Recognition of Double Helical DNA by Purine Oligonucleotides via Triple Helix Formation

Thesis by

Peter A. Beal

under the direction of

Prof. Peter B. Dervan

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1994

(Submitted December 17, 1993)

Acknowledgements

I would like to express my gratitude to my advisor, Professor Peter B. Dervan. This work would not have been possible without Peter's guidance and support throughout my stay at Caltech. It was a true privilege to work for someone with Peter's enthusiasm, intellect and vision.

I would also like to thank my committee members, Professors Dennis Dougherty, Douglas Rees and Andrew Myers, for their help and advice during my quest for a Caltech Ph.D.

I also thank all the members of the Dervan group for their influence during my graduate career. I couldn't have completed any of the work discussed in this thesis if it weren't for their help in doing the experiments and thinking about the results. The Dervan group is an amazingly talented bunch of people and I will always be proud to say that I was a part of it. I'd like to thank my labmates John Josey, Huw Nash, Paul Gillespie, Scott Priestley, and Will Greenberg for making 330 Church a great place to go everyday to work and learn. Thank you Tom Povsic and Scott Strobel for your friendships, teaching me how to run gels and for our trips to Dodger Stadium. I thank Matt Taylor for his friendship, lunches at Eddie's Market and excursions to Old Town.

Thank you Stew Fisher and Susan Haney for being such great friends and roommates. I know you'll have a wonderful life together and I hope we will always remain close.

I don't think I could have made it through graduate school if it weren't for the constant support I received from my parents, brothers and sisters and I thank them for their love and encouragement.

Finally, I would like to thank Sheila David for her love, companionship, advice and support while I was doing this work. I know I will cherish the relationship we started at Caltech for the rest of my life.

To Mom, Dad, Suz, Pat, Margaret, Ida, Hank, Laura and Willie

and

To Sheila

Abstract

Oligonucleotide-directed triple helix formation is one of the most versatile methods for the sequence specific recognition of double helical DNA. Chapter 2 describes affinity cleaving experiments carried out to assess the recognition potential for purine-rich oligonucleotides via the formation of triple helices. Purine-rich oligodeoxyribonucleotides were shown to bind specifically to purine tracts of double helical DNA in the major groove antiparallel to the purine strand of the duplex. Specificity was derived from the formation of reverse Hoogsteen G•GC, A•AT and T•AT triplets and binding was limited to mostly purine tracts. This triple helical structure was stabilized by multivalent cations, destabilized by high concentrations of monovalent cations and was insensitive to pH. A single mismatched base triplet was shown to destabilize a 15 mer triple helix by 1.0 kcal/mole at 25°C. In addition, stability appeared to be correlated to the number of G•GC triplets formed in the triple helix. This structure provides an additional framework as a basis for the design of new sequence specific DNA binding molecules.

In work described in Chapter 3, the triplet specificities and required strand orientations of two classes of DNA triple helices were combined to target double helical sequences containing all four base pairs by alternate strand triple helix formation. This allowed 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 depended on whether the binding site sequence was 5'-$(purine)_m(pyrimidine)_n$-3' or 5'-$(pyrimidine)_m(purine)_n$-3′.

In Chapter 4, the ability of oligonucleotide-cerium(III) chelates to direct the transesterfication of RNA was investigated. Procedures were developed

for the modification of DNA and RNA oligonucleotides with a hexadentate Schiff-base macrocyclic cerium(III) complex. In addition, oligoribonucleotides modified by covalent attachment of the metal complex through two different linker structures were prepared. The ability of these structures to direct transesterification to specific RNA phosphodiesters was assessed by gel electrophoresis. No reproducible cleavage of the RNA strand consistent with transesterification could be detected in any of these experiments.

Table of Contents

List of Tables and Figures

Chapter 3

Chapter 1

Introduction

Deoxyribonucleic acid (DNA) is the primary genetic material controlling the inheritance of traits and is the template coding for all the proteins and RNA necessary for cellular function.[1] DNA typically adopts a regular right-handed double helical secondary structure (B form), although other conformational families exist.[2] Structural heterogeneity in DNA arises from a local variation in the sequence of base pairs. Therefore, the sequence of double helical DNA provides an estimate for its structure at that location. Because of the essential functional roles of DNA and its regularity in secondary structure, it is an attractive target for the chemist who wishes to design molecules which bind specifically to an important biological receptor and control its function.[3] In addition, synthetic reagents which bind to defined sites and cleave the DNA molecule at those sites complement existing restriction enzymes for the manipulation of DNA fragments in cloning and sequencing protocols by offering the possibility of greater specificity.[4] For these reasons, the generation of DNA binding molecules which can specifically recognize predetermined sequences under physiologically relevant solution conditions is a worthwhile endeavor and the primary goal of the studies in this thesis.

The DNA double helix. Analysis of double helical DNA structure can provide insight into how sequence specific recognition might be achieved. Duplex DNA arises through the formation of hydrogen bonded base pairs between two polydeoxyribonucleotide strands creating a cylindrical grooved

structure.[5] The major and minor grooves have different hydrogen bonding sites and van der Waals surfaces depending on the identity of the base pair at that site (Figure 1.1). For instance, the major groove at a GC base pair contains two hydrogen bond acceptors (guanine-N7 and O6) and a hydrogen bond donor (cytosine-NH_2-4). In contrast, the major groove at an AT base pair has the hydrogen bond acceptors adenine-N7 and thymine-O4, the hydrogen bond donor adenine-NH_2-6 and the thymine C5 methyl group. A given DNA sequence can therefore be distinguished by the array of functional groups it displays in the grooves of the double helix (Figure 1.2).

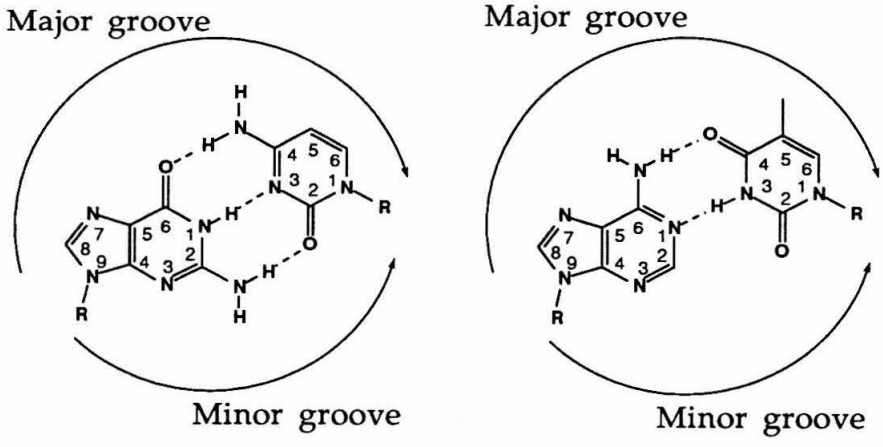


Figure 1.1 Watson Crick base pairs.

The design of new sequence specific binding molecules is aided by principles learned from structural analyses of complexes of DNA with proteins and natural products. It is apparent from these studies that sequence specific DNA binding arises from the ability of these molecules to form specific hydrogen bonds or van der Waals contacts with functional groups in the grooves, by Coulombic attraction to the negatively charged phosphodiesters and through intercalation of aromatic functional groups

Figure 1.2 Model of the structure of B form double helical DNA of the sequence (5'-CGCGAATTCGCG-3')$_2$.[6] Atoms of the bases are colored individually by atom type (O=red, N=blue, C=green, and H=white). The phosphodiester deoxyribose backbones are colored light blue. Each sequence of double helical DNA can be distinguished by the array of functional groups it displays in the grooves.

between the DNA bases. Some or all of these binding modes are used in combination to accomplish high affinity sequence specific binding.

Sequence specific DNA binding proteins. Sequence-specific recognition of double helical DNA by proteins is achieved through the assembly of structural motifs within the protein which place amino acid side chains in the correct relative positions to interact through hydrogen bonding or van der Waals contacts to functional groups in the grooves or by salt bridges to the negatively charged phosphodiester backbone.[7] Several motifs have been identified which serve this purpose including the helix turn helix,[8] the leucine zipper[9] and the zinc finger.[10] Side chain interactions with DNA, such as the binding to GC base pairs by arginine or histidine found for the Zif268 zinc finger protein-DNA complex, show how hydrogen bonding to recognition sites in the grooves can be used to engender sequence specificity (Figure 1.3).[10]

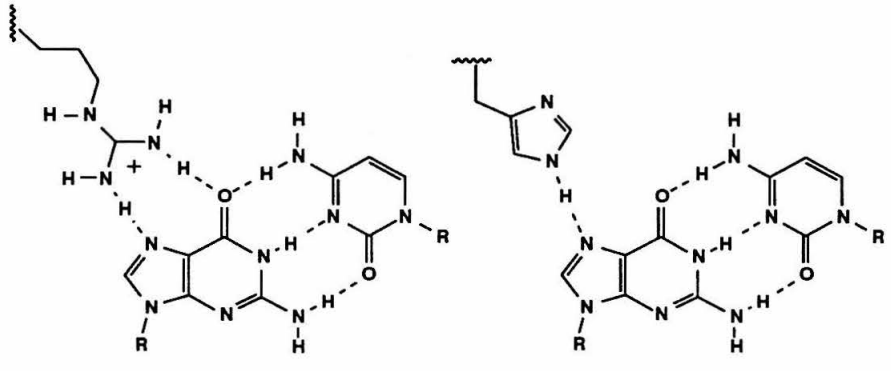


Figure 1.3 Specific amino acid-base contacts found in the Zif268 zinc finger DNA binding protein-DNA complex.[10] Secondary structural elements in DNA binding proteins position amino acid side chains for interaction with bases and phosphodiesters.

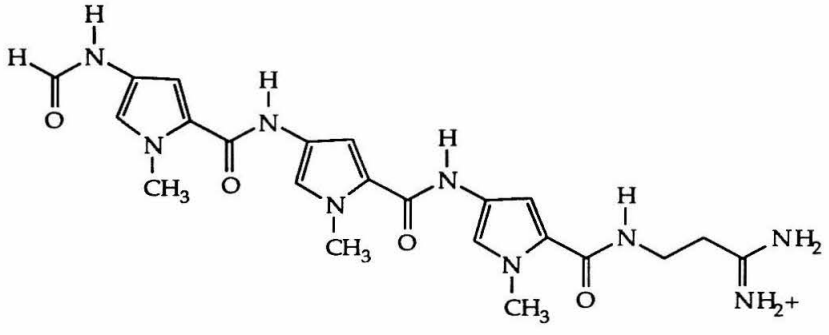
In addition, aliphatic amino acids such as alanine have been shown to make van der Waals contacts with thymine methyl groups. Arginine is often

found ion paired with DNA phosphodiesters. Also, aromatic residues such as phenylalanine can interact with the duplex through stacking by inserting between adjacent base pairs.

These studies provide information about how DNA structure can direct sequence specific binding, however, it appears that properly folded secondary structural elements are necessary to orient the side chains of DNA binding proteins to interact with the necessary DNA functional groups. In several cases the binding domains of DNA binding proteins have been synthesized and shown to bind specifically to the native binding site.[11] However, the design of *new proteins* that recognize sequences different from the native binding site requires not only the proper selection of amino acids for recognition but also the ability to control folding such that the side chains are positioned properly. Therefore, protein design is a challenging approach to the generation of new sequence specific DNA binding molecules. It remains to be seen if truly new sequence specificities can be engineered into designed DNA binding proteins.

DNA binding natural products. Several natural products have been isolated that are capable of binding to DNA sequence specifically.[12] These include the minor groove binders distamycin (Figure 1.4), netropsin, and Hoechst 33258, the minor groove binding intercalators actinomycin and echinomycin and the ene-diyne family of antitumor antibiotics. Where structural studies have been reported, these molecules are shown to bind by intercalation and/or specific hydrogen bonding and van der Waals contacts in the minor groove. For instance, x-ray crystallographic[13] and NMR[14] studies show that distamycin binds either as a 1:1 complex with DNA or as a 2:1 side by side antiparallel dimer at sites containing four to five consecutive A,T base pairs. Specificity is derived from the formation of bifurcated hydrogen bonds

between the carboxamide NH's in the natural product with adenine N3 and thymine O2 in the minor groove. Additional specificity arises from the shape complemenarity between the crescent shaped natural product and the minor groove at A,T-rich sites.



Distamycin

Figure 1.4 Structure of the minor groove binding natural product distamycin

The sequence selectivity demonstrated by distamycin has been exploited to design new molecules with new recognition properties. The 1:1 model for distamycin led to the design of oligomeric distamycin analogs capable of recognizing longer tracts of A,T rich DNA.[15] The presence of the 2-amino group of guanosine in the minor groove at G,C base pairs presented a steric block to these molecules and therefore binding was limited to A,T sites. However, substitutions for the terminal *N*-methylpyrrole ring of distamycin by 1-methylimidazole or pyridine generated synthetic molecules capable of recognizing the double helical sequence 5'-TGTCA-3' as 2:1 side by side dimers.[16] The systematic alteration of the specificities of minor groove binding natural products through synthesis continues to be a productive method for generating new sequence specific DNA binding molecules.

<u>**Oligonucleotide-directed triple helix formation.**</u> Currently the most

versatile method for the sequence specific recognition of DNA is through triple helix formation. Synthetic oligonucleotides are capable of binding to tracts of intact double helical DNA through the formation of a local triple helical structure.[17,18] Specificity arises from the formation of base triplets, where hydrogen bonding occurs between bases in the oligonucleotides and the duplex base pairs. The ability to target a broad range of DNA sequences, the high stabilities of the resulting triple helices and the sensitivity to mismatches makes this a powerful technique for binding single sites in large segments of double helical DNA. However, sequence limitations as well as limitations on the solution conditions under which these structures can form do exist. Overcoming these limitations by the design of modified oligonucleotides is a current area of intense investigation.[19]

Triple helical nucleic acid structures were discovered over 30 years ago when Felsenfeld, Davies and Rich demonstrated that poly (U) bound to poly (A) in a two to one stoichiometry in the presence of $MgCl_2$.[20] Shortly thereafter, poly (C) was shown to form a two to one complex with guanosine oligoribonucleotides at low pH.[21] Further work on polymeric nucleic acids demonstrated that poly(dT-dC) or poly(U-C) associated with poly(dG-dA)•poly(dC-dT) at acidic pH to form a three stranded complex.[22] Formation of isomorphous T(U)•AT(U) and C+GC triplets were proposed to rationalize the binding mode for the third strand in these complexes (Figure 1.5). The apparent requirement for protonation of cytosine explained the necessity for acidic pH for formation of triple helices containing this base. X-ray diffraction studies with fibers of poly(U)2poly(A) and poly(dT)2poly(dA) suggested an A' helical conformation with the third strand bound parallel to the purine strand of the underlying duplex and *anti* conformations for the bases.[23] The

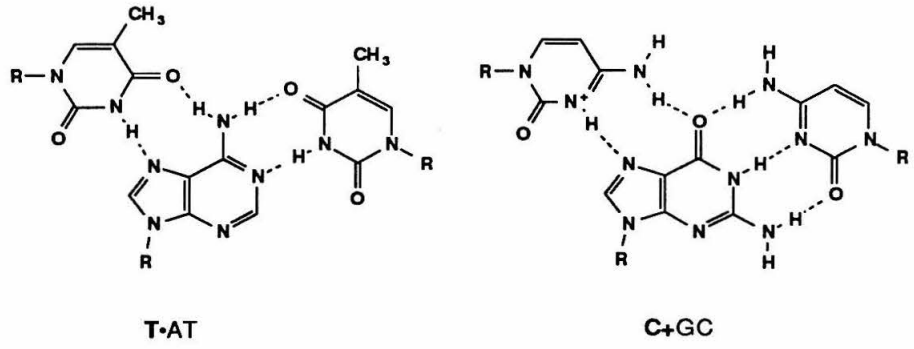


Figure 1.5 Isomorphous T•AT and C+GC triplets

affinity cleaving experiments of Moser and Dervan demonstrated unambiguously that pyrimidine oligonucleotides bind in the major groove parallel to the purine strand of double helical DNA.[17a] This work also showed that pyrimidine oligonucleotides equipped with the cleaving function EDTA•Fe could bind to a 15 bp site within a 4 kbp plasmid and cleave the DNA at that site.

The ability of pyrimidine oligonucleotides to bind uniquely to sites of 15-16 bp suggested that these molecules might have the necessary sequence selectivity to bind to one site in genomic DNA.[24] Indeed, pyrimidine oligonucleotides have been used to mediate single site specific cleavage of human chromosomal DNA.[4] Triple helix formation has also been shown to block access to DNA by DNA binding proteins such as methylases, transcription factors and polymerases *in vitro*, suggesting that they might be functional in controlling gene expression *in vivo*.[17e,18j]

Triple helices formed by pyrimidine oligonucleotides have been studied extensively in recent years. ^{1}H NMR spectroscopy has confirmed the structures of the hydrogen bonded base triplets and has been used to determine sugar puckering in the deoxyribose backbones.[25] A kinetic analysis

of the binding of a 21 mer pyrimidine oligonucleotide suggests that triple helix formation is a slow process (pseudo first order rate constant of 1.8±0.2 x 10^3 $M^{-1}s^{-1}$), but the formed complex is long lived (half life of ~ 12 hours).[17i] Thermodynamic studies indicate that stability is highly dependent on solution conditions, with binding free energies varying from 6 to 10 kcal/mole depending on the conditions chosen.[17q,r] Oligonucleotide binding has been shown to be sensitive to single base mismatches and the extent of destabilization is dependent on the identity of the mismatched triplet.[17f,q] The triple helical structures are stabilized by polycations such as spermine or $Co(NH_3)_6^{+3}$.[17a] The formation of triple helices with C+GC triplets requires acidic pH.[17a,r] Extension of the pH range for triple helix formation at GC rich sites has been accomplished through methylation of cytosine at C5[17c] and through the use of nonnatural base triplets.[17m,19g-i] Oligonucleotide binding appears to be limited to purine tracts of duplex DNA, since stabilizing base triplets form only with GC and AT base pairs. This sequence limitation has been partially alleviated through the use of G•TA triplets,[17f] nonnatural base triplets[17m,19k,l] and alternate strand triple helix formation by 3'-3' or 5'-5' linked pyrimidine oligonucleotides.[17j, 19a,b,d] However, a general solution to the recognition of any DNA sequence by triple helix formation still requires effective methods for specifically recognizing TA and CG base pairs within this structure.

The triple helices formed by pyrimidine oligonucleotides constitute one of two different families of triple helical structures. A second class of triple helices is introduced and discussed in Chapter 2. This structure differs from that formed by pyrimidine oligonucleotides in the base composition of the third strand, the relative orientations and positions of the three strands

and the base triplet interactions. Formation of this structure is also dependent on solution conditions and duplex sequence.

DNA affinity cleaving and footprinting. Studies of DNA binding molecules such as triple helix forming oligonucleotides are aided by the techniques of DNA footprinting[26] and affinity cleaving[27] (Figure 1.6). Properties such as binding site size and orientation of the DNA binding molecule can be determined through the use of these experiments. Also, with the development of quantitative DNase I footprinting[28] and quantitative affinity cleaving,[17q] binding free energies can be determined for DNA binding molecules under a variety of solution conditions.

DNA affinity cleaving has become a standard technique for assessing binding properties of DNA binding molecules. This method involves the covalent attachment of a nonspecific DNA cleaving function to a sequence specific DNA binding molecule to generate a sequence specific DNA cleaving molecule. EDTA•Fe is useful as the nonspecific cleaving function, since it generates a diffusible oxidant (presumably hydroxyl radical) which initiates cleavage of the deoxyribose backbone at nearby locations in the bound complex. This cleavage is sequence independent and therefore information about the sequence selectivity and binding orientation of the covalently attached molecule can be determined by locating the cleavage sites using sequencing gels. The groove location of the EDTA•Fe in the bound complex can be determined from analysis of the pattern of cleavage bands. If EDTA•Fe is located in the major groove of right handed double helical DNA, the major sites of cleavage are asymmetrically shifted in the 5' direction, whereas if the cleaving function is located in the minor groove, the major cleavage sites are 3' shifted (Figure 1.7). Also, it has been shown that the amount of cleavage detected is proportional to the fractional occupancy of the binding site.[17q]

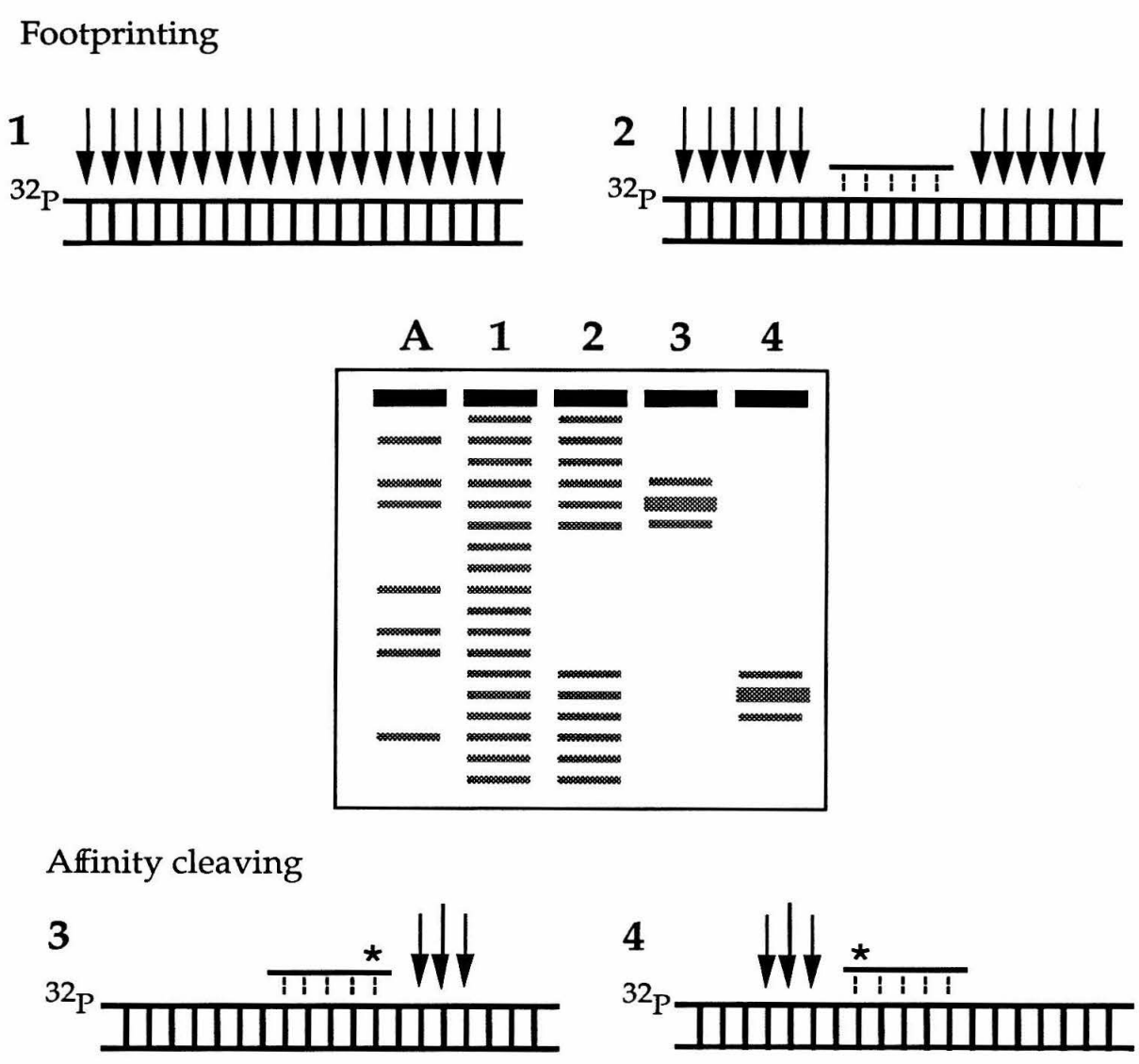


Figure 1.6 Schematic description of DNA footprinting and affinity cleaving. Footprinting involves treatment of ^{32}P labeled DNA with nonspecific cleavage reagents in the presence and absence of the DNA binding molecule. The cleavage products are separated by polyacrylamide gel electrophoresis along with the products of sequencing reactions. Autoradiography reveals the location and size of the binding site as positions protected from cleavage (Middle, lanes **1** and **2**). Affinity cleaving requires the use of DNA binding molecules covalently attached to a nonspecific DNA cleaving moiety (*). The binding site location and binding orientation are indicated by the location of the cleaved nucleotides (Middle, lanes **3** and **4**).

