAGAGA TTCTTTTTTCGAGGAGGGA GAGAGA
```

D)

Oligo 5 (DMS)

```
                      *
5'-CTCTCT AAGAAAAAGCTCCTCCCT CTCTCT
3'-GAGAGA TTCTTTTTTCGAGGAGGGA GAGAGA
                         ** ***
```

Fig 7

639 bp

^{32}P

5'- CTCTCT TCCCTCCTCGAAAAAGAA CTCTCT-3'
3'- GAGAGA AGGGAGGAGCTTTTTCTT GAGAGA-5'

6   5'-*TGGGAGGAGCTTTTTCTT-3'

7   5'-*TGGGAGGAGTCTTTTTCTT-3'

8   5'-*TGGGAGGAGTTCTTTTTCTT-3'

9   5'- TGGGAGGAGTTCTTTTTCTT*-3'

10   5'-*T GGGAGGAG -3'

11   5'- CTTTTTCTT*-3'

12   5'- TGGGAGGAGTTCTTTTTCTT-3'

Fig 5

Oligo

	1	2	3	4	5	6	7	8	9
6	-	-	-	-	-	+	-	-	-
7	-	-	-	-	-	-	+	-	-
8	-	-	-	-	-	-	-	+	-
9	-	-	-	-	-	-	-	-	+
10	-	+	-	+	-	-	-	-	-
11	-	-	+	+	-	-	-	-	-

639 bp

Triple Helix Site

8

5' 3'
| |
T-A
T-A
A-T
A-T
C-G
T-A
C-G
T-A
C-G
T-A
T-A•T* 5'
C-G•G
C-G•G
C-G•G
T-A•A
C-G•G
C-G•G
T-A•A
T-C-G•G
C-G-C•T
T•A-T
T•A-T
T•A-T
T•A-T
T•A-T
G•G-C
T•A-T
T•A-T 3'
C-G
T-A
C-G
T-A
C-G
T-A
A-T
G-C
A-T

2

5' 3'
| |
T-A
T-A
A-T
A-T
C-G
T-A
C-G
T-A
C-G
T-A
T-A•T 5'
C-G•G
C-G•G
C-G•G
T-A•A
C-G•G
C-G•G
T-A•A
T-C-G•G
C-G-C•T
T•A-T
T•A-T
T•A-T
T•A-T
T•A-T
G•G-C
T•A-T
*T•A-T 3'
C-G
T-A
C-G
T-A
C-G
T-A
A-T
G-C
A-T

Fig 9

DNase I + + + + + +

10 μM Oligo 12 · · + + · ·

1.0 μM Oligo 12 · · · · + +

1 2 3 4 5 6 7 8 9 10

639 bp

Triple Helix Site

Fig 10

MPE•Fe + + + +
Oligo 12 - - + +

1 2 3 4 5 6 7 8

639 bp

Triple Helix Site

Fig 11

A)

Oligo 8

```
5'-C T C T C T|T C C C T C C T C G A A A A A G A A|C T C T C T
3'-G A G A G A|A G G G A G G A G C T T T T T T C T T|G A G A G A
```

B)

Oligo 9

```
5'-C T C T C T|T C C C T C C T C G A A A A A G A A|C T C T C T
3'-G A G A G A|A G G G A G G A G C T T T T T T C T T|G A G A G A
```

C)

Oligo 12 (DNase I)

```
5'-C T C T C T|T C C C T C C T C G A A A A A G A A|C T C T C T
3'-G A G A G A|A G G G A G G A G C T T T T T T C T T|G A G A G A
```

D)

Oligo 12 (MPE•Fe)

```
5'-C T C T C T|T C C C T C C T C G A A A A A G A A|C T C T C T
3'-G A G A G A|A G G G A G G A G C T T T T T T C T T|G A G A G A
```

Fig.