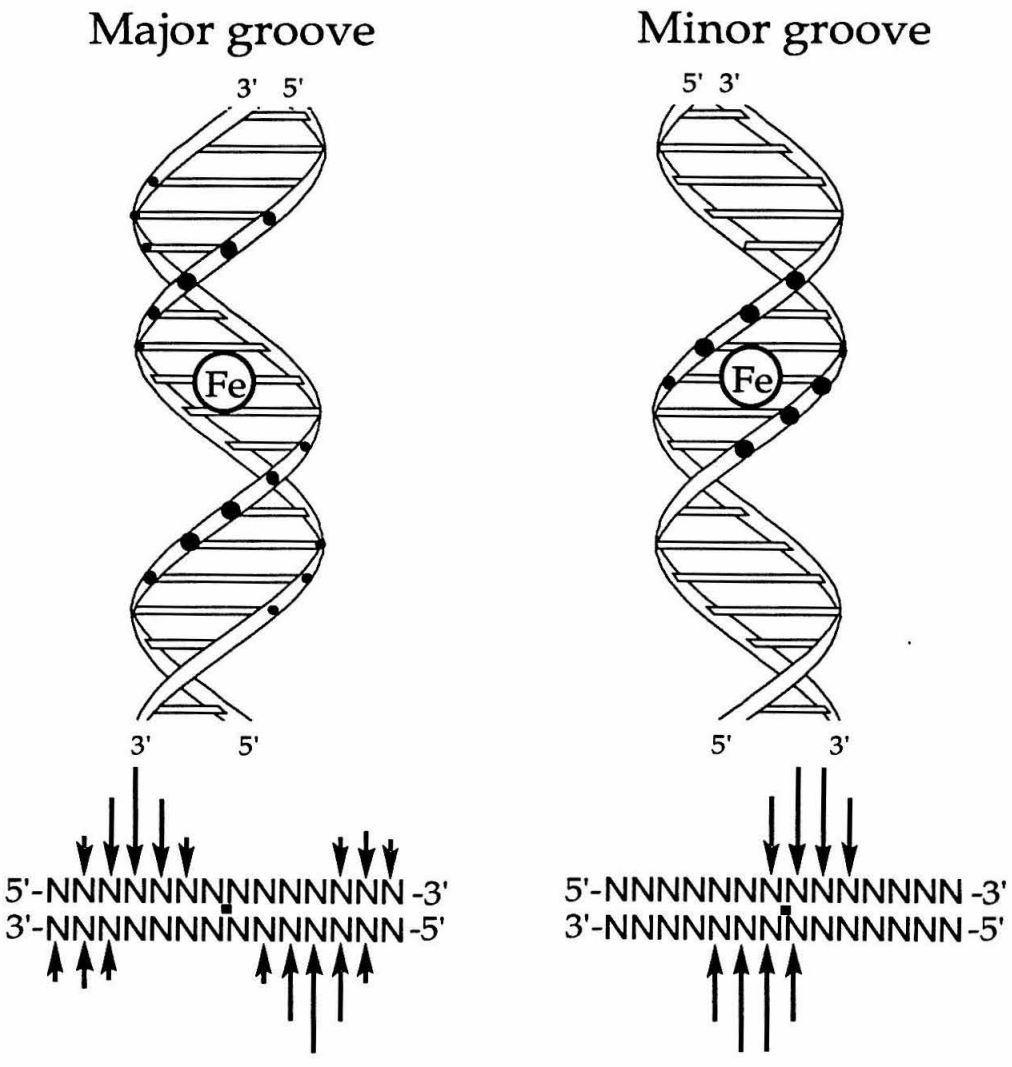


Figure 1.7 Cleavage patterns generated by EDTA•Fe located in the major or minor grooves of right handed double helical DNA. Circles and arrows indicate cleavage sites.

This allows for a qualitative assessment of differences in binding affinities for DNA binding molecules modified with EDTA•Fe by measuring differences in cleavage efficiencies. In addition, use of the quantitative affinity cleavage titration can yield the binding free energies for triple helix forming oligonucleotides.[17q]

Oligonucleotides are readily functionalized with EDTA•Fe for affinity cleaving studies through automated DNA synthesis using a protected thymidine-EDTA(T*) phosphoramidite (Figure 1.8).[29]

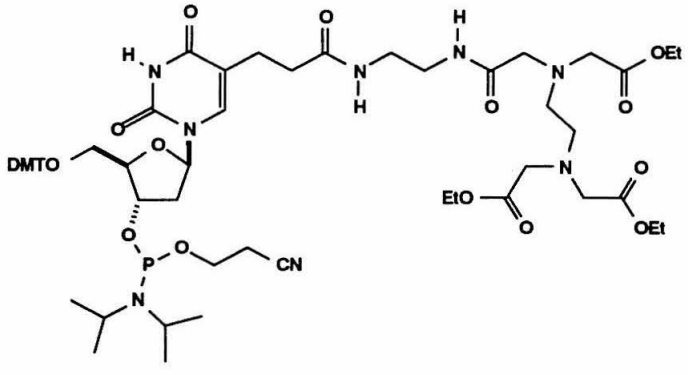


Figure 1.8 Structure of the thymidine-EDTA (T*) phosphoramidite

Description of this work. Chapter 2 of this thesis introduces a second structural motif for the recognition of double helical DNA based on triple helix formation with purine oligonucleotides. Affinity cleaving experiments were carried out to define properties of this triple helical structure such as relative strand orientations, specificity and sensitivity to solution conditions and sequence constraints. Results from these studies led to the publication of reports in *Science*[30] and *Nucleic Acids Research*.[31] In Chapter 3, this triple helical structure is used in tandem with the pyrimidine triple helix to bind sequences of double helical DNA containing all four base pairs by alternate strand triple helix formation. A study on the alternate strand triple helix

including most of the data presented in this chapter was reported in the *Journal of the American Chemical Society*.[32] Chapter 4 describes some exploratory studies on the use of oligonucleotides covalently attached to a hexadentate Schiff-base macrocyclic cerium(III) complex for the sequence specific transesterification of RNA. This project is ongoing in the Dervan laboratories.

References and Notes

1. Watson, J.D.; Hopkins, N.H.; Roberts, J.W.; Steitz, J.A.; Weiner, A.M.; *Molecular Biology of the Gene* (Benjamin/Cummings, Menlo Park, CA, 1987).
2. Saenger, W.; *Principles of Nucleic Acid Structure* (Springer-Verlag, New York, 1984).
3. Neilson, P.E.; *Bioconj. Chem.* **1991**, 2, 1.
4. Strobel, S.A.; Doucette-Stamm, L.A.; Riba, L.; Housman, D.E.; Dervan, P.B.; *Science* **1991**, 254, 1639.
5. Watson, J.D.; Crick, F.H.C.; *Nature* **1953**, 171, 737.
6. Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R.E.; *Nature* **1980**, 287, 755.
7. Schleif, R.; *Science* **1988**, 241, 1182.
8. Pabo, C.O.; Sauer, R.T.; *Ann. Rev. Biochem.* **1984**, 53, 293.
9. Ellenberger, T.E.; Brandl, C.J.; Struhl, K.; Harrison, S.C.; *Cell* **1992**, 71, 1223.
10. Pavletich, N.P.; Pabo, C.O.; *Science* **1991**, 252, 809.
11. (a) Bruist, M.F. et al.; *Science* **1987**, 235, 777. (b) Sluka, J.P.; Horvath, S.J.; Bruist, M.F.; Simon, M.I.; Dervan, P.B.; *Science* **1987**, 238, 1129. (c)

Oakley, M.G.; Dervan, P.B.; *Science* **1990**, 248, 847. (d) Talanian, R.V.; McKnight, C.J.; Kim, P.S.; *Science* **1990**, 249, 769.

12. (a) Neidle, S.; *CRC Crit. Rev. Biochem.* **1984**, 17, 73. (b) Lee, M.D., Ellestad, G.A., Borders, D.B.; *Acc. Chem. Res.* **1991**, 24, 235.

13. Coll, M.; Frederick, C.A.; Wang, A. H-J.; Rich, A.; *Proc. Natl. Acad. Sci. USA* **1987**, 84, 8385.

14. Pelton, J.G.; Wemmer, D.E.; *J. Am. Chem. Soc.* **1990**, 112, 1393.

15. Youngquist, R.S.; Dervan, P.B.; *Proc. Natl. Acad. Sci. U.S.A.* **1985**, 82, 2565.

16. (a) Wade, W.S.; Mrksich, M.; Dervan, P.B.; *J. Am. Chem. Soc.* **1992**, 114, 8783. (b) Mrksich, M.; Wade, W.S.; Dwyer, T.J.; Deierstanger, B.H.; Wemmer, D.E.; Dervan, P.B.; *Proc. Natl. Acad. Sci. USA* **1992**, 89, 7586.

17. (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) Horne, D.A.; Dervan, P.B.; *J. Am. Chem. Soc.* **1990**, 112, 2435. (k) Strobel, S.A.; Dervan, P.B.; *Nature* **1991**, 350, 172. (l) Distefano, M.D.; Shin, J.A.; Dervan, P.B.; *J. Am. Chem. Soc.* **1991**, 113, 5901. (m) Koh, J.S.; Dervan, P.B.; *J. Am. Chem. Soc.* **1992**, 114, 1470. (n) Griffin, L.C.; Kiessling, L.L.; Beal, P.A.; Gillespie, P.; Dervan, P.B.; *J. Am. Chem. Soc.* **1992**, 114, 7977. (o) Distefano, M.D.; Dervan, P.B.; *J. Am. Chem. Soc.* **1992** 114, 11006. (p)

Maher, L.J.; Dervan, P.B.; Wold, B.; *Biochemistry* **1992**, 31, 70. (q) Singleton, S.F.; Dervan, P.B.; *J. Am. Chem. Soc.* 1992, **114**, 6956. (r) Singleton, S.F.; Dervan, P.B.; *Biochemistry* **1992**, 31, 10995. (s) Distefano, M.D.; Dervan, P.B.; *Proc. Natl. Acad. Sci. USA* **1993**, 90, 1179. (t) Colocci, N.; Distefano, M.D.; Dervan, P.B.; *J. Am. Chem. Soc.* **1993**, 115, 4468.

18. (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. (h) Mergny, J-L. et al.; *Biochemistry* **1991**, 30, 9791. (i) Yoon, K.; Hobbs, C.A.; Koch, J.; Sardaro, M.; Kutny, R.; Weis, A.; *Proc. Natl. Acad. Sci. USA* **1992**, 89, 3840. (j) Maher, L.J.; *Biochemistry* **1992**, 31, 7587. (k) Mergny, J.L. et al.; *Science* **1992**, 256, 1681. (l) Thuong, N.T.; Helene, C.; *Angew. Chem. Int. Ed. Engl.* **1993**, 32, 666.

19. (a) Ono, A.; Chen, C-N.; Kan, L-S.; *Biochemistry* **1991**, 30, 9914. (b) McCurdy, S.; Froehler, B.; *Nucleosides and Nucleotides* **1991**, 10, 287. (c)

Callahan, D.E.; Trapane, T.L.; Miller, P.S.; T'so, P.O.P.; Kan, L-S.; *Biochemistry* **1991**, 30, 1650. (d) Froehler, B.C.; Terhorst, T.; Shaw, J-P.; McCurdy, S.N.; *Biochemistry* **1992**, 31, 1603. (e) Escude, C. et al.; *C.R. Acad. Sci. Paris* **1992**, 315, 521. (f) Shimizu, M.; Konishi, A.; Shimada, Y.; Inoue, H.; Ohtsuka, E.; *FEBS* **1992**, 302, 155. (g) Miller, P.; Bhan, P.; Cushman, C.; Trapane, T.; *Biochemistry* **1992**, 31, 6788. (h) Ono, A.; Ts'o, P.O.P.; Kan, L-S.; *J. Org. Chem.* **1992**, 57, 3225. (i) Krawczyk, S.H.; Milligan, J.F.; Wadwani, S.; Moulds, C.; Froehler, B.C.; Matteucci, M.; *Proc. Natl. Acad. Sci. USA* **1992**, 89, 3761. (j) Kim, S-G.; Tsukahara, S.; Yokoyama, S.; Takaku, H.; *FEBS* **1992**, 314, 29. (k) Huang, C-Y,; Cushman, C.D.; Miller, P.S.; *J. Org. Chem.* **1993**, 58, 5048. (l) Huang, C-Y.; Miller, P.S.; *J. Am. Chem. Soc.* **1993**, 115, 10456.

20. Felsenfeld, G.; Davies, D.R.; Rich, A.; *J. Am. Chem. Soc.* **1985**, 107, 5528.

21. Lipsett, M.N.; *Biochem. Biophys. Res. Commun.* **1963**, 11, 224.

22. (a) Morgan, A.R.; Wells, R.D.; *J. Mol. Biol.* **1968**,37, 63. (b) Lee, J.S.; Johnson, D.A.; Morgan, A.R.; *Nucleic Acids Res.* **1979**, 6, 3073.

23. (a) Arnott, S.; Bond, P.J.; *Nature New Biol.* **1973**, 244, 99. (b) Arnott, S.; Selsing, E.; *J. Mol. Biol.* **1974**, 88, 509. (c) Arnott, S.; Bond, P.J.; Selsing, E.; Smith, P.J.C.; *Nucleic Acids Res.* **1976**, 3, 2459.

24. Dervan, P.B.; *Science* **1986**, 232, 464.

25. (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. (g) Macaya, R.F.; Gilbert, D.E.; Malek, S.; Sinsheimer, J.S.; Feigon, J.; *Science* **1991**, 254, 270. (h)

Radhakrishnan, I.; Patel, D.J.; *J. Am. Chem. Soc.* **1992**, 114, 6913. (i) Macaya, R.F.; Schultz, P.; Feigon, J.; *J. Am. Chem. Soc.* **1992**, 114, 781.

26. (a) Galas, D. J.; Schmitz, A., *Nucleic Acids Res.* **1978**, 5, 3157. (b) Van Dyke, M.W.; Dervan, P.B.; *Cold Spring Harbor Symp. Quant. Biol.* **1982**, 47, 347. (c) Uchida, K.; Pyle, A.M.; Morii, T.; Barton, J.K.; *Nucleic Acids Res.* **1989**, 24, 10259. (d) Kohwi-Shigematsu, T.; Kohwi, Y.; *Methods in Enzy.* **1992**, 212, 155.

27. Taylor, J.S.; Schultz, P.G.; Dervan, P.B.; *Tetrahedron* **1984**, 40, 457.

28. (a) Brenowitz, M.; Senear, D.; Shea, M.; Ackers, G.K.; *Proc. Natl. Acad. Sci. USA* **1986**, 83, 8462. (b) Brenowitz, M.; Senear, D.F.; Shea, M.A.; Ackers, G.K.; *Methods in Enzy.* **1986**, 130, 132.

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

30. Beal, P.A.; Dervan, P.B.; *Science* **1991**, 251, 1360.

31. Beal, P.A.; Dervan, P.B.; *Nucleic Acids Res.* **1992**, 20, 2773.

32. Beal, P.A.; Dervan, P.B.; *J. Am. Chem. Soc.* **1992**, 114, 4976.

Chapter 2

A Second Structural Motif for the Recognition of Double Helical DNA by Triple Helix Formation

Introduction

Early work on homopolymers of DNA and RNA suggested that in addition to triple helical nucleic acid structures comprised of two pyrimidine strands and one purine strand (a pyrimidine motif), there existed structures comprised of two purine strands and one pyrimidine strand (a purine motif). Shortly after Felsenfeld, Davies and Rich discovered that poly(U) bound to poly(A) to form a stable 2:1 complex in the presence of $MgCl_2$, Lipsett showed that GpGpG bound to poly(C) in a 2:1 stoichiometry.[1] It was proposed that the $(GpGpG)_2$(poly(C)) structure was triple helical and was generated through the formation of G•GC triplets (Figure 2.1). Later, Marck and Theile demonstrated that the deoxyribo homopolymers poly(dG) and poly(dC) could form $(poly(dG))_2$(poly(dC) at pH 8.[2]

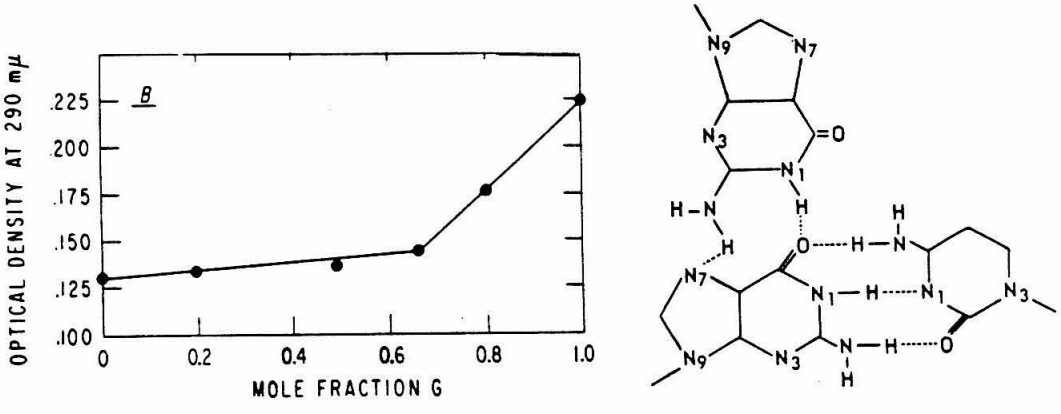


Figure 2.1 (Left) UV mixing curve for GpGpG + poly(C) (Right) G•GC base triplet proposed by Lipsett.[1]

More recently, Fresco reported that poly(A) of specific degrees of polymerization bound to (poly(A))(poly(U)) to form $(poly(A))_2(poly(U))$. Interestingly, only poly(A) of length 28-150 formed this triplex structure.[3]

Studies of purine tracts in plasmid DNA also suggested that triple helices with two purine strands formed. Kohwi-Shigematsu carried out extensive chemical reactivity studies on a poly(dG)poly(dC) tract under torsional stress in supercoiled plasmid DNA.[4] The results indicated that this sequence can adopt two nonB DNA conformations depending on the solution conditions (Figure 2.2). In the absence of Mg^{+2} and at acidic pH, the nucleotides in the 5′ half of the dG strand and at the center of the dC strand reacted with reagents specific for single stranded DNA. However, in the presence of Mg^{+2}, the nucleotides in the 5' half

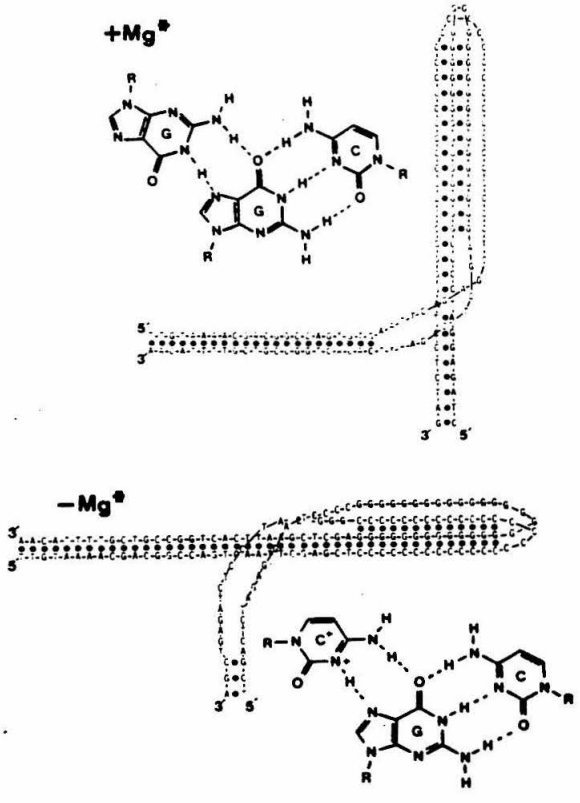


Figure 2.2 Models for nonB DNA conformations adopted by poly(dG) tract in plasmid DNA proposed by Kohwi-Shigematsu.[4]

of the dC strand and at the center of the dG strand were reactive. Two triple helices were proposed to explain these data; one with the dC strand folded back parallel to the dG strand of the WC duplex forming H-form DNA with C+GC triplets, the other with the dG strand folded back upon itself in an antiparallel fashion forming G•GC triplets.

Evidence that a purine-rich oligonucleotide could bind site specifically to duplex DNA was first provided by Hogan and co-workers when they demonstrated that a guanosine-rich 27 mer bound to a duplex sequence present in the c-*myc* promoter (Figure 2.3).[5] Binding of this oligonucleotide to the promoter repressed transcription of the c-*myc* gene *in vitro*, demonstrating the potential utility of this structure. The binding of this oligonucleotide was proposed to be due to the formation of G•GC and A•AT triplets with the third purine-rich strand parallel to the purine strand of the duplex (Figure 2.3). Although these authors did not present hydrogen bonding models to explain the sequence specific binding, it was suggested that the triplets formed may be similar to the base triplets found stabilizing tertiary structure in tRNAs. The third strand orientation was based on that found for the $(poly(dT))_2(poly(dA))$ triple helix.

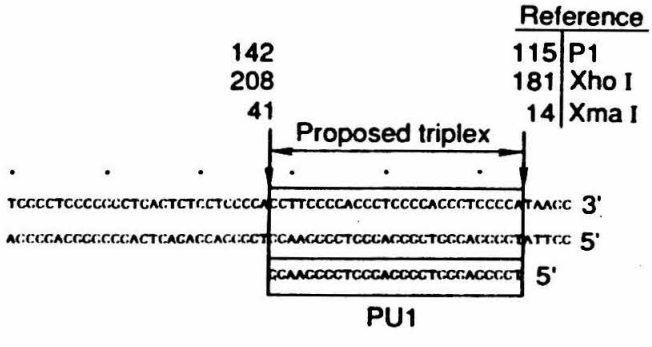


Figure 2.3 Sequence of 27 mer G rich oligonucleotide PU1 designed by Hogan to bind via triple helix formation in the *c-myc* promoter.[5]

Triplet Models. The sequence specificity of triple helix formation derives from the formation of hydrogen bonded base triplets. It is therefore of primary importance to understand the nature of these triplets. In the earliest work on the formation of purine motif triple helices, Lipsett proposed the formation of G•GC triplets with the third strand guanosine bound to a GC base pair by hydrogen bonding (Figure 2.1).[1] In this model, the N2 exocyclic amino group and the N1 NH of the third strand guanosine were hydrogen bonded to the guanosine of the Watson-Crick (WC) base pair at N7 and O6 respectively. This type of G•G interaction places the sugars of the nucleotides in similar relative positions as seen for the Hoogsteen T•A interaction and has therefore been referred to as a Hoogsteen G•G interaction (Figure 2.4).[6] The Hoogsteen G•G interaction is found in guanine quartets(G-quartets), which form with telomeric DNA sequences.[7]

Kohwi-Shigematsu proposed a different G•GC triplet with the N2 exocyclic amino group and N1 NH of the third strand G bound to the G in the WC base pair at O6 and N7, respectively (Figure 2.2).[4] This G•G interaction places the sugars in relative positions similar to that in the reverse Hoogsteen T•A interaction and is referred to as reverse Hoogsteen (Figure 2.4).[6] Reverse Hoogsteen G•G hydrogen bonding is found in crystal structures of guanine and guanosine and has been calculated to be the most stable G•G interaction in chloroform.[8,9] In addition, Williams et al. showed by NMR that the reverse Hoogsteen bonded G•GC triplet was more stable than the Hoogsteen bonded G•GC triplet in chloroform.[10] Importantly, the crystal structure of yeast $tRNA^{phe}$ revealed that the tertiary fold of this tRNA has a reverse Hoogsteen bonded ^{m7}G46•G22•C13 triplet, with the strands containing the ^{m7}G46 and G22 in an antiparallel orientation (Figure 2.5).[11] Klug et al. note that this triplet is

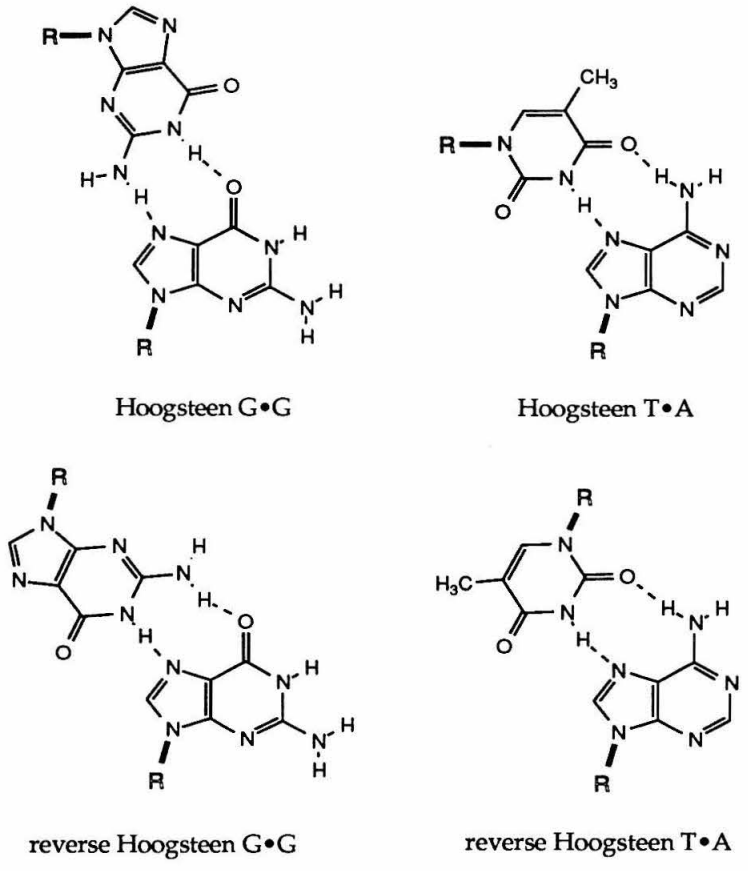
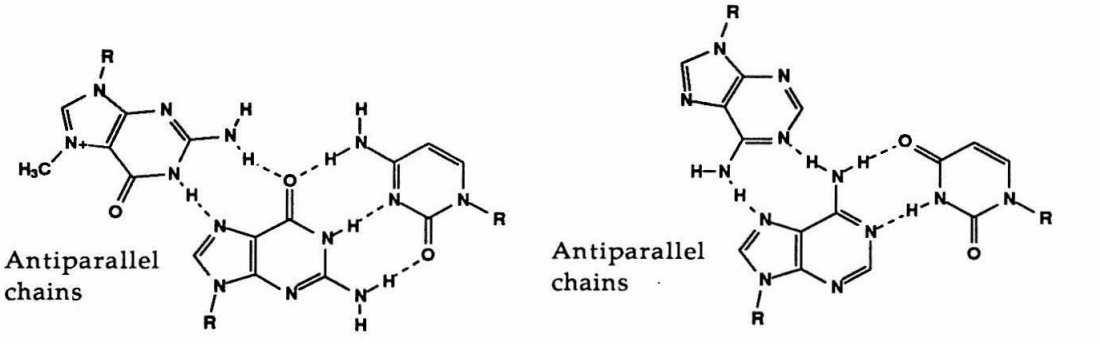


Figure 2.4 Guanine-guanine hydrogen bonding schemes; termed Hoogsteen or reverse Hoogsteen based on their relation to the thymine-adenine hydrogen bonding identified by Hoogsteen in crystals of 9-methyladenine and 1-methylthymine (upper right).[6]

highly conserved in class I type tRNAs with only yeast $tRNA^{Arg}$ having different nucleotides at positions 13, 22 and 46. It was suggested that $tRNA^{Arg}$, which has a U13, A22, and A46 sequence, might maintain the proper folding by forming an A46•A22•U13 triplet, which is nearly isomorphous to the ^{m7}G46•G22•C13 triplet in $tRNA^{phe}$ (Figure 2.5).



^{m7}**G46**•G22C13; found in $tRNA^{Phe}$ **A46**•A22U13; predicted for $tRNA^{Arg}$

Figure 2.5 Base triplets in tRNA

Therefore, the sequence specificity for purine motif triple helix formation can be rationalized based on models for G•GC and A•AU triplets. In addition, models for Hoogsteen and reverse Hoogsteen G•GC triplets have sugar-phosphodiester backbone positions similar to Hoogsteen and reverse Hoogsteen bonded T•AT triplets, suggesting that T•AT and G•GC triplets may be compatible.

Results and Discussion

Orientation. At the initiation of this work, the relative strand orientations in the purine motif triple helix had not been determined. Hogan and co-workers had designed an oligonucleotide to bind via G•GC and A•AT triplets with the third strand parallel to the purine rich strand of the duplex target.[5] However,

Maher et al. were unable to confirm this parallel purine-rich model when they could not detect binding by a purine oligonucleotide to a homopurine target site.[12] Moreover, Griffin and Dervan found that guanine in an otherwise pyrimidine oligonucleotide bound to a TA base pair rather than a GC base pair, suggesting that two different structural motifs may exist for triple helical recognition.[13] Additionally, Kohwi-Shigematsu had proposed a model for a G•GC containing triple helix in a supercoiled plasmid with the two purine strands antiparallel.[4]

Since both a parallel and antiparallel orientation of the third strand had been proposed in the literature for purine motif triple helices, it was necessary to define this parameter before further studies with this structure could be undertaken. Affinity cleaving experiments were designed to address this issue.

We chose as our initial double helical binding site the 19 bp purine sequence, 5′-AGGGAGGGGAGGGGAGGGA-3 within a 39 bp oligonucleotide duplex. Since this is a symmetrical sequence, which is identically read from 3′ to 5′ or 5′ to 3′, it could, in principle, support two triple helical structures with the third strand bound parallel or antiparallel to the WC purine strand. Considering the models for possible base triplet interactions in the purine motif, we reasoned that in addition to guanine recognition of GC base pairs, AT base pairs might be recognized by adenines or thymines in the third strand. Given this specificity hypothesis, oligonucleotides **1** and **2** of sequence composition 5′-T*GGGXGGGGXGGGGXGGGT-3′(where X = T or A) were prepared by automated synthesis with thymidine-EDTA (T*) at each 5′ end such that DNA binding could monitored by the affinity cleaving method (Figure 2.6).[14]

In the presence of Fe(II) and DTT, these oligonucleotides cleaved both strands of the target duplex near the 3′ end of the purine tract (2 μM, pH=7.8, 24°C). The cleavage maximum on each strand was shifted 2-5 bases in the 5′

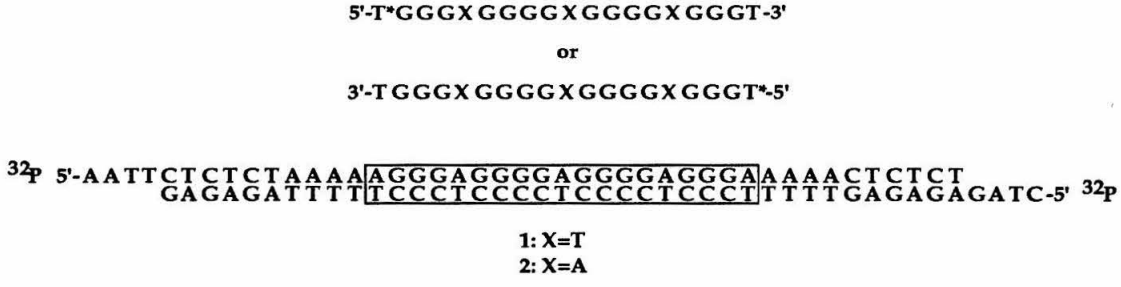


Figure 2.6 Sequences of oligonucleotidesEDTA **1** and **2** with symmetrical duplex target sequence (boxed) for determination of preferred third strand orientation

Figure 2.7 (**A** and **B**) Autoradiograms of 20% denaturing polyacrylamide gels used to separate cleavage products. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM) and Fe(II) (5 μM) with the ^{32}P-labeled oligomer duplex [~ 20,000 cpm] in a solution of TrisOAc pH=7.8 (50 mM), NaCl(10 mM), spermine, and calf thymus DNA(100 μM in base pairs) and then incubating at 24° C for 1 hour. Cleavage reactions were initiated by the addition of dithiothreitol (DTT) (4 mM) and allowed to proceed for 12 hours at 24° C. The reactions were stopped by freezing and lyophilization, and the cleavage products were analyzed by gel electrophoresis. (**A**, Lanes 1 to 8) Duplexes containing 5' end labeled d($A_2T_2(CT)_3A_5G_3AG_4AG_3A_5(CT)_3$) (**B**, Lanes 1 to 8) Duplexes containing 5' end labeled d(CTAG$(AG)_3T_5C_3TC_4TC_4T_5(AG)_3$) (Lanes **A**1 and **B**1) Products of A specific cleavage reactions.[15] (Lanes **A**2 and **B**2) Controls showing intact 5' end labelled duplexes obtained after treatment according to the cleavage reactions in the absence of oligonucleotide-EDTA•Fe(II) (Lanes 3-5) Cleavage products produced by oligonucleotide-EDTA•Fe(II) **1** (Lanes 6-8) Cleavage products produced by oligonucleotide-EDTA•Fe(II) **2** (Lanes 3,6) Spermine at 10 μM (Lanes 4,7) Spermine at 100 μM (Lanes 5,8) Spermine at 1 mM.

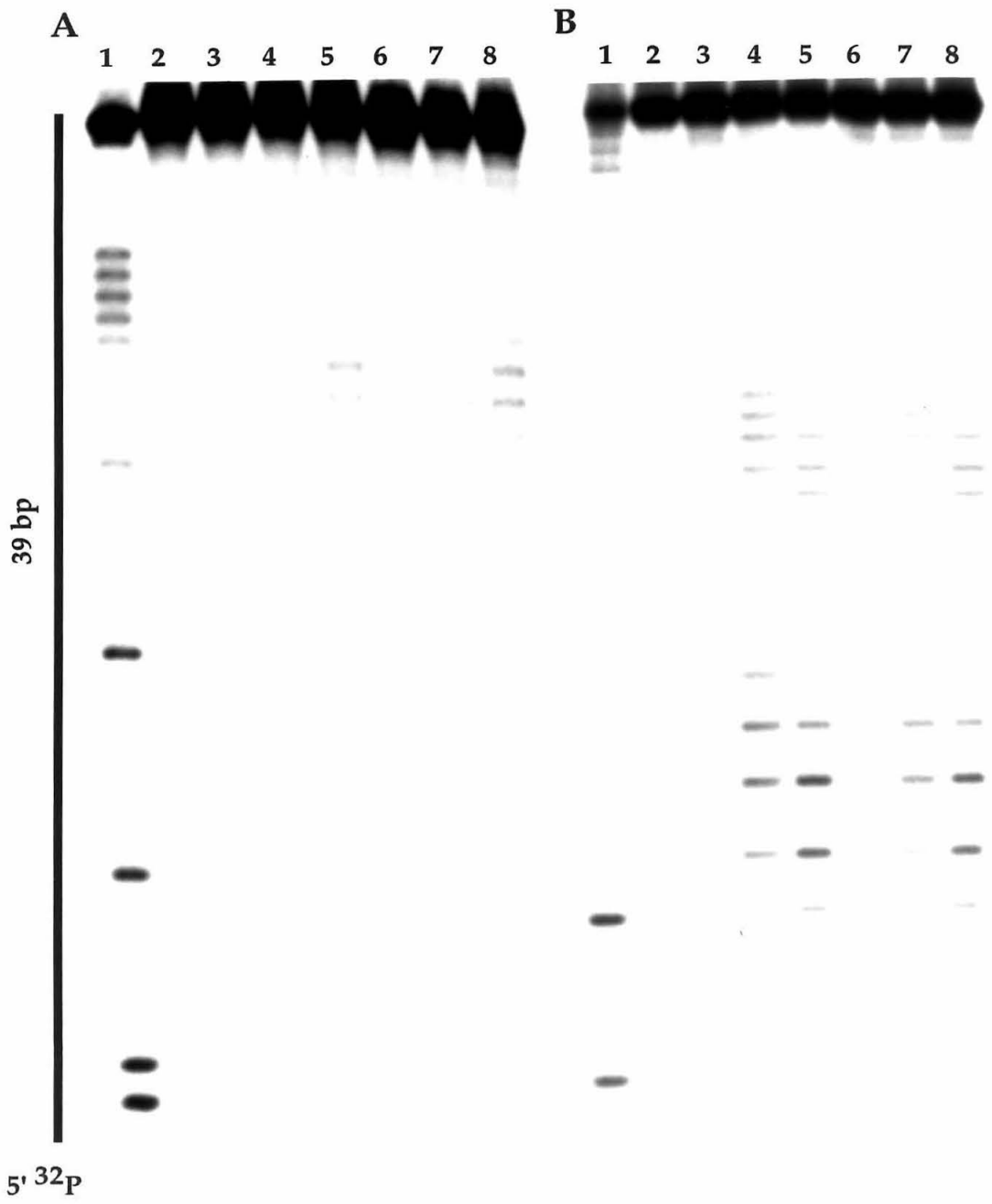

A
1 2 3 4 5 6 7 8
B
1 2 3 4 5 6 7 8
39 bp
5' 32P

Figure 2.8 Histogram of DNA cleavage pattern generated by **1** derived from densitometry of the autoradiogram in Figure 2.7.

direction from the last base pair in the binding site (Figures 2.7 and 2.8). The efficiency of cleavage was dependent on the concentration of the polycation spermine with maximum cleavage occurring at 1 mM.

The location of the cleavage sites and asymmetry of the pattern showed definitively that oligonucleotides **1** and **2** bound in the major groove *antiparallel* to the WC purine strand (Figure 2.8). This result ruled out strand displacement (D-looping) as the mode of recognition. No cleavage products corresponding to a parallel orientation could be detected. Triple helical complexes are known to be stabilized by added divalent and multivalent cations including spermine.[16] The observation that the cleavage efficiencies were dependent on the concentration of spermine is consistent with formation of triple helices by **1** and **2**.

The discovery that these G-rich 19 mers bound the double helix with an antiparallel orientation suggested that the 27 mer oligonucleotide designed by Hogan to bind the c-*myc* promoter may also have bound in an antiparallel fashion. This is possible due to the pseudosymmetry of the c-*myc* promoter sequence (Figure 2.3). In fact, when orientation was explicitly determined, all G-rich oligonucleotides designed to bind the c-*myc* promoter did so with an antiparallel orientation.[6b] Several subsequent reports have confirmed that purine-rich third strands bind antiparallel to the purine strand of duplex DNA.[17]

A•AT and T•AT triplets. From the intensities of the cleavage bands in Figure 2.7, it appeared that **1** and **2** were binding the target site with nearly equal affinities. Because of the sequence design, the formation of G•GC base triplets dominates these binding events. However, from simple modeling of potential hydrogen bonding interactions, it seemed reasonable that the formation of A•AT and T•AT triplets could also contribute to the overall stabilities of the complexes. To test this hypothesis, oligonucleotide **3**, of sequence 5′-

T*GGGCGGGGCGGGGCGGGT-3', was prepared and assayed for binding at the target site along with **1** and **2**. In a triple helix formed by oligonucleotide **3**, three T•AT(**1**) or A•AT(**2**) interactions are replaced by three C•AT interactions. If the binding of these oligonucleotides was controlled only by the formation G•GC base triplets, **1, 2** and **3** would bind with equal affinities.

The comparison of the binding of **1, 2** and **3** was made on a 648 bp *HindIII/SspI* restriction fragment from the plasmid pPBAG19, which has the 19 bp target sequence located 35 bp from the *HindIII* site (Figure 2.9). Cleavage of the 3' end ^{32}P labeled restriction fragment by oligonucleotidesEDTA•Fe **1-3** could only be detected at the target site near the 3' end of the purine tract (0.1 to 1.0 μM oligonucleotide, 0.1 mM spermine, 10 mM NaCl, pH=7.8) (Figure 2.10). This single locus of cleavage bands indicated that the oligonucleotides bound specifically at the target site in this 648 bp fragment. The absolute cleavage efficiencies differed depending on the identity of the base in the third strand opposite the three internal AT base pairs in the binding site. The relative cleavage efficiencies at 1 μM oligonucleotide concentration were 1.0 (X=T), 0.56(X=A) and 0.13 (X=C). Therefore, for this particular sequence and under these conditions, the contribution to the cleavage efficiency from the formation of three A•AT triplets differed from that of three T•AT triplets by less than a factor of two. However, when C was positioned opposite three AT base pairs, the cleavage efficiency decreased by nearly a factor of eight. We believed that the differences in cleavage efficiencies was evidence for specific stabilizing interactions, likely hydrogen bonding, in the T•AT and A•AT triplets. Undoubtedly, the relative affinities of these triplets will depend on sequence composition.

Models for G•GC, A•AT and T•AT triplets. Within the constraints of our experimental data that the third strand was antiparallel to the purine-rich WC

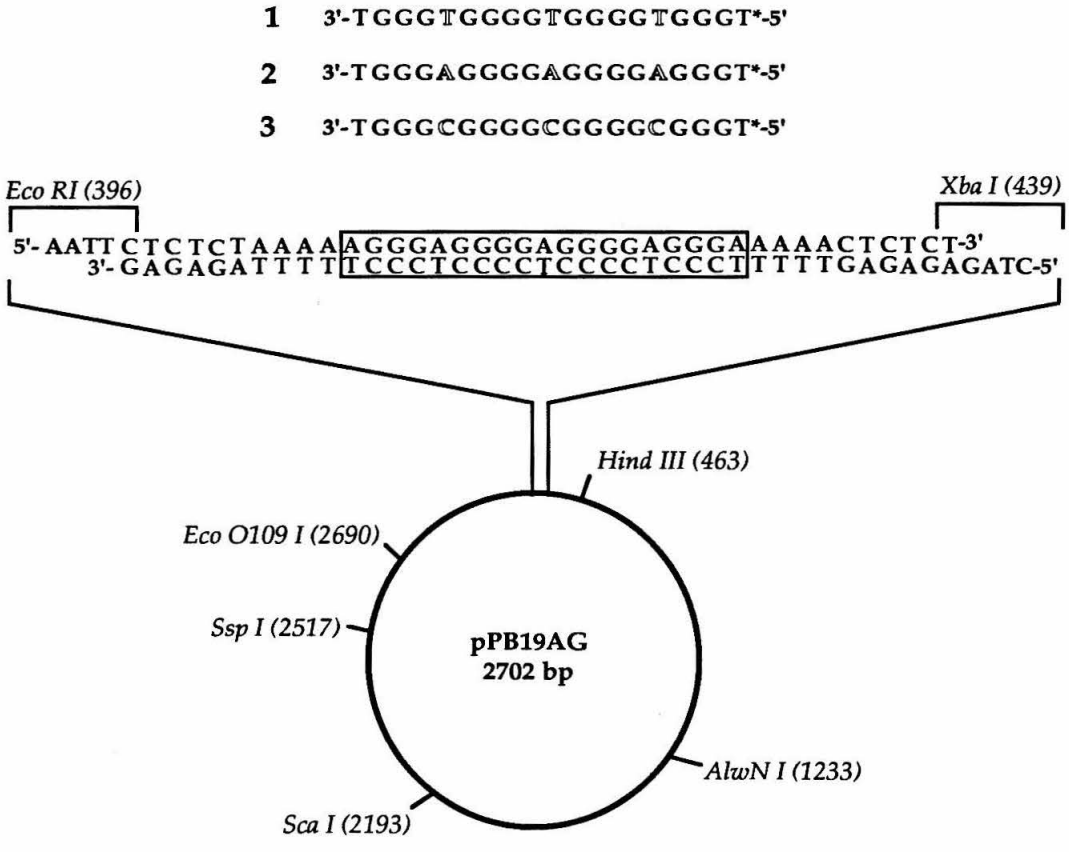
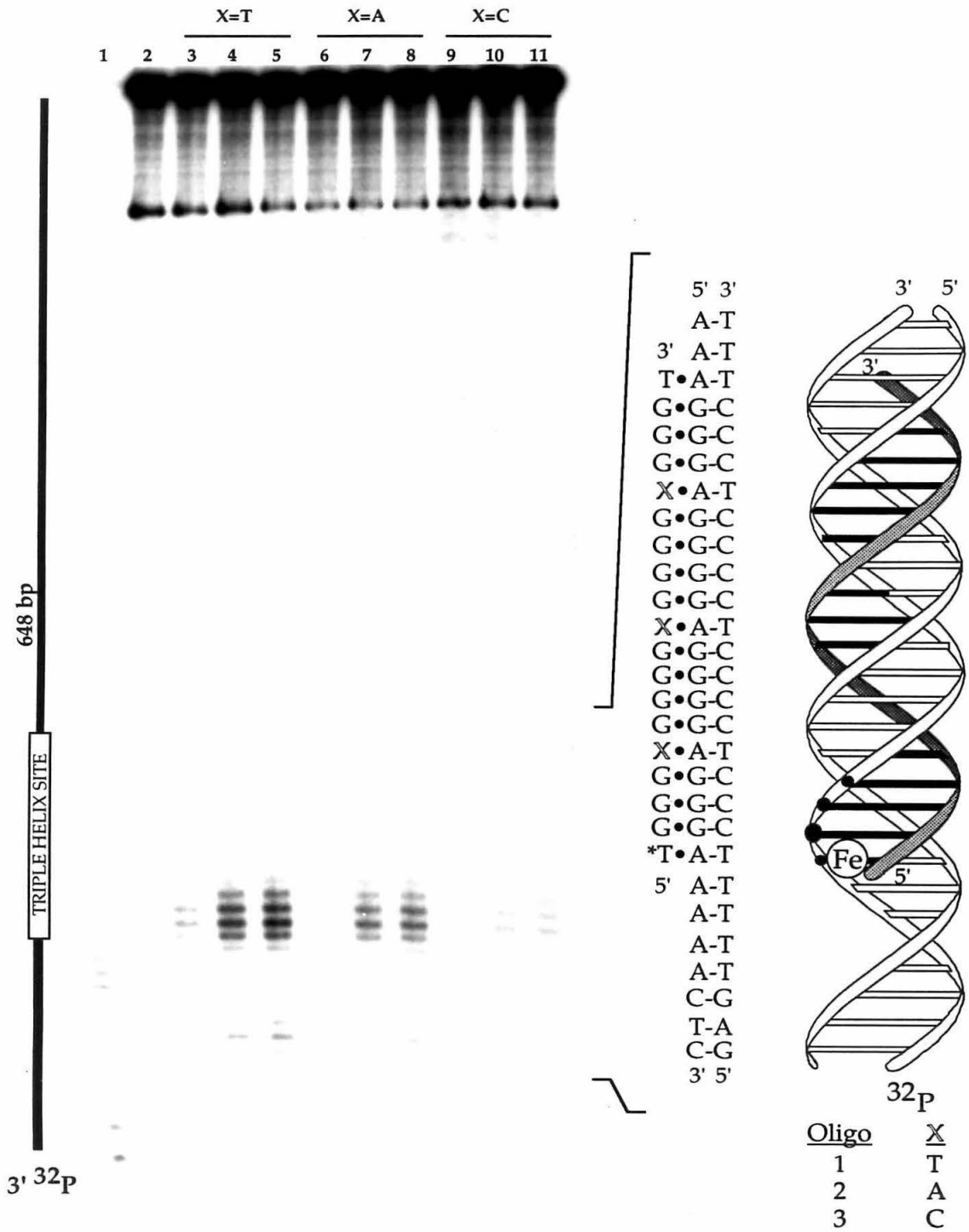
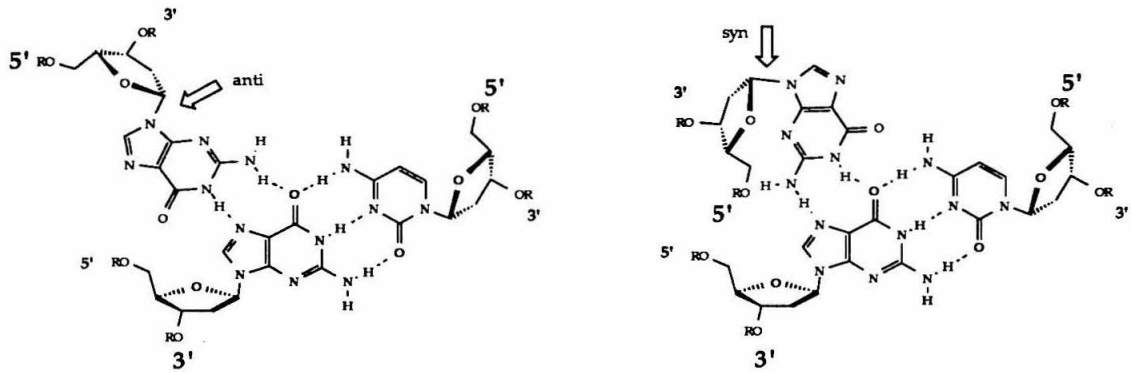


Figure 2.9 (Top) Sequences of oligonucleotidesEDTA **1-3**. (Bottom) Schematic diagram of pPBAG19, constructed from the ligation of a synthetic oligonucleotide duplex into the *EcoRI /XbaI* site of pUC19.

Figure 2.10 (Left) Autoradiogram of an 8% polyacrylamide sequencing gel showing cleavage efficiency dependence on sequence composition of oligonucleotides **1-3**. The cleavage reactions were performed on the HindIII/Ssp I restriction fragment of plasmid pPBAG19 ^{32}P labelled at the 3' end. The reactions were carried out by combining a mixture of oligonucleotide-EDTA and 2.5 equivalents of $Fe(NH_4)_2(SO_4)_2 \bullet 6H_2O$ with labelled restriction fragment [~ 100 nM in bp, ~ 10,000 cpm] in a solution of tris acetate pH=7.8 (50 mM), NaCl (10 mM), spermine (100 μM), and calf thymus DNA (100 μM in bp) and then incubating at 24° C for 1 H. Cleavage reactions were initiated by the addition of dithiothreitol (DTT) (4 mM) and allowed to proceed for 12 H at 24° C. (Lane 1) Products of an A specific cleavage reaction.[15] (Lane 2) Control showing intact 3' end labelled restriction fragment. (Lanes 3-11) Cleavage products produced by oligonucleotides-EDTA•Fe(II) of general sequence 5'-T*$G_3XG_4XG_4XG_3$T-3'. (Lanes 3-5) X=T, (Lanes 6-8) X=A (Lanes 9-11) X=C. Concentrations of oligonucleotide-EDTA•Fe(II) **1-3**: (Lanes 3,6,9) **1-3** at 0.1 μM, (Lanes 4,7,10) **1-3** at 0.5 μM, (Lanes 5,8,11) **1-3** at 1.0 μM. Cleavage efficiencies were quantitated using storage phosphorimaging plates with a Molecular Dynamics 400S PhosphorImager. **(Right)** Ribbon model and sequence of local triple helical complex between oligonucleotides-EDTA•Fe(II) and the target sequence. Circles represent backbone positions of cleavage. Size of circles represent extent of cleavage.


X=T
X=A
X=C
1 2 3 4 5 6 7 8 9 10 11
648 bp
TRIPLE HELIX SITE
3' 32P
5' 3'
A-T
3' A-T
T•A-T
G•G-C
G•G-C
G•G-C
X•A-T
G•G-C
G•G-C
G•G-C
G•G-C
X•A-T
G•G-C
G•G-C
G•G-C
G•G-C
X•A-T
G•G-C
G•G-C
G•G-C
*T•A-T
5' A-T
A-T
A-T
A-T
C-G
T-A
C-G
3' 5'
3' 5'
3'
Fe
5'
32P
Oligo X
1 T
2 A
3 C

strand, models of possible hydrogen bonding patterns for base triplets place the phosphodiester-deoxyribose backbone in different locations in the major groove depending on whether the base in the third strand is in the *anti* or *syn* conformation (Figure 2.11). If the third strand base were in the *anti* conformation, the likely hydrogen bonding interaction would be reverse



anti G•GC (reverse Hoogsteen) syn G•GC (Hoogsteen)

Figure 2.11 Models for G•GC triplets. Given the third strand is antiparallel to the purine strand of the duplex, likely hydrogen bonding patterns depend on the third strand guanosine glycosidic bond conformation.

Hoogsteen generating a structure with the phosphodiester-deoxyribose backbone centrally located in the major groove of the double helix. The *syn* conformation required Hoogsteen hydrogen bonding and placed the backbone in a similar location in the major groove as found with the pyrimidine motif. Although both Hoogsteen and reverse Hoogsteen models for the G•GC and T•AT triplets could be drawn where the third strand base was bound to the underlying base pair by two hydrogen bonds, only the reverse Hoogsteen model for A•AT had two hydrogen bonds (Figure 2.12). Also, although both *syn* and *anti* conformations are adopted by purine nucleosides, the *anti* conformation greatly dominates for

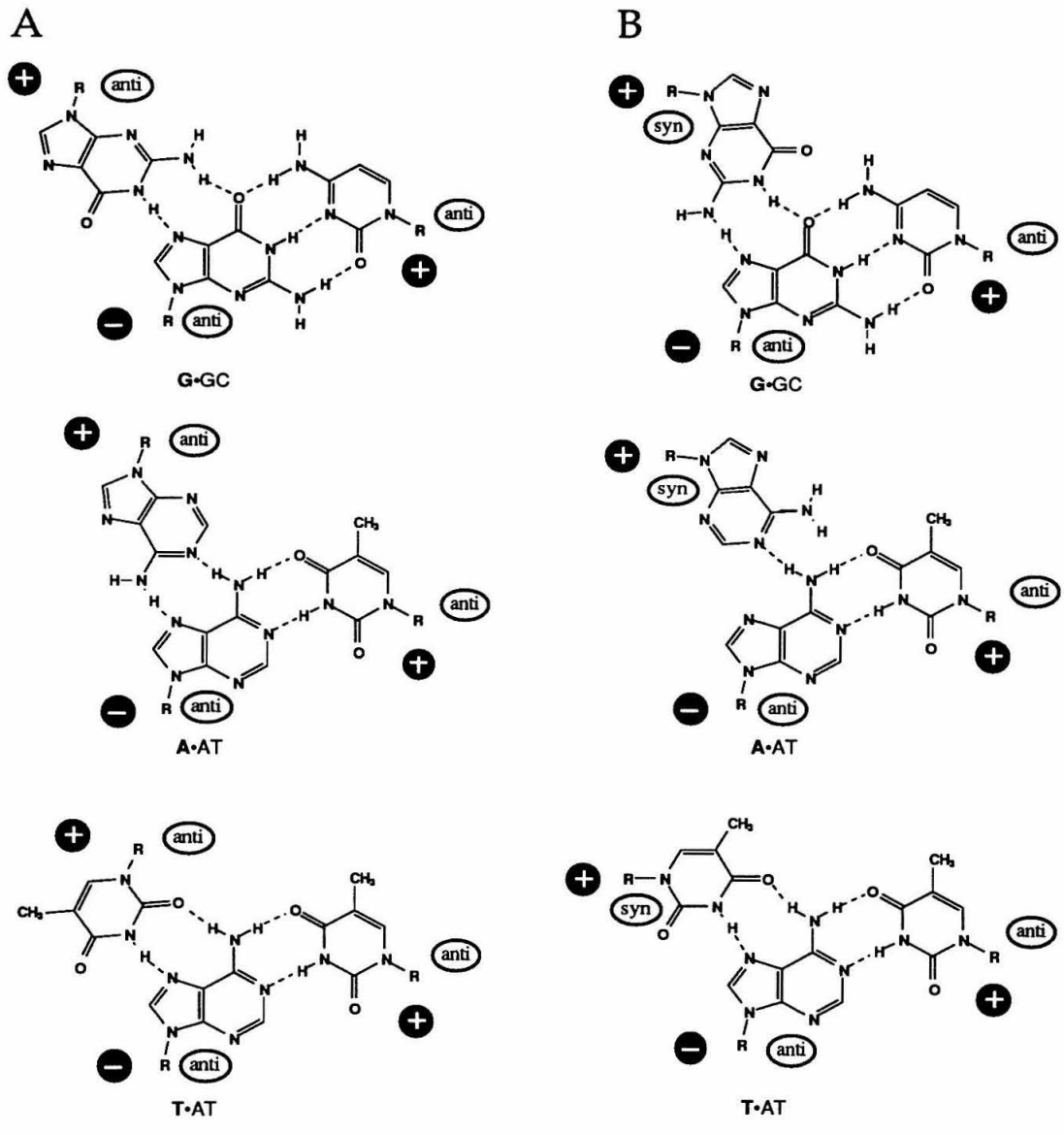


Figure 2.12 (**A**) Models for G•GC , A•AT and T•AT triplets within a triple helix motif where the third strand is antiparallel to the purine WC strand and bases are in the *anti* conformation (reverse Hoogsteen triplets). (**B**) Triplet models where the third strand is antiparallel to the WC purine strand and the bases are in the *syn* conformation (Hoogsteen triplets). Plus and minus indicate relative polarities of the strands.

pyrimidine nucleosides, making a *syn* T•AT triplet less likely.[18] Hence, we tentatively favored the model with *anti* conformations and reverse Hoogsteen hydrogen bonding, placing the phosphodiester-deoxyribose backbone near the center of the major groove. This triple helix structure differed from the pyrimidine motif not only in triplet specificities but also in the location and orientation of the third strand and, therefore, constituted a second structural motif for triple helix formation.[19]

Groove Location of EDTA•Fe. When an oligonucleotide binds in the major groove of double helical DNA, this groove is partitioned into two new grooves: one between the third strand and the purine strand of the double helix (groove A in Figure 2.13) and one between the third strand and the pyrimidine strand (groove B in Figure 2.13). The two models for T•AT triplets (Hoogsteen and reverse Hoogsteen) place the third strand thymine C5 methyl in one or the other of these two grooves. The reverse Hoogsteen T•AT places the C5 methyl in groove A, the Hoogsteen T•AT has it in groove B.

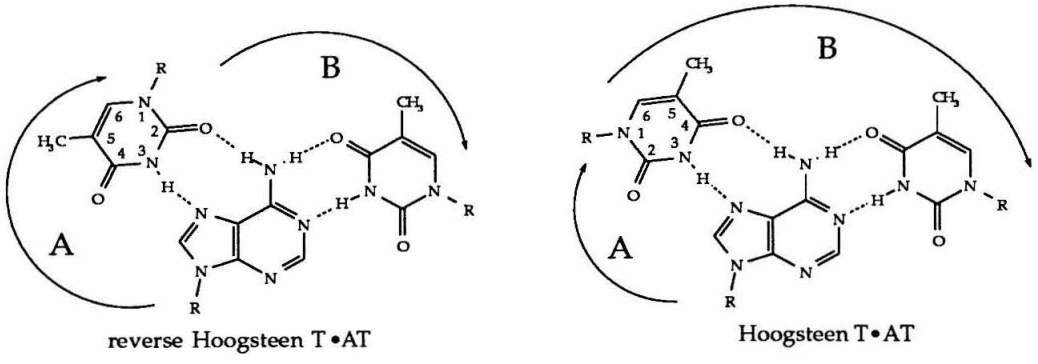


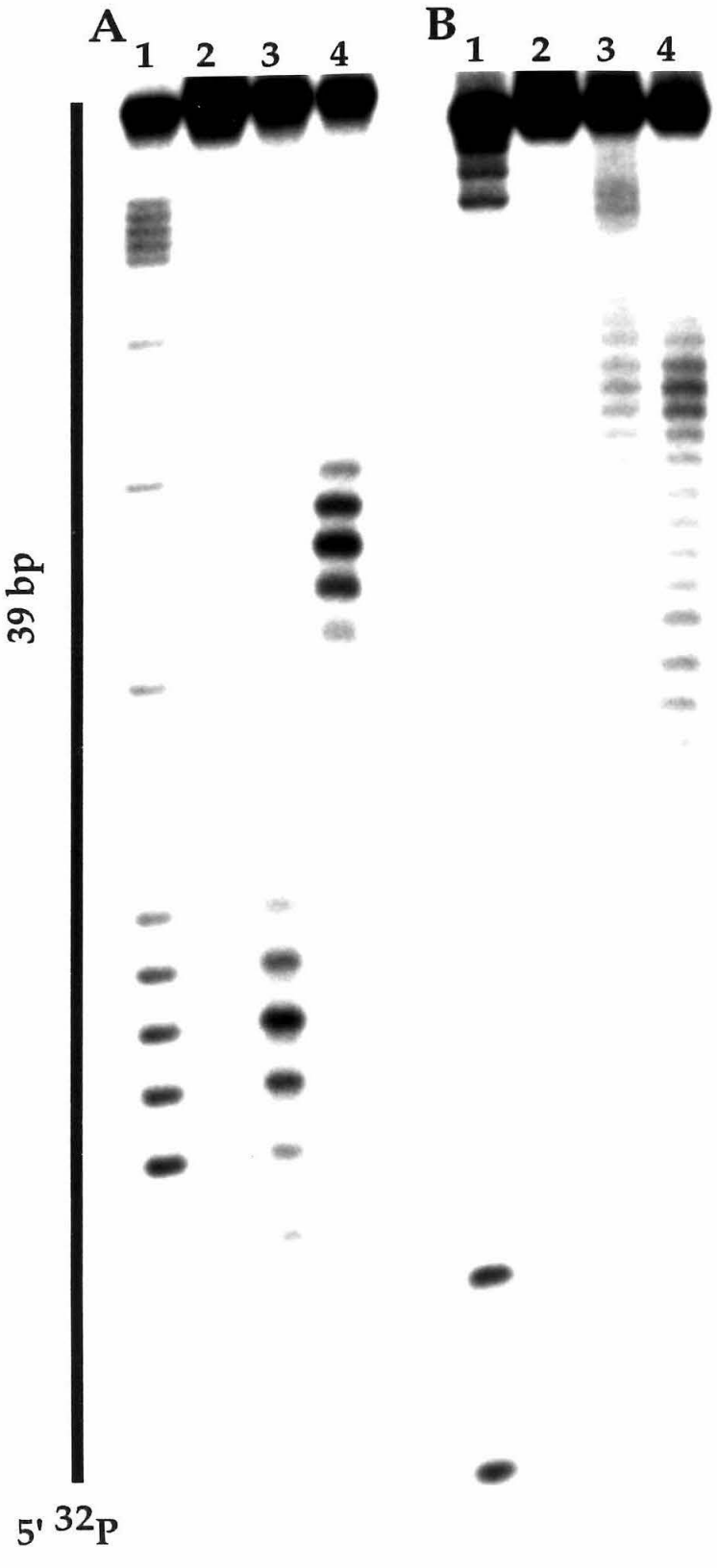
Figure 2.13 Triple helix formation partitions the duplex major groove into two new grooves (A and B). The reverse Hoogsteen T•AT triplet places the third strand thymine C5 methyl in groove A. The Hoogsteen T•AT triplet places this group in groove B.

Moser and Dervan have shown that pyrimidine oligonucleotides containing thymidineEDTA at internal positions cleave only the pyrimidine strand of the double helix efficiently; the purine strand is protected from cleavage.[16] Strobel and Dervan have also shown that pyrimidine oligonucleotides with T* at their 3′ ends only cleave the pyrimidine strand efficiently.[20] These observations are resultant of the right-handed helical nature of the triple helix and the Hoogsteen T*•AT triplet in the pyrimidine motif, which places the EDTA•Fe, linked to the base at the C5 position, in groove B between the third strand and the pyrimidine strand of the double helix. The purine strand is protected from cleavage by a steric blockade provided by the third strand sugar-phosphate backbone.

Our preferred model for the triplets in the purine motif included a reverse Hoogsteen T•AT, placing the C5 methyl in groove A, between the third strand and the purine strand of the duplex. Given the results discussed, this model predicted that for a purine-rich oligonucleotide with T* at an internal position or at the 3′ end, only the purine strand of the duplex would be cleaved efficiently; i.e., the pyrimidine strand would be protected.

To determine if this prediction was correct, oligonucleotidesEDTA•Fe **4** and **5** were prepared which differed from **1** only in the position of the thymidineEDTA (T*). The cleavage generated by oligonucleotide **4**, with T* at the center position, and oligonucleotide **5**, with T* at the 3′ position was analyzed on a 39 bp 5′ end ^{32}P labeled oligonucleotide duplex. The autoradiogram of the resulting cleavage products is shown in Figure 2.14. Histograms of the cleavage patterns for **4** and **5** are shown in Figure 2.15 along with that generated by **1**. The cleavage sites for **4** and **5** are located at positions consistent with an antiparallel orientation of the third strand, with cleavage occurring at the 5′ end of the purinetract for **5**, which has T* at its 3′ end. The major cleavage sites for both **4**

Figure 2.14 (**A** and **B**) Autoradiograms of 20% denaturing polyacrylamide gels used to separate cleavage products. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM) and Fe(II) (5 μM) with the ^{32}P-labelled oligomer duplex [~ 50,000 cpm] in a solution of tris-acetate pH=7.4 (50 mM), NaCl(10 mM), spermine (1 mM), and calf thymus DNA(100 μM in base pairs) and then incubating at 24° C for 1 hour. Cleavage reactions were initiated by the addition of dithiothreitol (DTT) (4 mM) and allowed to proceed for 12 hours at 24° C. The reactions were stopped by freezing and lyophilization, and the cleavage products were analyzed by gel electrophoresis. (**A**, Lanes 1 to 4) Duplexes containing 5' end labelled d($A_2T_2(CT)_3A_5G_3AG_4AG_3A_5(CT)_3$) (**B**, Lanes 1 to 4) Duplexes containing 5' end labelled d($CTAG(AG)_3T_5C_3TC_4TC_4T_5(AG)_3$) (Lanes **A**1 and **B**1) Products of A specific cleavage reactions.[15] (Lanes **A**2 and **B**2) Controls showing intact 5' end labelled duplexes. (Lanes **A**3 and **B**3) Cleavage products produced by oligonucleotide-EDTA•Fe(II) **5**. (Lanes **A**4 and **B**4) Cleavage products produced by oligonucleotide-EDTA•Fe(II) **4**.


A
1
2
3
4
B
1
2
3
4
39 bp
5' 32P

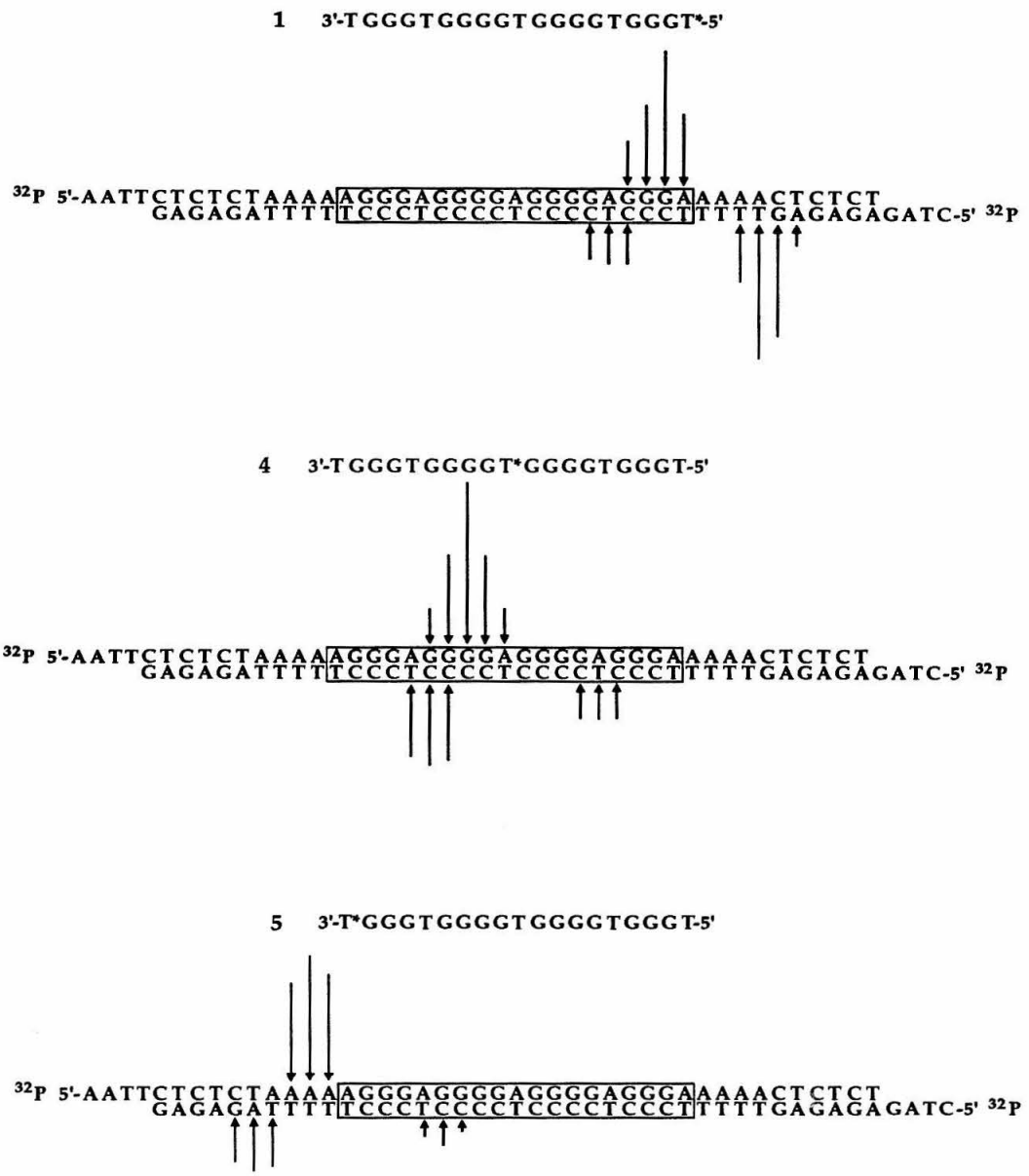


Figure 2.15 Sequences of oligonucleotidesEDTA **1, 4** and **5**, where T* indicates location of thymidineEDTA. Histograms of DNA cleavage patterns generated by **1, 4** and **5** on the 39 bp oligonucleotide duplex.

and **5** are on the purine strand 2-3 bases to the 5′ side of the AT base pair bound by T*. Less efficient cleavage occurs on the pyrimidine strand. Importantly, cleavage on the pyrimidine strand 4-5 bases to the 5′ side of the location of the T*•AT triplet is much less efficient for **4** and **5** than for **1**. These results are consistent with protection of the pyrimidine stand from cleavage as would be predicted for a reverse Hoogsteen T*•AT triplet.

Figure 2.16 shows the cleavage sites generated by **4** superimposed on ribbon models depicting a purine motif triple helix with reverse Hoogsteen or Hoogsteen triplets. The main sites are clustered on one face of the helix proximal to the groove between the third strand and the purine strand of the duplex. This cleavage pattern is clearly most consistent with the model with a reverse Hoogsteen T*•AT triplet with the T* in the *anti* conformation placing the EDTA•Fe in this groove. If the T*•AT triplet were Hoogsteen, the main cleavage sites would be expected on the pyrimidine strand shifted 5′ 2-3 bases from the position of the T*, which is not the observed result. These results are consistent with our preferred model for base triplets in the purine motif with third strand bases in the *anti* conformation and reverse Hoogsteen hydrogen bonding.

NMR Studies. Subsequent to these studies, Radhakrishnan et al. reported an NMR spectroscopic study of an intramolecular purine motif triple helix containing T•AT and G•GC triplets.[17c] This study showed that the third strand bases had *anti* glycosidic bond conformations and the base triplet alignments were reverse Hoogsteen, supporting our preferred model. For instance, no strong NOEs were detected from the sugar H1′ to purine H8 protons indicating that no purines in the structure were in the *syn* conformation. Also, NOEs were observed between guanosine imino protons in the third strand and the guanosine H8 and H2′, H2″ sugar protons in the duplex, consistent with

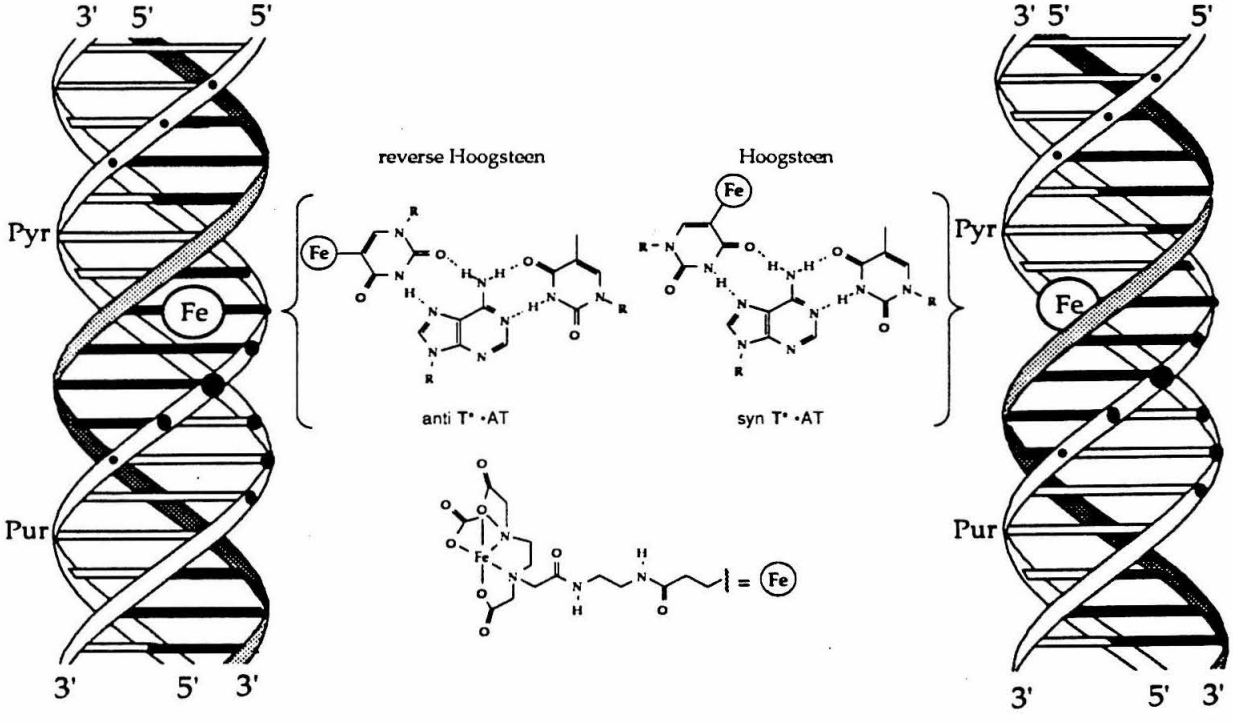


Figure 2.16 Ribbon models of the triple helix formed by **4** with a reverse Hoogsteen (left) or Hoogsteen T*•AT triplet (right). The reverse Hoogsteen triplet requires an *anti* conformation for the T*; the Hoogsteen triplet requires a *syn* T*. Cleavage sites are designated by black circles with the size of the circle indicating the relative cleavage efficiency at that site. The cleavage sites are most consistent with the model on the left.

reverse Hoogsteen pairing where these protons are close in space (see Figure 2.11). Recently, this group also published a solution structure of an intramolecular purine motif triple helix using a combined NMR and molecular dynamics approach (see Figure 2.17 for sequence).[17g] Figure 2.18 depicts the G•GC and T•AT triplets from this NMR structure, showing the reverse Hoogsteen hydrogen bonding interactions and *anti* glycosidic bond conformations. A view perpendicular to the helical axis looking into the groove between the third strand and the purine strand of the duplex is given in Figure 2.19. Highlighted are the C5 methyl of the central thymidine in the third strand with the nucleotide positions cleaved most efficiently by the central T* in the related triple helix formed by oligonucleotideEDTA **4**. The proximity of the thymine methyl group and these nucleotides in this NMR structure correlates well with the positions cleaved by T* in the structure formed by the binding of **4** (Figures 2.15 and 2.16). Another noteworthy feature of this structure is the discontinuity in the third strand backbone conformation at the 5′-TpG-3′ steps resultant from the fact that the T•AT and G•GC triplets are not completely isomorphous. This deformation may indicate that alternating 5′-$(ApG)_n$-3′ sequences will be difficult to target by purine motif triple helix formation (Figure 2.20).

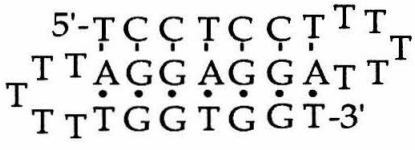


Figure 2.17 Sequence of the intramolecular purine motif triple helix for which a solution structure was recently determined by a combined NMR and molecular dynamics approach.[17g]

Figure 2.18 (overleaf) G•GC and T•AT triplets from the Radhakrishnan and Patel NMR structure.[17g] Duplex base pairs are blue; third strand bases are red.

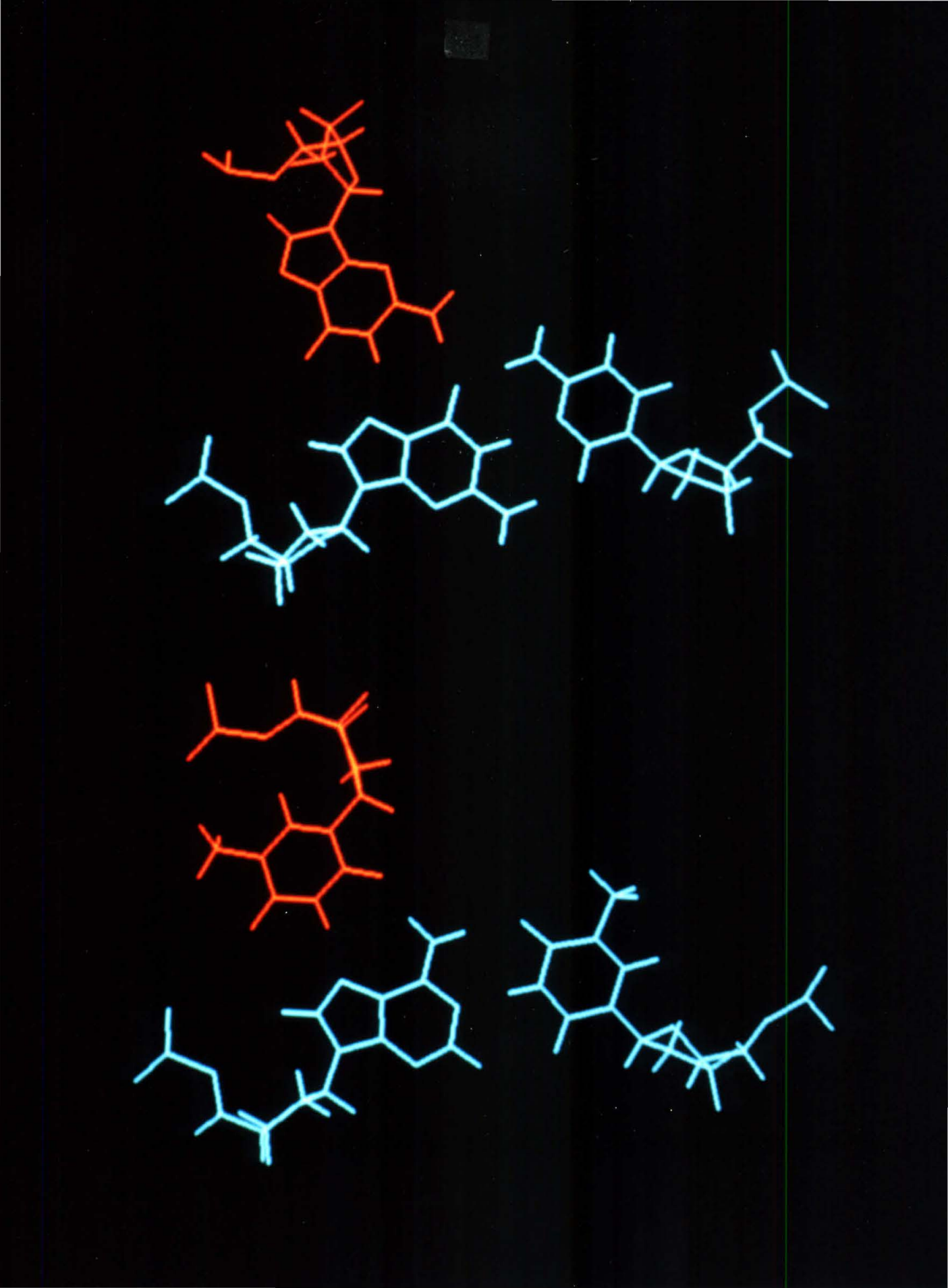


Figure 2.19 View of triple helix perpendicular to the helical axis looking into the groove between the third strand and the purine strand of the duplex.[17g] The duplex strands are colored blue and the third strand is red. The C5 methyl of the central thymidine in the third strand is highlighted in green along with the nucleotide positions (yellow) cleaved most efficiently by the central thymidineEDTA in the related triple helix formed by **4** (Figures 2.15 and 2.16).

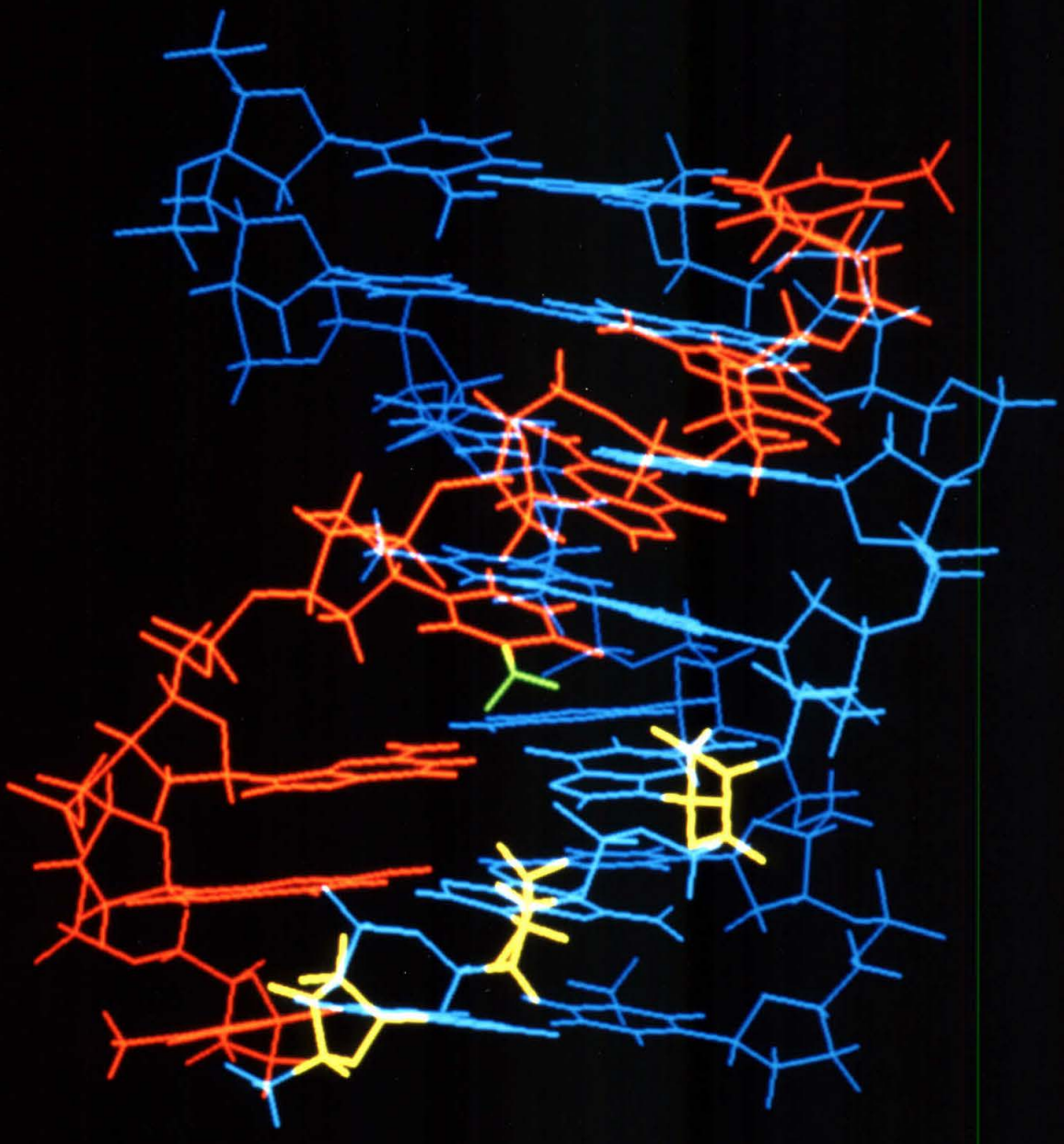
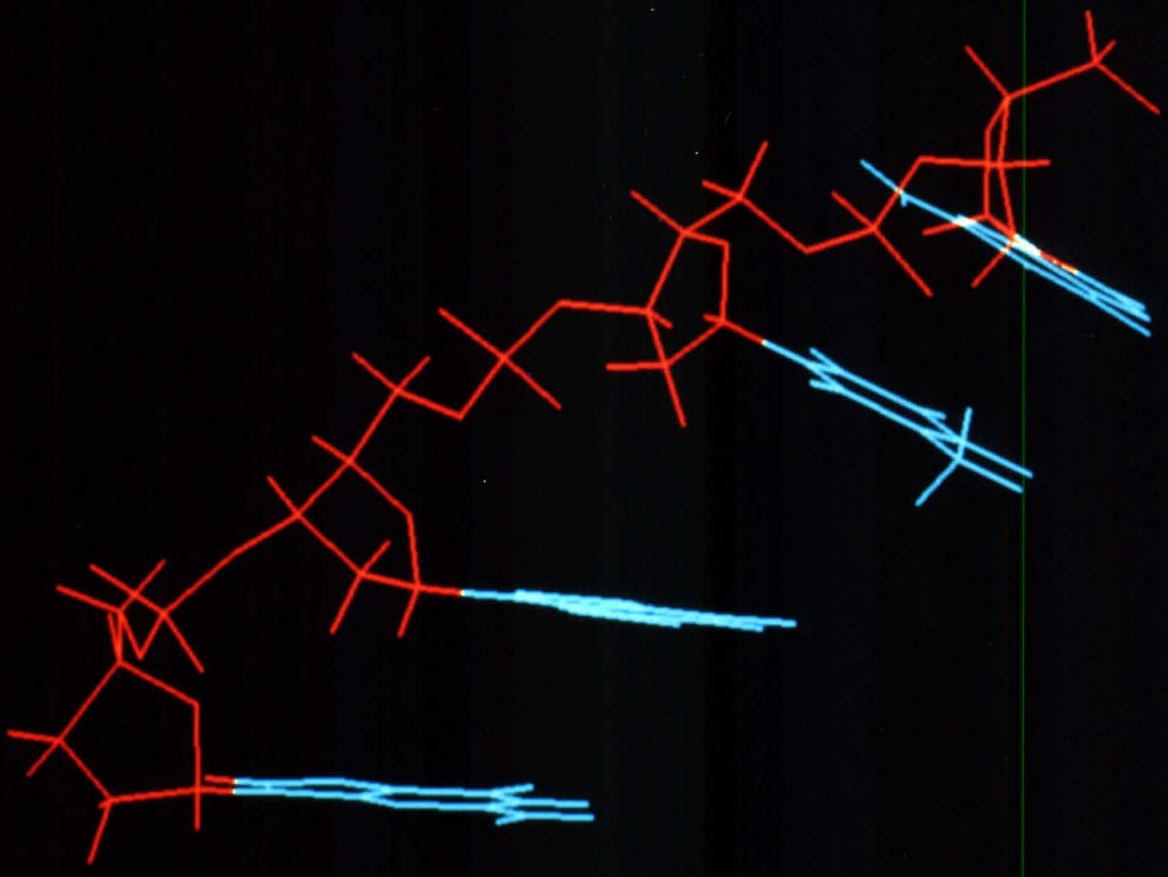


Figure 2.20 View of the third strand 5′-GTGG-3′ segment showing the backbone conformation discontinuity resulting in poor stacking at 5′-TpG-3′ step.[17g] The phosphodiester deoxyribose backbone is colored red; the bases are blue.



Site-specific double strand cleavage of plasmid size DNA. Oligonucleotide **1** was shown to bind to its 19 bp target sequence within a 39 bp oligonucleotide duplex (Figure 2.7) and a 648 bp restriction fragment (Figure 2.10). To determine if this G-rich oligonucleotide was capable of binding specifically to its double helical target sequence within the context of a larger piece of DNA, a uniquely end labeled 2.5 kbp fragment was generated from plasmid pPBAG19 (Figure 2.21). This fragment was obtained by digestion with Eco O109 I followed by 3′ end labeling and digestion with SspI. This procedure generates a linear labeled fragment with the 19 bp target sequence 0.4 kb from the ^{32}P labeled end.

Oligonucleotide **1** (0.6-5 μM, 10 mM NaCl, 0.1 mM spermine, pH=7.8) induced a double strand break in this fragment generating one major cleavage product 0.4 kbp in length, as determined by comparison to DNA size markers (Figure 2.22). No other cleavage bands could be detected. Therefore, **1** bound sequence specifically to a single site within 2.5 kbp of double helical DNA. The ability of pyrimidine oligonucleotides to bind sequence specifically via triple helix formation has been exploited to target single sites in megabase fragments of DNA.[25] It remains to be seen if purine-rich oligonucleotides can be similarly employed in very large DNA fragments.

The effect of added ions. The stabilities of pyrimidine motif triple helices are greatly dependent on solution conditions including the presence of multivalent and/or divalent cations.[17] Early examples of the formation of triple helices containing pur•pur•pyr triplets also required the addition of cations, usually Mg^{+2}.[4,5] The formation of a triple helix by oligonucleotide **1** was shown to be dependent on the concentration of the multivalent cation spermine (Figure 2.7). To determine if other cations might serve to stabilize this triple helix, the

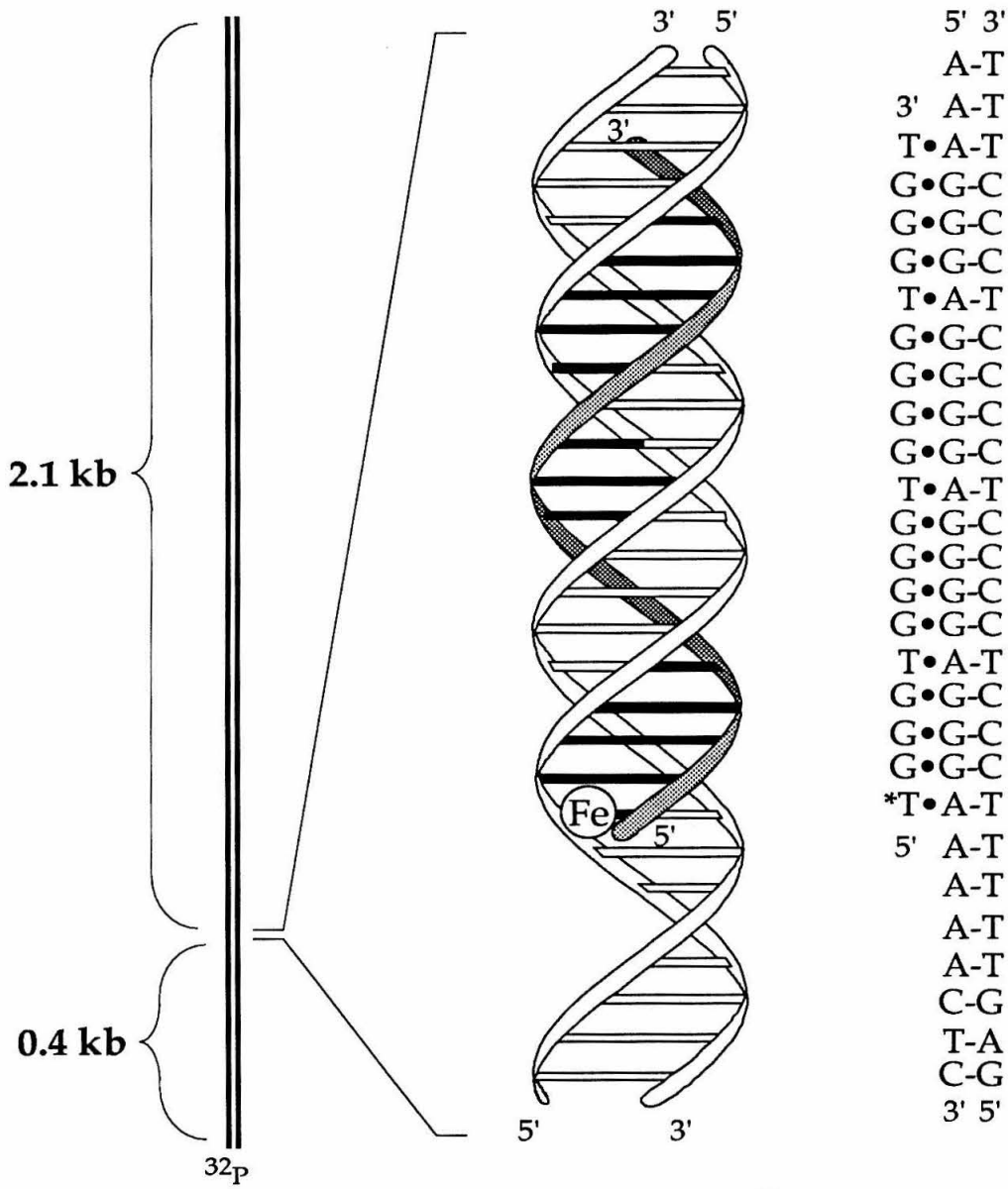
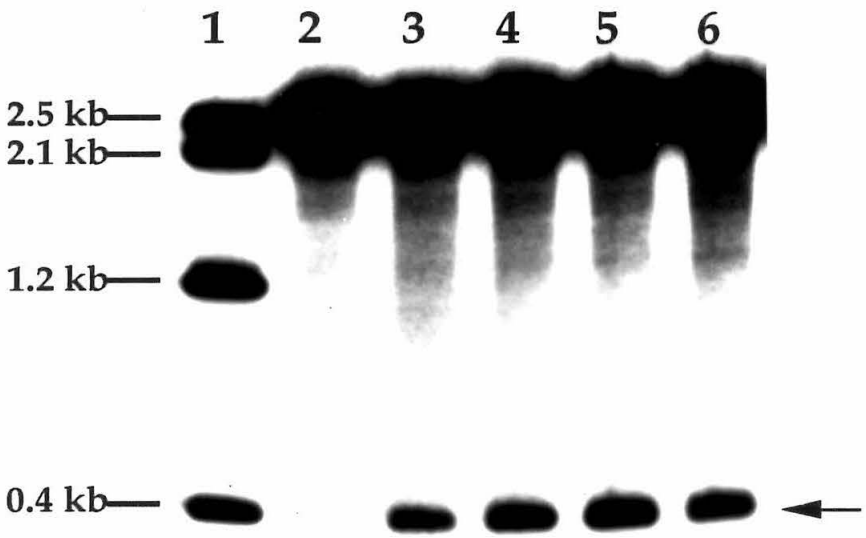


Figure 2.21 (Left) Location of triple helix site in a ^{32}P labeled 2.5 kb fragment from plasmid pPBAG19 (see Fig. 2.8). (Right) A schematic diagram and sequence of the triple helix.

Figure 2.22 Autoradiogram of 1% agarose gel used to separate double strand cleavage products. Plasmid pPBAG19 was linearized with Eco O109 I and 3' end labeled with ^{32}P. The resulting fragment was cleaved with Ssp I generating a 2.5 kb single end labeled fragment. This labeled fragment was dissolved in buffer containing NaCl, TrisOAc, spermine and calf thymus DNA and was mixed with oligonucleotideEDTA **1**, previously equilibrated with 2.5 eq. of Fe(II). After incubation for 1 H at 24°C, the reactions were initiated by the addition of DTT. The final concentrations in the reactions were as follows: 50 mM TrisOAc, 10 mM NaCl, 100 μM calf thymus DNA, and 4 mM DTT. The cleavage reactions were allowed to proceed for 12 H, at which time they were stopped by precipitation with ethanol and the cleavage products were analyzed by gel electrophoresis. (Lane 1) DNA size markers obtained by digestion of the end labeled Eco O109 I/Ssp fragment from pPBAG19 with Sca I, AlwN I and Xba I: 2529 (undigested DNA), 2205, 1245 and 451 bp. (Lane 2) Control containing no oligonucleotide-EDTA•Fe(II); (Lanes 3-6) oligonucleotide-EDTA•Fe(II) **1**; (Lane 3) **1** at 0.62 μM (Lane 4) **1** at 1.2 μM (Lane 5) **1** at 2.5 μM (Lane 6) **1** at 5.0 μM. Arrow at right indicates single cleavage band.


1
2
3
4
5
6
2.5 kb
2.1 kb
1.2 kb
0.4 kb

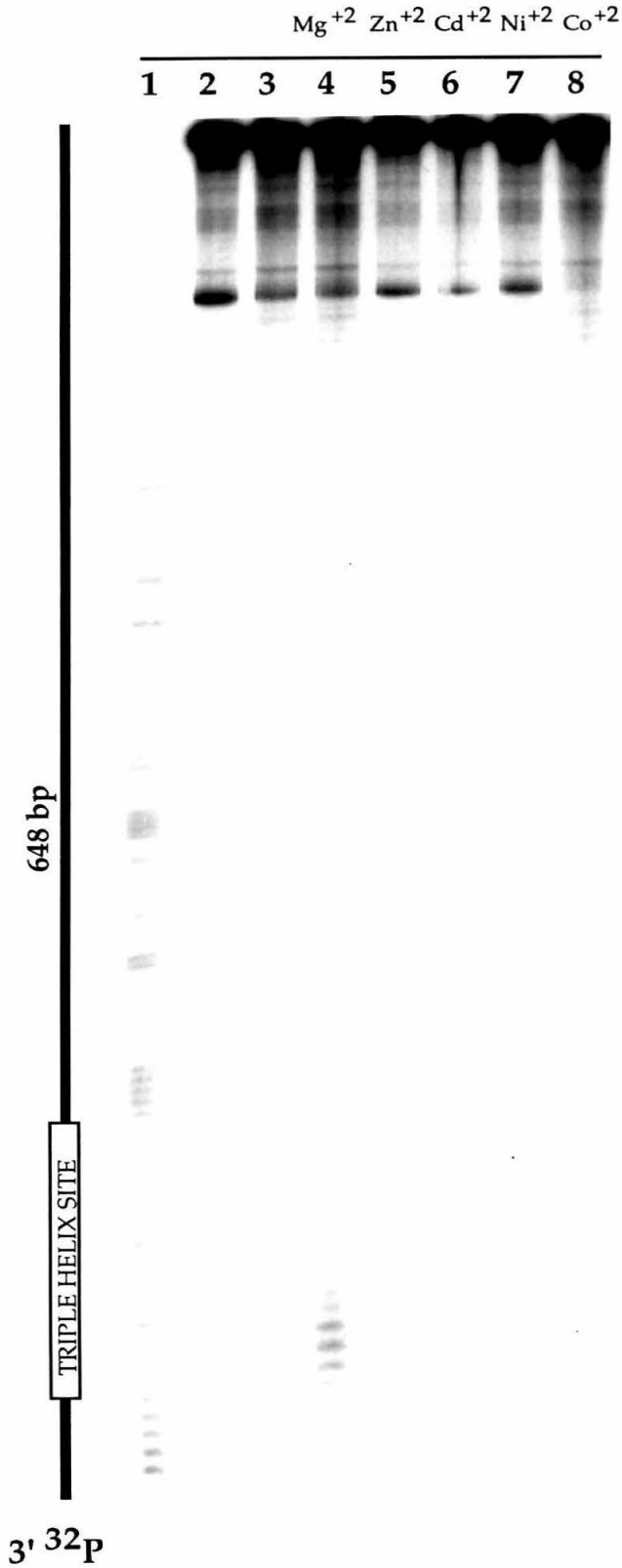
ability of **1** to cleave at its 19 bp target sequence in the *HindIII/SspI* fragment from pPBAG19 (see Fig. 2.10 for sequences) was tested in the absence of spermine and the presence of 5 mM Mg^{+2}, Zn^{+2}, Cd^{+2}, Ni^{+2}, and Co^{+2}.

No cleavage could be detected in the absence of added cations (Figure 2.23, lane 3). Specific cleavage was detected in the presence of Mg^{+2} indicating that both spermine and Mg^{+2} serve to stabilize the triple helix formed by **1** (Figure 2.23, lane 4). However, cleavage could not be detected in the presence of the other divalent ions tested. This may indicate that these metal ions are not capable of stabilizing the triple helix. However, since these ions have been shown to decrease the efficiency of the EDTA•Fe cleavage reaction, no firm conclusions should be drawn on their ability to stabilize the purine motif triple helix based on these experiments alone.[22] In fact, reports have appeared in the recent literature on a Zn^{+2} stabilized triple helical structure in supercoiled plasmid DNA containing alternating A•AT and G•GC triplets.[23]

Increasing the Na^+ concentration in the presence of spermine decreases the equilibrium binding constant for a triple helix forming pyrimidine oligonucleotide.[24] This result has been rationalized by invoking a competition between Na^+ and spermine for binding sites on the triple helix with Na^+ unable to stabilize the triple helix as efficiently.

For guanine-rich oligonucleotides, the presence of high concentrations of monovalent ions could be even more detrimental to triple helix formation since these oligonucleotides form stable secondary structures in the presence of certain monovalent ions.[25] These structures are stabilized by the G-quartet hydrogen bonding arrays which stack in a helical fashion with cations located between the G-quartets, K^+ being more efficient than Na^+ at stabilizing these structures (Figure 2.24). The formation of Na^+or K^+ stabilized G-quartet structures by G-rich oligonucleotides potentially poses an obstacle to purine motif triple helix

Figure 2.23 Autoradiogram of an 8% polyacrylamide sequencing gel used to separate cleavage products. The cleavage reactions were performed on the HindIII/Ssp I restriction fragment of plasmid pPBAG19 ^{32}P labeled at the 3' end. The reactions were carried out by combining a mixture of oligonucleotide-EDTA **1** and 2.5 equivalents of $Fe(NH_4)_2(SO_4)_2 \bullet 6H_2O$ with labeled restriction fragment [~ 100 nM in bp, ~ 30,000 cpm] in a solution of TrisOAc pH=7.4 (50 mM), NaCl (10 mM), MCl_2, where M=Mg, Zn, Cd, Ni, or Co (5 mM) and calf thymus DNA (100 μM in bp) and then incubating at 24° C for 1 H. Cleavage reactions were initiated by the addition of dithiothreitol (DTT) (4 mM) and allowed to proceed for 12 H at 24° C. (Lane 1) Products of an A specific cleavage reaction.[15] (Lane 2) Control showing intact 3' end labelled restriction fragment. (Lanes 3-8) Cleavage products produced by oligonucleotides-EDTA•Fe(II) **1** (Lane 3) No added MCl_2 (Lane 4) M=Mg (Lane 5) M=Zn (Lane 6) M=Cd (Lane 7) M=Ni (Lane 8) M=Co.


Mg+2 Zn+2 Cd+2 Ni+2 Co+2
1 2 3 4 5 6 7 8
648 bp
TRIPLE HELIX SITE
3' 32P

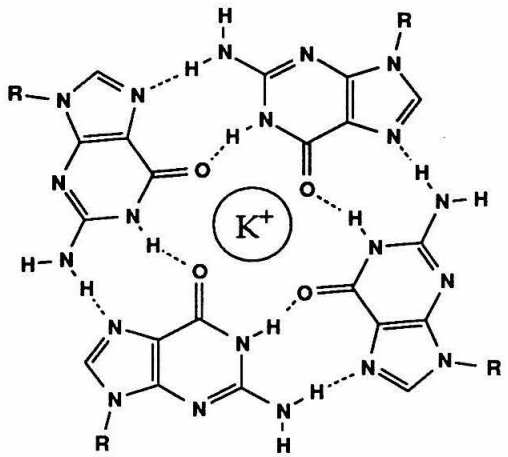


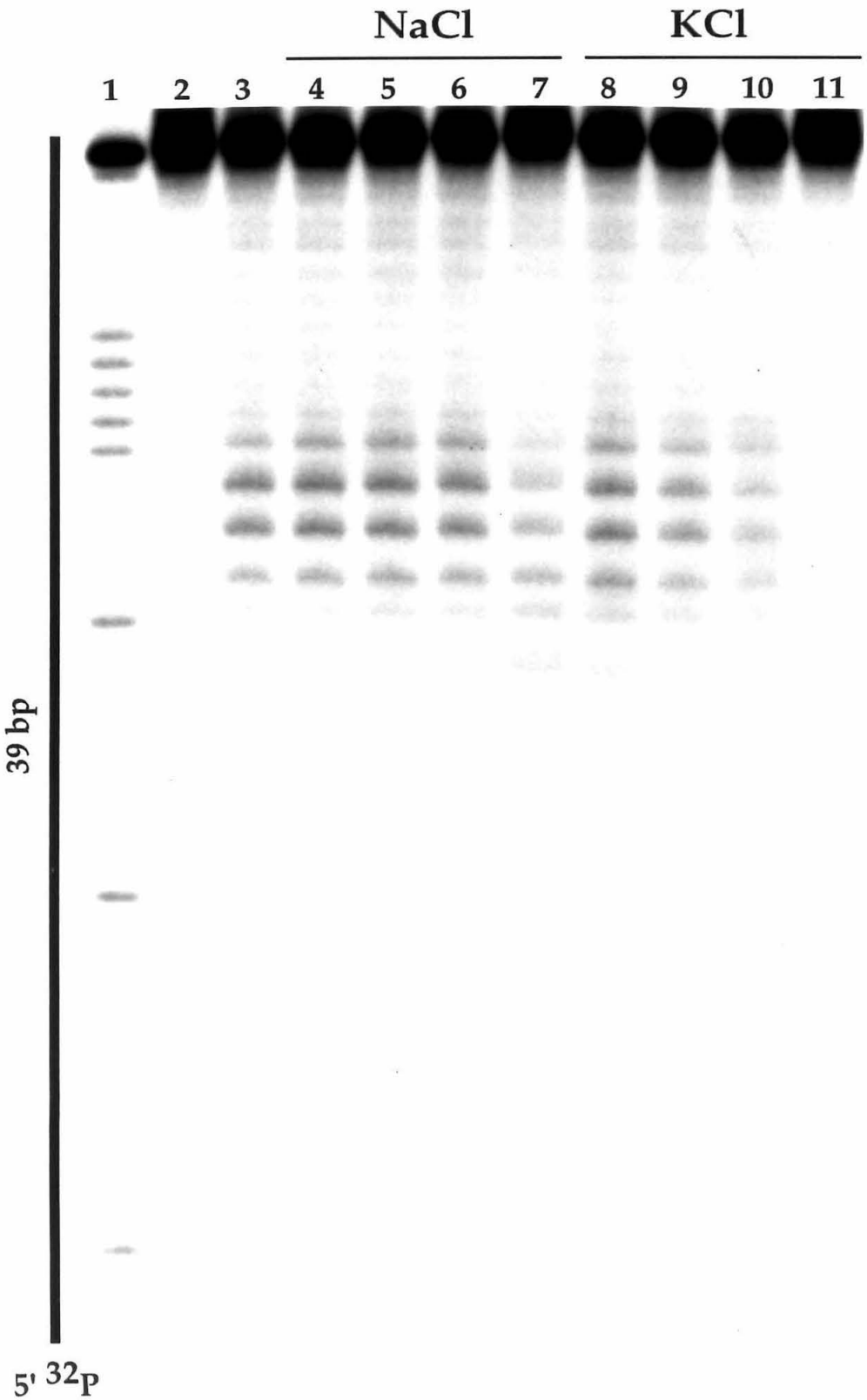
Figure 2.24 G quartet stabilized by a potassium ion.

formation.

As expected from these considerations, cleavage by **1** at its binding site in the 39 bp oligonucleotide duplex (see Fig. 2.8 for sequences) is inhibited by increasing NaCl and KCl concentration in the presence of 1 mM spermine. The inhibitory effect is realized at a lower concentration of KCl than NaCl, suggesting that an oligonucleotide secondary structure which is preferentially stabilized by K^+ over Na^+ could be important (Figures 2.25 and 2.26).

The effect of abundant cellular ions such as K^+, Na^+, Mg^{+2}, and spermine on the stability of the purine motif triple helix must be considered before the full potential of this structure for *in vivo* targeting of DNA sequences can be evaluated. These experiments provide the qualitative assessment that Mg^{+2} and spermine are stabilizing ions, whereas K^+ and Na^+ can be destabilizing. However, a thorough analysis of the effect of these ions on the kinetics and thermodynamics of purine motif triple helix formation has not yet been reported.

Figure 2.25 Autoradiogram of a 20% denaturing polyacrylamide gel used to separate cleavage products. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM) and Fe(II) (5 μM) with the ^{32}P-labeled 39 bp oligonucleotide duplex [~ 40,000 cpm] in a solution of TrisOAc pH=7.8 (50 mM), spermine (1 mM), and calf thymus DNA(100 μM in base pairs), NaCl or KCl at the concentration indicated below and then incubating at 24° C for 1 hour. Cleavage reactions were initiated by the addition of dithiothreitol (DTT) (4 mM) and allowed to proceed for 12 hours at 24° C. The reactions were stopped by freezing and lyophilization, and the cleavage products were analyzed by gel electrophoresis. (Lanes 1 to 11) Duplexes containing 5' end labelled d($A_2T_2(CT)_3A_5G_3AG_4AG_3A_5(CT)_3$) (Lane 1) Products of an A specific cleavage reactions[15] (Lane 2) Control showing intact 5' end labeled duplex. (Lanes 3-11) Cleavage products produced by oligonucleotide-EDTA•Fe(II) **1** (Lane 3) Cleavage products generated in the absence of a monovalent chloride (Lane 4) NaCl at 0.1 mM (Lane 5) NaCl at 1.0 mM (Lane 6) NaCl at 10 mM (Lane 7) NaCl at 100 mM (Lane 8) KCl at 0.1 mM (Lane 9) KCl at 1.0 mM (Lane 10) KCl at 10 mM (Lane 11) KCl at 100 mM.


NaCl
KCl
1
2
3
4
5
6
7
8
9
10
11
39 bp
5' 32P

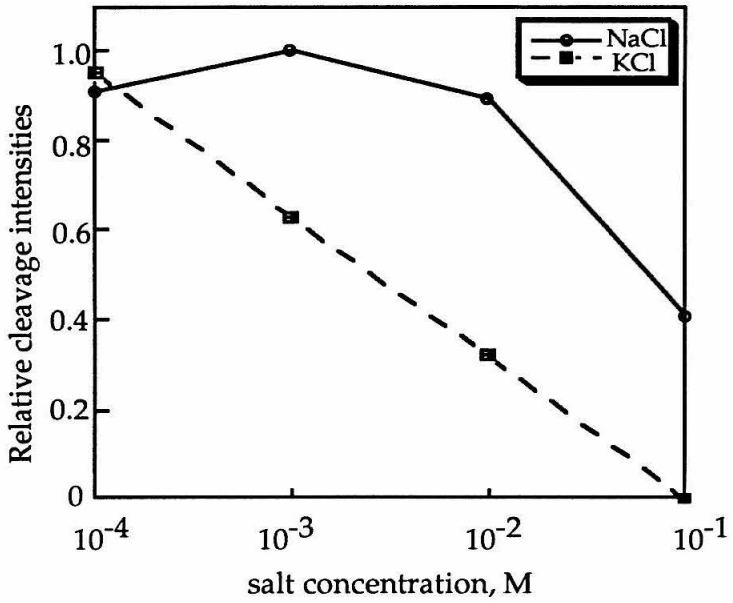


Figure 2.26 Comparison of the relative cleavage intensities for **1** in the presence of increasing concentrations of NaCl and KCl as determined by densitometry of autoradiogram in Figure 2.23.

The effect of the 16 common base triplets at a single variable position. It was shown that three A•AT or three T•AT base triplets stabilized a 19 mer purine motif triple helix to a greater extent than three C•AT triplets.[20] However, the effect of each of the 16 common natural base triplets on the stability of the purine motif triple helix was not known. For this reason, four 15 mer oligonucleotides were prepared. These oligonucleotides differed in sequence at a single position (Z), where Z = A, G, C, or T, for **6-9** respectively, and were equipped with thymidine-EDTA (T*) at each 3' end such that binding could be monitored by affinity cleaving. Because the EDTA•Fe moiety generates a diffusible oxidant yielding a Gaussian, sequence-independent cleavage pattern, the amount of cleavage observed is proportional to the fractional occupation of the target site by the oligonucleotide-EDTA•Fe.[30] Therefore, the relative

stabilities of the triple helices formed by oligonucleotides **6-9** with duplexes differing in sequence at a single variable position (XY = AT, GC, CG, and TA) opposite the variant third strand base were assessed by comparing the relative cleavage efficiencies (Figure 2.27).

The oligonucleotides **6-9** cleaved the target duplexes at the 5' end of the purine tract indicating that they bound in an orientation antiparallel to the WC purine strand (Figure 2.28). As discussed previously, cleavage efficiencies were dependent on ionic conditions, particularly the concentration of spermine. Under conditions of high spermine concentration (0.5-1.0 mM), **6-9** cleaved the four duplexes equally. However, more stringent conditions could be found under which differences in cleavage efficiencies were detected based on a single triplet change (100 nM oligonucleotide-EDTA, 100 μM spermine, 20 mM NaCl, pH 7.4). Under these conditions, the three combinations ZXY=AAT, TAT and GGC showed the most efficient cleavage (Figure 2.28, Lanes 3, 6 and 8 respectively). No other strong interactions could be detected. Some weak interactions were observed for ZXY= CAT, AGC and TCG (Figure 2.28, Lanes 5, 7, and 14 respectively). The relative cleavage efficiencies allowed for a qualitative ranking of the relative stabilities of the 16 triplet combinations at a single variable position within this sequence context (Figure 2.29 A and B).[27]

Models for the 16 common base triplets. The NMR spectroscopic studies of Radhakrishnan et al. indicated that the glycosidic bond conformations in a purine motif triple helix were *anti*.[17c,f,g] Therefore, we have considered possible two dimensional models for the 16 common triplets with the third strand base in an *anti* conformation. These models are shown in Figure 2.30. For each of the strong triplets A•AT, G•GC and T•AT, the third strand base is presumed to form two hydrogen bonds to the purine of the WC base pair. The three weaker base triplet interactions C•AT, A•GC and T•CG might be rationalized in part by

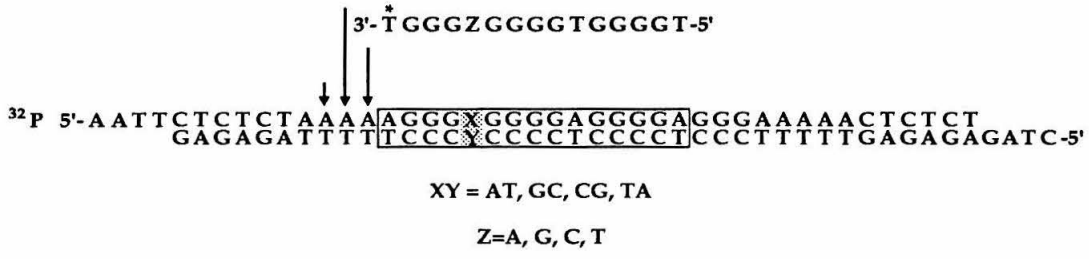


Figure 2.27 Sequences of oligonucleotidesEDTA **6-9** are shown, where T* indicates the position of thymidineEDTA. The oligonucleotides differ at one position, where Z= A, G, C and T for **6**, **7**, **8**, and **9** respectively. The target duplexes are indicated with a common variable base pair position. The box indicates the double stranded sequence bound by **6-9**. The WC base pair (AT, GC, CG or TA) opposite the variant base is shaded. The arrows show sites of cleavage and heights indicate the relative cleavage efficiency at the indicated bases.

Figure 2.28 Autoradiogram of a 20% denaturing polyacrylamide gel used to separate affinity cleavage products. (Lanes 1-18) Duplexes containing 5' end labeled 5'-AATTCTCTCTAAAAAGGGXGGGGAGGGGAGGGAAAAACTC TCT-3'. (Lane 1) Products of an adenine specific sequencing reaction for duplex with X=A.[15] (Lane 2) Intact 5' labeled duplex obtained after incubation under the conditions of the cleavage reactions in the absence of oligonucleotide-EDTA•Fe(II). (Lanes 3-18) DNA cleavage products produced by oligonucleotide-EDTA•Fe(II) **1-4**; **1** (Z=A) (lanes 3, 7, 11, and 15); **2** (Z=G) (lanes 4, 8, 12, and 16); **3** (Z=C) (lanes 5, 9, 13, and 17); **4** (Z=T) (lanes 6, 10, 14, and 18). XY=AT (Lanes 3-6). XY=GC (Lanes 7-10). XY=CG (Lanes 11-14). XY=TA (Lanes 15-18).

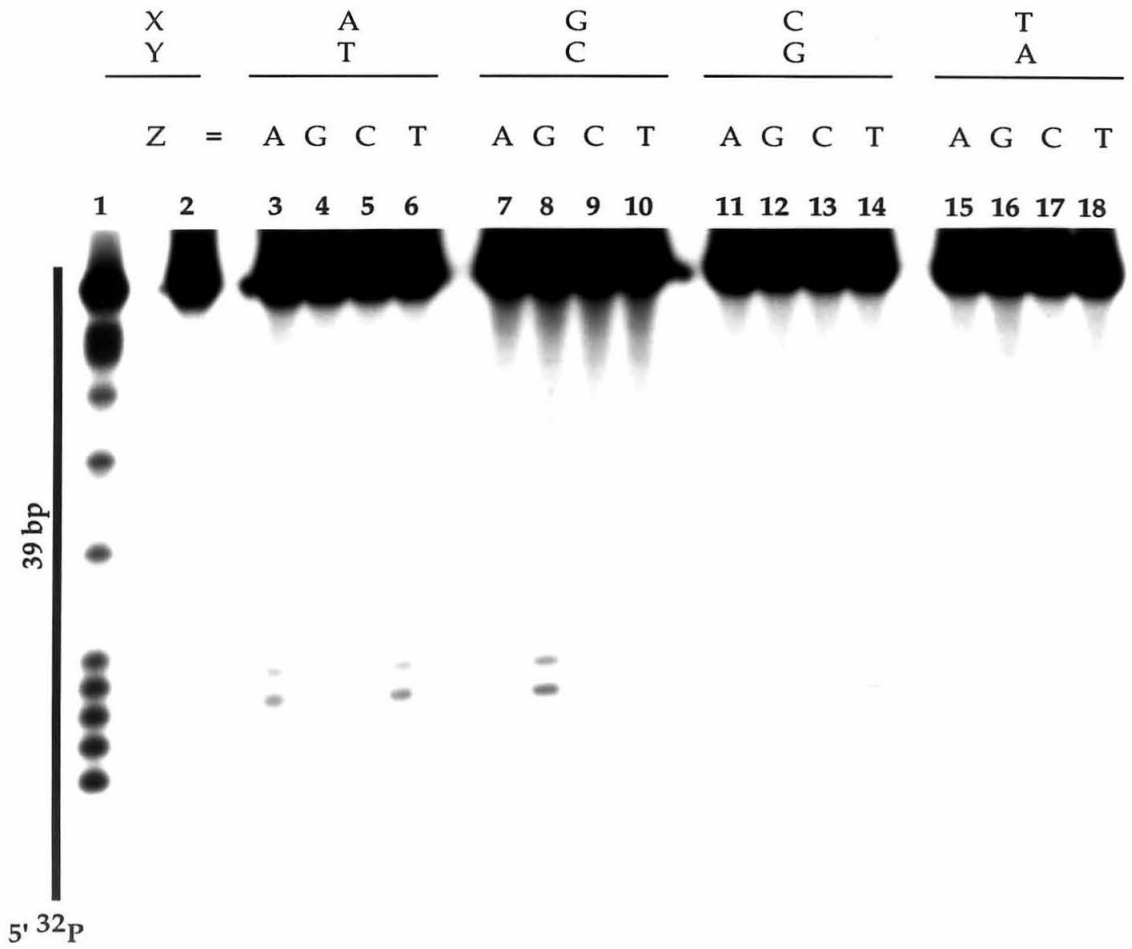
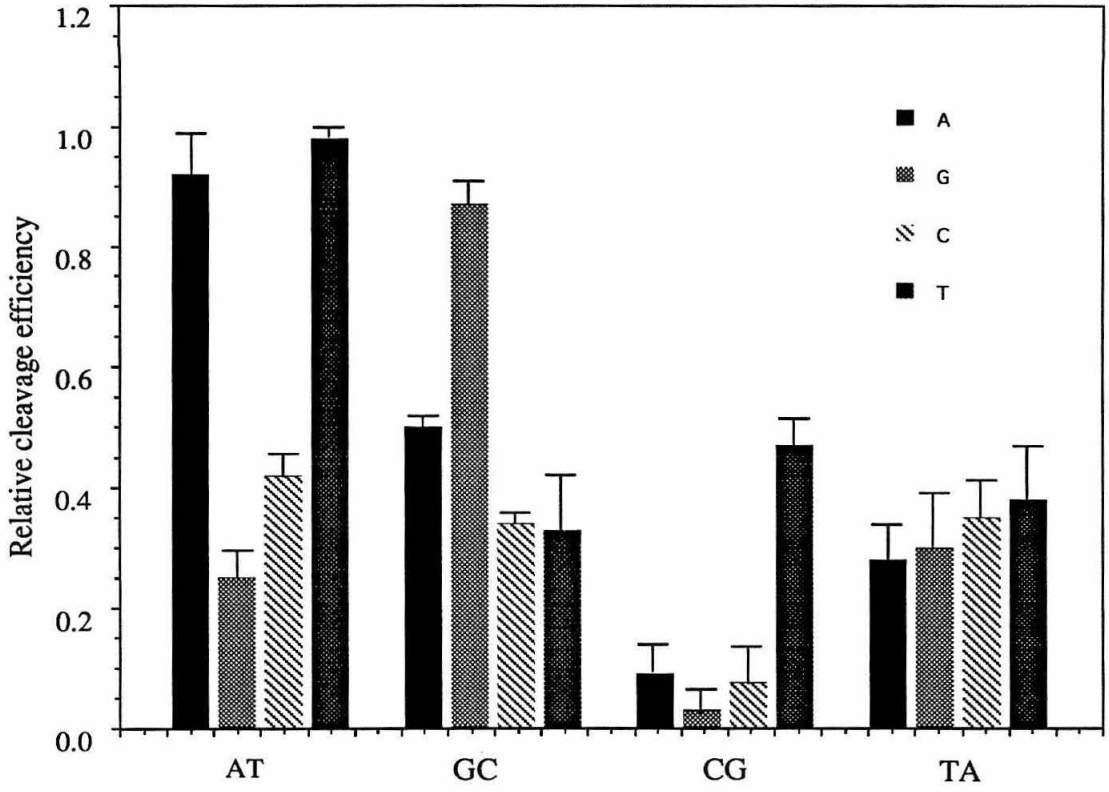

X
Y
A
T
G
C
C
G
T
A
Z = A G C T A G C T A G C T A G C T
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
39 bp
5' 32P

Figure 2.29 (A) Bar graph of relative cleavage efficiencies determined by quantitative analysis using phosphor storage autoradiography of oligonucleotidesEDTA **6-9** on duplexes containing each of the four base pairs at the single variable position. Values represent the mean ±SD for two determinations. (B) Relative strengths of base triplets in the purine motif determined from relative cleavage efficiencies. (+++) >80%; (++) 80-60; (+) 60-40; (-) <40 % relative to the most efficient combination.

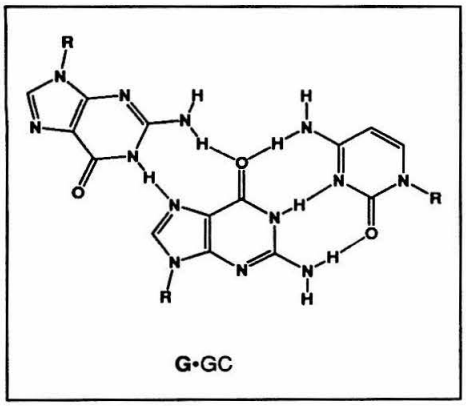
A



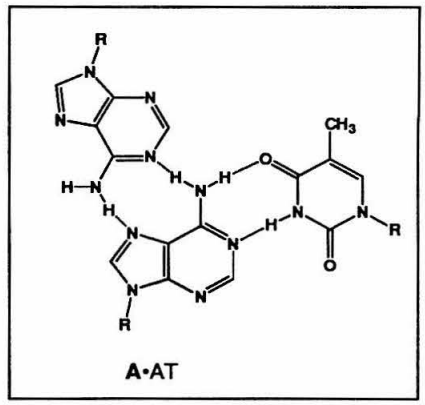
B

W C duplex (XY)	Third strand base (Z)			
	A	G	C	T
AT	+++	-	+	+++
GC	+	+++	-	-
CG	-	-	-	+
TA	-	-	-	-

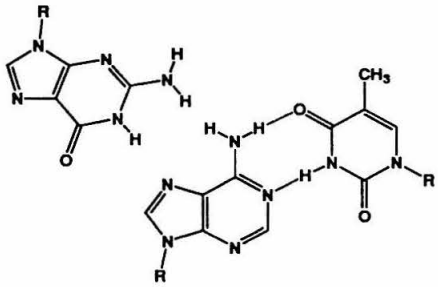
Figure 2.30 Models depicting the 16 possible natural base triplets for the purine motif triple helix. The third strand base has been positioned to maximize complementarity of hydrogen bonding surfaces while maintaining a backbone position similar to the G•GC, A•AT or T•AT triplet shown boxed. Hydrogen bonds are shown in triplets for which stabilizing interactions were detected; two for strong triplets, one for weak triplets. All bases are depicted with *anti* conformations about the glycosidic bonds.[17g]



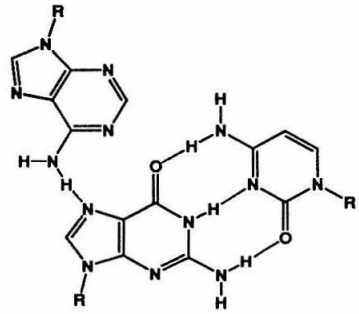
G•GC



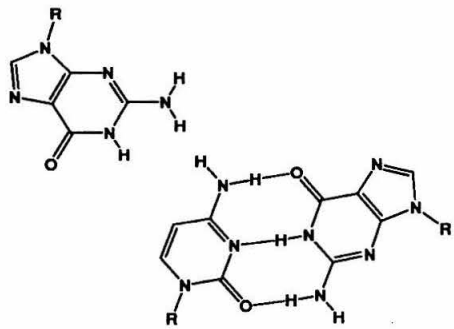
A•AT



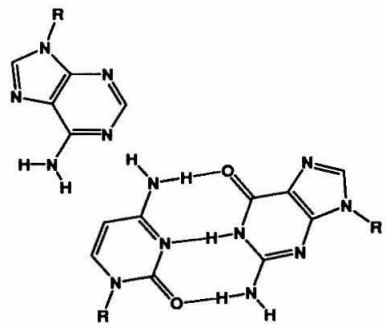
G•AT



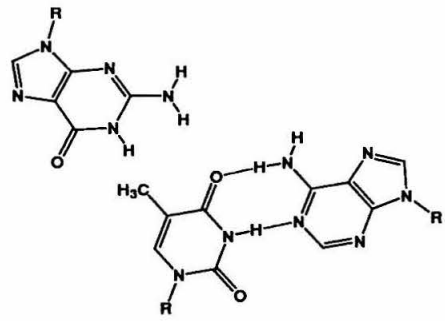
A•GC



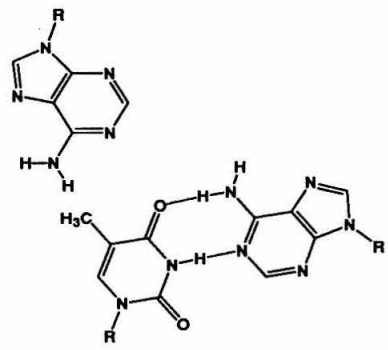
G•CG



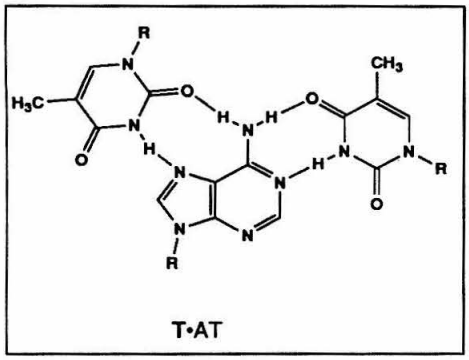
A•CG



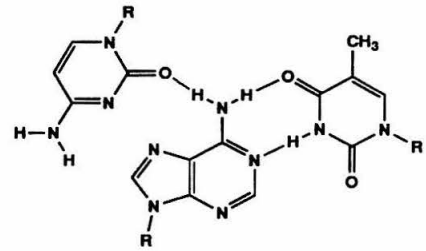
G•TA



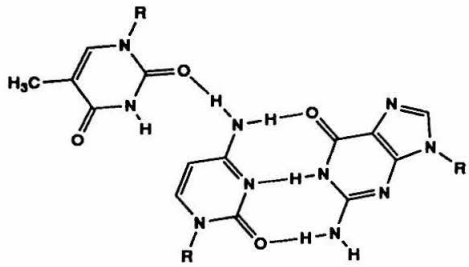
A•TA



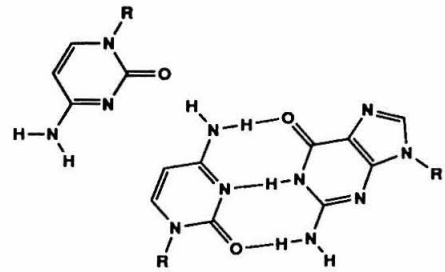
T•AT



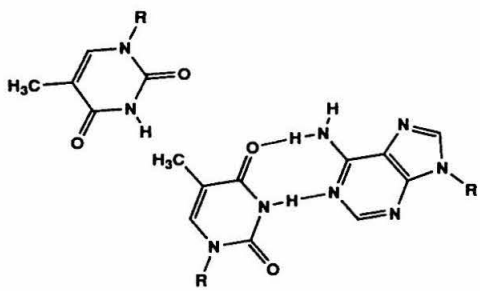
C•AT



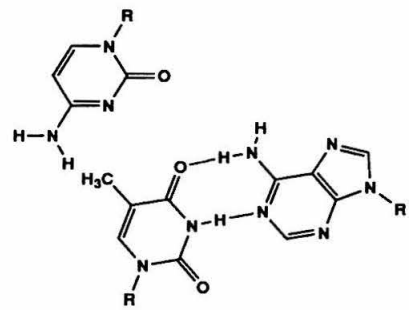
T•CG



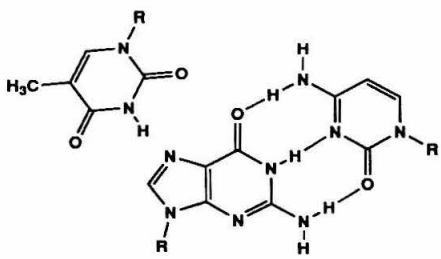
C•CG



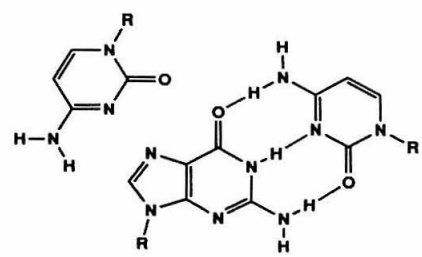
T•TA



C•TA



T•GC



C•GC

the formation of a single hydrogen bond in each case between the third strand base and WC base pair, while maintaining a backbone position similar to that of the G•GC, A•AT or T•AT triplets.

Because the most stable base triplets form between third strand bases and purines of the Watson Crick GC or AT base pairs, the best potential double helical target sites are purine tracts. A limited number of CG base pairs may possibly be tolerated, if thymidine in the third strand is used (T•CG triplets). However, since thymidine binds tightly to the AT base pair (T•AT triplets), it is not a solution to specific recognition of CG. Also, no stabilizing interactions were detected between the third strand bases and the TA base pair. This lack of specific, highly stabilizing interactions at the CG or TA base pair demonstrated a sequence limitation to the formation of this structure. It should be emphasized that these triplet stabilities were determined only for the sequences shown in Figure 2.27, where the variable triplet is flanked by two G•GC triplets. The relative strengths of base triplets may vary with flanking sequence, as is observed in the pyrimidine motif triple helix.[28]

Since the common DNA bases were shown not to form specific, stabilizing interactions with TA or CG base pairs, research directed toward the design of heterocycles which could specifically recognize TA or CG was initiated. Stilz and Dervan reported that 2'-deoxynebularine bound to CG and AT base pairs with nearly equal affinity and work is currently under way in the Dervan laboratories to develop nebularine analogs with increased affinity and specificity for CG base pairs (Figure 2.31).[29] However, there have been no reported structural leads for the specific recognition of TA base pairs in the purine triple helix motif.

Thermodynamic analysis of the effect of a mismatched triplet. For a more quantitative assessment of the effect of mismatched triplets, equilibrium

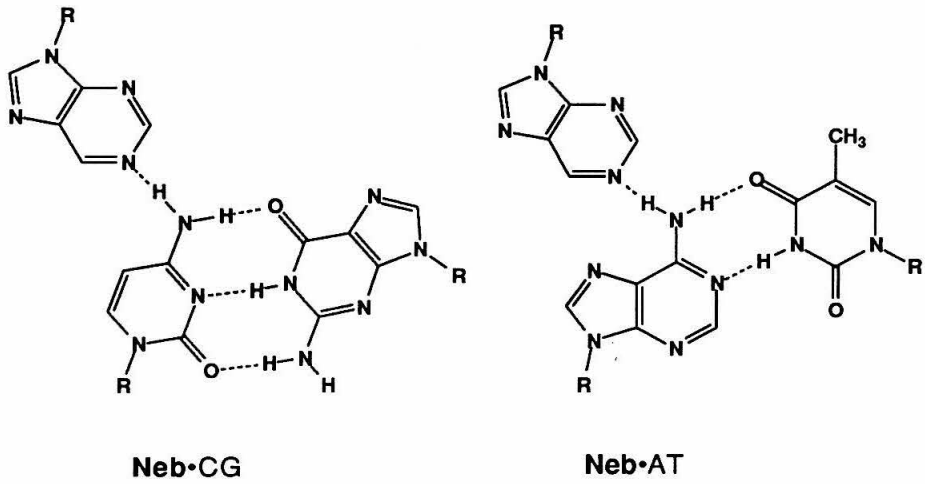


Figure 2.31 Models representing the Neb•CG and Neb•AT base triplets formed between 2′-deoxynebularine and the WC duplex base pairs CG and AT.[29]

binding constants were determined for a triple helix with 15 matched triplets and one with a single mismatched triplet. The quantitative affinity cleavage titration was used for this purpose. A detailed description of this procedure has been published.[26] Briefly, the formation of cleavage products (P_{cl}) from the binding of an oligonucleotideEDTA•Fe (O) to a double helical target site (D) to form a triple helix (T) followed by the affinity cleavage reaction can be expressed by:

$$O + D \overset{K_T}{\rightleftharpoons} T \xrightarrow{k_{cl}} P_{cl} \qquad (1)$$

where K_T is the equilibrium binding constant for the oligonucleotide and k_{cl} is the pseudo-first order rate constant for the cleavage reaction. K_T can be

expressed as a function of Θ, the fractional occupancy of the duplex target site, and $[O]_{eq}$, the equilibrium concentration of the oligonucleotide by:

$$K_T = \frac{\Theta}{1 - \Theta} \frac{1}{[O]_{eq}} \qquad (2)$$

The apparent fractional occupancy of the duplex site (Θ_{app}) has been shown to be related to the intensity of cleavage bands (I_{site}) at the site by the expression:

$$\Theta_{app} = I_{site} I_{sat}^{-1} \qquad (3)$$

where I_{sat} is the intensity of the cleavage bands when $\Theta=1$.[26] Also, with ^{32}P labeled duplex DNA, very low duplex concentrations are used and the approximation that $[O]_{eq}=[O]_{tot}$ is reasonable, since even at $\Theta=1$, the concentration of bound O is negligible compared to $[O]_{tot}$. Therefore, substituting $I_{site}I_{sat}^{-1}$ and $[O]_{tot}$ into equation 2 allows for the expression of K_T in terms of I_{site}, I_{sat}, and $[O]_{tot}$:

$$I_{site} = I_{sat} \frac{K_T [O]_{tot}}{1 + K_T [O]_{tot}} \qquad (4)$$

Thus, by measuring the intensity of cleavage bands at the target site as a function of $[O]_{tot}$, an empirical binding isotherm can be generated and from it K_T can be determined by nonlinear least squares analysis.

This method was used to determine the equilibrium binding constant for oligonucleotide **7** (where Z=G) at its 15 bp duplex target sequence located on the 648 bp *HindIII/SspI* restriction fragment from the plasmid pPB19GC (where XY=GC) (Figure 2.32).

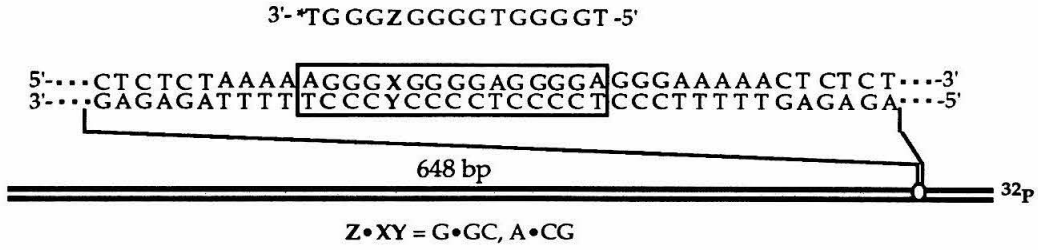


Figure 2.32 Sequences of oligonucleotides **6** and **7**, where Z=A and G, respectively, and the triple helix site present on the 648 bp *HindIII/SspI* fragments from pPB19GC and pPB19CG, where XY = GC and CG, respectively.

Shown in Figure 2.33 is an autoradiogram of the polyacrylamide gel used to separate the cleavage products generated by **7** at this target sequence as a function of oligonucleotide concentration. At 100 μM spermine, 10 mM NaCl, pH=7.5, the equilibrium binding constant where ZXY=GGC was determined to be 8.5(±0.9) x10^6 M^{-1}, corresponding to a binding free energy (ΔG_T) of 9.4(±0.1) kcal/mole (Figure 2.34, Table I). Similarly, the binding constant for oligonucleotide **6** (where Z=A) at the duplex target sequence located on the *HindIII/SspI* fragment of pPB19CG (where XY=CG) was determined. Under the same conditions as above, the equilibrium binding constant where ZXY=ACG is 1.6(±0.1)x10^6 M^{-1}, equal to a binding free energy (ΔG_T) of 8.4(±0.1) kcal/mole (Figure 2.34, Table I). This corresponds to a decrease in the binding constant by a factor of 5 or a destabilization of this triple helix by 1.0 kcal/mole from the presence of the mismatched A•CG triplet relative to the matched G•GC triplet.

Singleton and Dervan have shown the pyrimidine motif triple helix is highly sensitive to mismatches, with a difference in binding free energy of 2.5-3.0 kcal/mole when comparing matched 15 mers with 15 mers containing a single mismatched triplet.[26] The destabilization of purine motif triple helices by base

Figure 2.33 Autoradiogram of an 8% denaturing polyacrylamide gel used to separate the cleavage products generated by 7•Fe(II) from a quantitative affinity cleavage experiment performed in association buffer (10 mM NaCl, 0.1 mM spermine tetrahydrochloride, 16 μM $Fe(NH_4)_2(SO_4)_2 \bullet 6H_2O$, 0.1 mM bp calf thymus DNA, 50 mM TrisOAc, pH=7.5) at 24°C on the 3′ end labeled HindIII/SspI fragment from pPBGC19. Bars on the left of the autoradiogram indicate the 15 bp triple helix site (white bar), the bands used to measure I_{site} (dark grey bar) and the bands used to measure I_{ref} (light grey bar). (Lane 1) Products of an adenine specific cleavage reaction.[15] (Lane 2) Intact 3′ end labeled fragment incubated under the reaction conditions in the absence of oligonucleotideEDTA•Fe (Lane 3-19) DNA affinity cleavage products produced by 7•Fe(II) at the concentrations indicated below: (Lane 3) 2 μM (Lane 4) 1 μM (Lane 5) 800 nM (Lane 6) 400 nM (Lane 7) 200 nM (Lane 8) 100 nM (Lane 9) 80 nM (Lane 10) 40 nM (Lane 11) 20 nM (Lane 12) 10 nM (Lane 13) 8 nM (Lane 14) 4 nM (Lane 15) 2 nM (Lane 16) 1 nM (Lane 17) 0.8 nM (Lane 18) 0.4 nM (Lane 19) 0.2 nM.

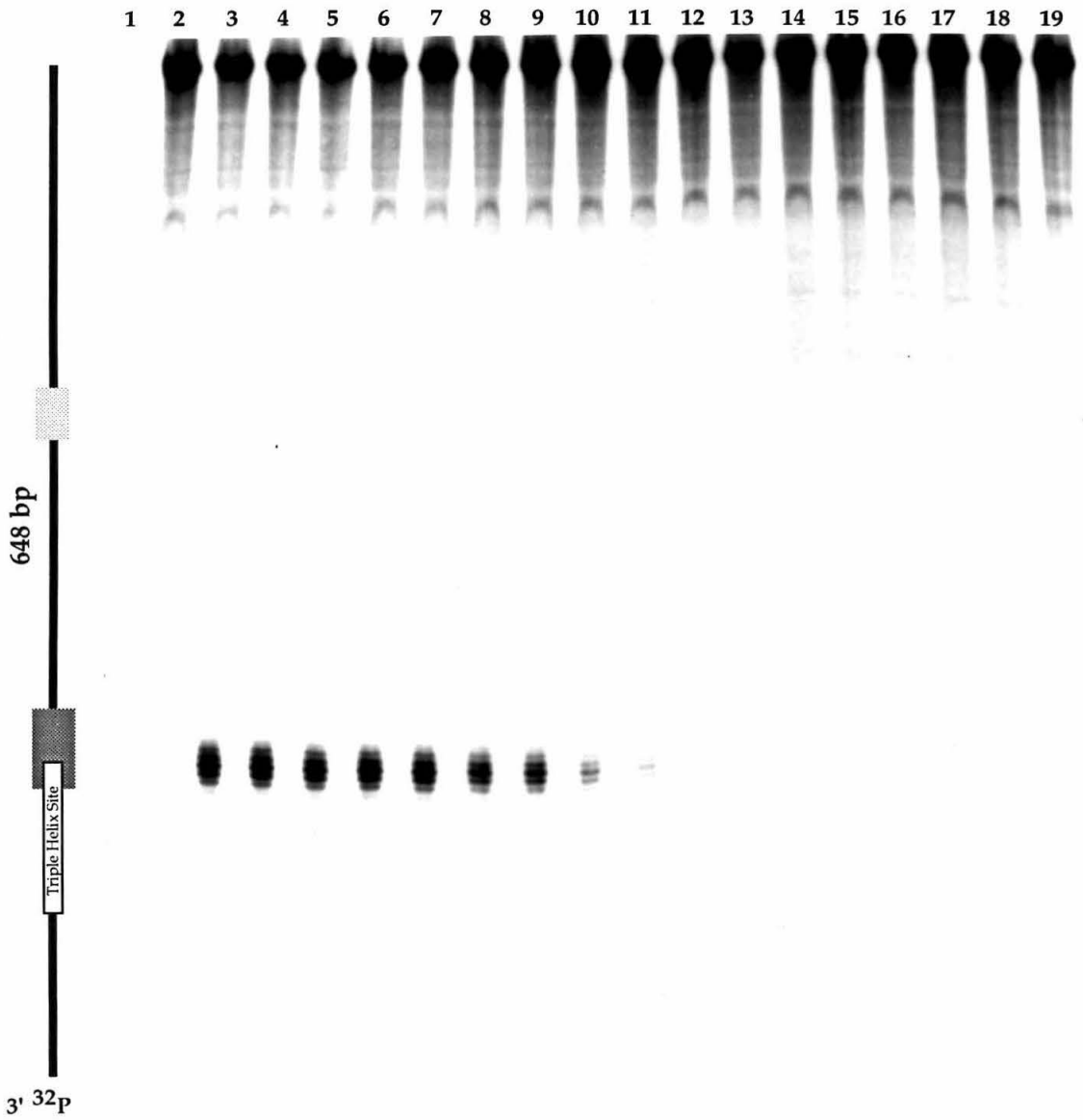

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
648 bp
Triple Helix Site
3' 32P

Figure 2.34 Data from quantitative affinity cleavage experiments involving oligonucleotidesEDTA **6** and **7** preequilibrated with labeled duplex DNA in association buffer at 24°C for 48 H followed by 4 H affinity cleavage reactions. The data points represent the average site specific cleavage signal intensities from four independent experiments. The sigmoidal curves show the titration binding isotherms plotted using the mean values of K_T for **6** and **7** (Table I) and equation 4. The data points were normalized using I_{sat} for each experiment and the binding curves were subsequently normalized using $I_{sat} = 1$ for equation 4.

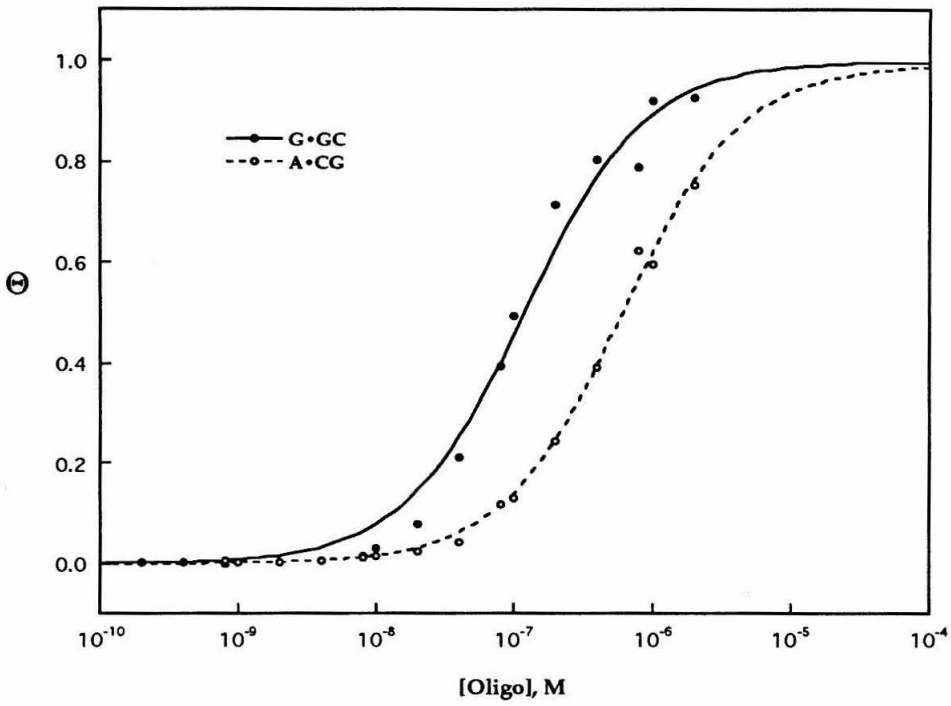


Table I: Association constants at pH = 7.5 and 25°C[a]

Z•XY	K_T (M^{-1})	ΔG_T (kcal•mol^{-1})
G•GC	8.5 (±0.9) x 10^6	-9.4 (±0.1)
A•CG	1.6 (±0.1) x 10^6	-8.4 (±0.1)

[a]Values listed are mean values measured from four affinity cleavage titrations performed in association buffer (10 mM NaCl, 0.1 mM spermine, 16 μM $Fe(NH_4)_2(SO_4)_2•6H_2O$, 0.1 mM bp calf thymus DNA, 50 mM TrisOAc, pH=7.5).

triplet mismatches will be dependent on the identity of the mismatch as seen from the qualitative ranking of the stability of the 16 common triplets in Figure 2.29. For one of the most destabilizing mismatches (A•CG), the energetic penalty was determined here to be 1.0 kcal/mole, suggesting that the purine motif triple helix may not be as sensitive to mismatches as is the pyrimidine motif triple helix. It is possible that mismatched triplets in the purine motif triple helix maintain a stabilizing stacking interaction, even though hydrogen bonding interactions are not optimal. This might result in greater stability for mismatched sequences.

The effect of G•GC content on triple helix stability. Recent reports on the formation of purine motif triple helices involved targeting duplex sequences rich in GC base pairs, including the sequences discussed so far in this work.[4,5,6b,17] Several groups have reported difficulty in forming triple helices comprised of mainly A•AT triplets.[3,6b,17b] In addition, T-rich oligonucleotides bind parallel to the purine strand at AT rich targets (pyrimidine motif triple helices).[16]

To determine if a minimum number of G•GC triplets were necessary for formation of a purine motif triple helix, we designed 15 bp target sequences varying the GC to AT ratio. Figure 2.35 shows the sequences of these double helical target sites present on restriction fragments generated from cloned plasmids. Duplex A has eight GC and seven AT base pairs, duplex B has six GC and nine AT base pairs and sequence C has four GC and 11 AT base pairs. OligonucleotidesEDTA **10-12** were designed to bind to these duplexes by the formation of G•GC and T•AT triplets and their binding properties were determined by affinity cleaving.

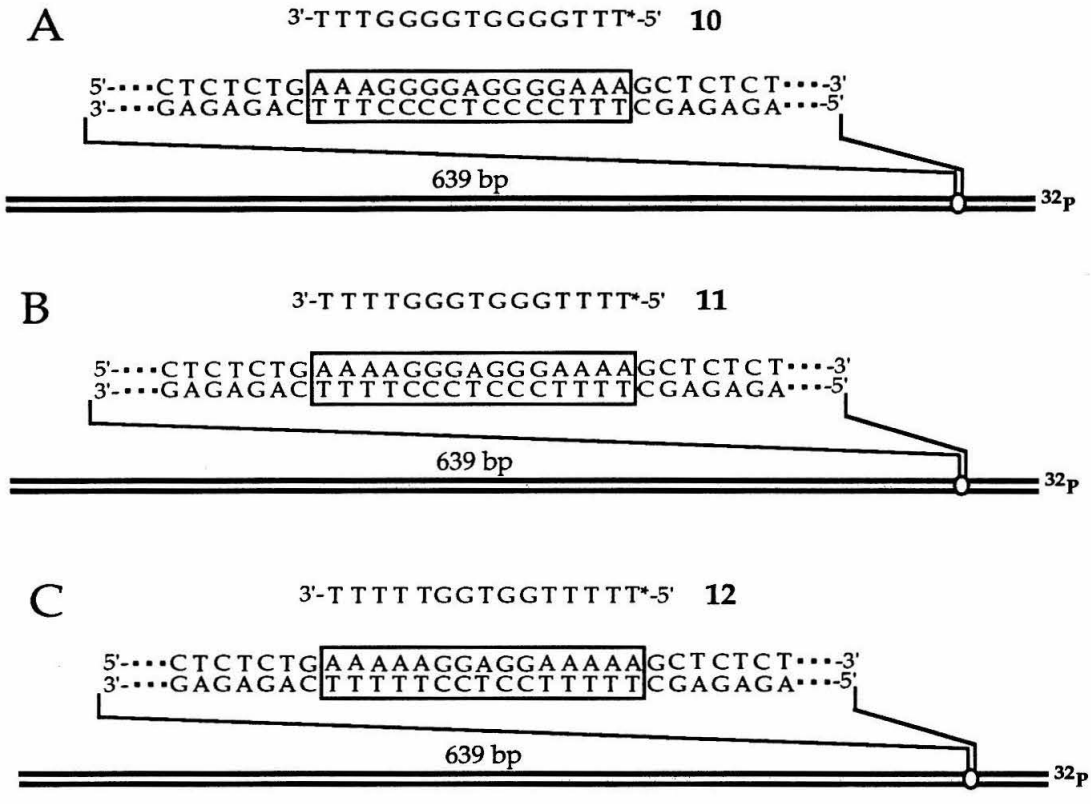
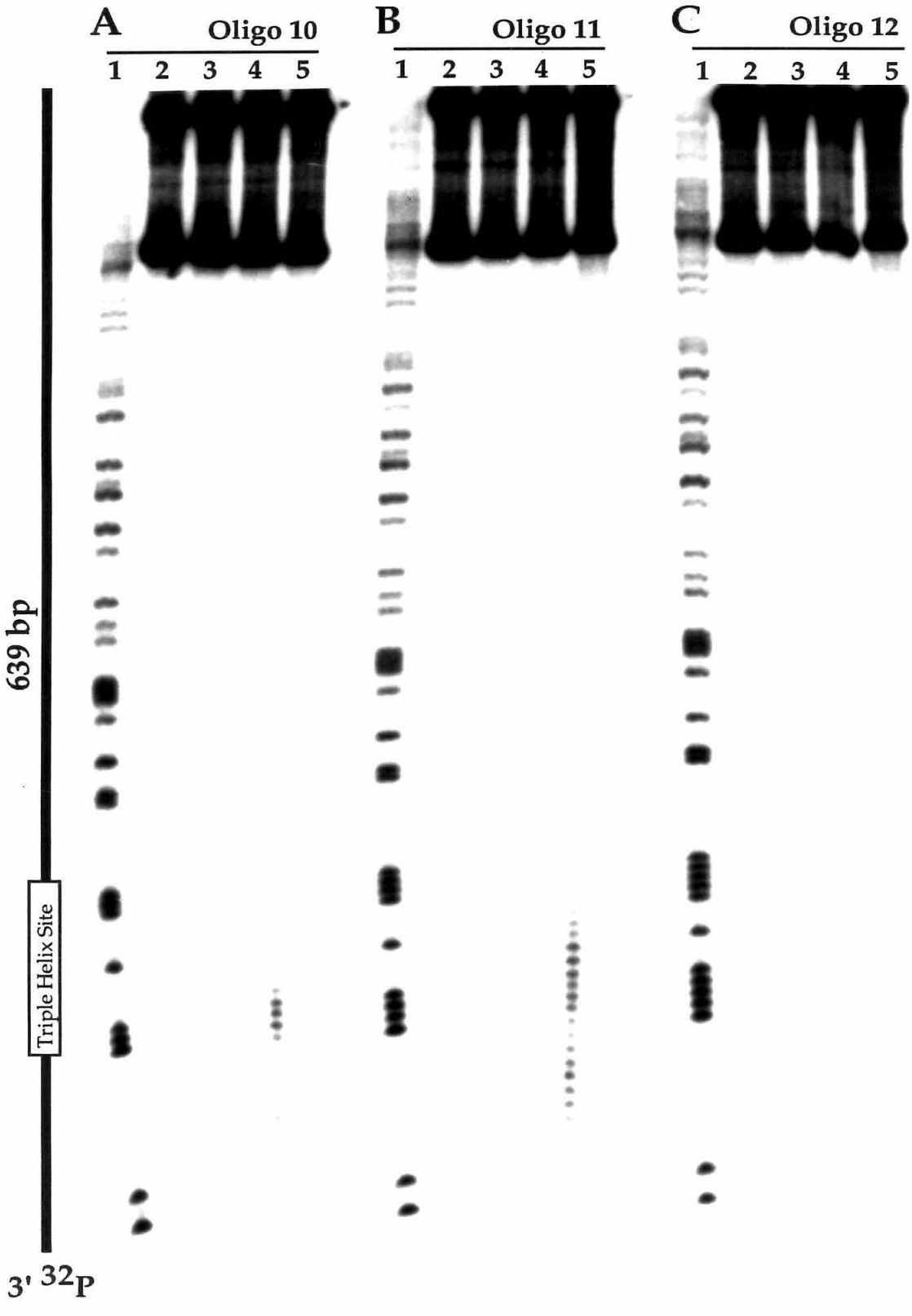


Figure 2.35 Sequences of oligonucleotidesEDTA **10, 11** and **12** and their double helical target sites **A, B** and **C** present on 639 bp restriction fragments from cloned plasmids. The resulting triple helices differ in G•GC content.

Figure 2.36 Autoradiogram of an 8% denaturing polyacrylamide gel used to separate the cleavage products generated by oligonucleotidesEDTA•Fe **10, 11** and **12** on 3′ end labeled restriction fragments containing duplex sequences **A, B** or **C** respectively (Figure 2.35). The reactions were carried out by combining a mixture of oligonucleotide-EDTA and 2.5 equivalents of $Fe(NH_4)_2(SO_4)_2 \bullet 6H_2O$ with labeled restriction fragment [~ 100 nM in bp, ~ 30,000 cpm] in a solution of TrisOAc pH=7.4 (50 mM), NaCl (10 mM), spermine (1 mM) and calf thymus DNA (100 μM in bp) and then incubating at 24° C for 1 H. Cleavage reactions were initiated by the addition of dithiothreitol (DTT) (4 mM) and allowed to proceed for 12 H at 24° C. **A**) Cleavage products generated by oligonucleotideEDTA•Fe **10** on duplex **A**. (Lane **A**1) Products of an adenine specific cleavage reaction.[15] (Lane **A**2) Intact 3′ end labeled fragment incubated under the reaction conditions in the absence of **10**. (Lane **A**3-**A**5) DNA affinity cleavage products produced by **10**•Fe(II) at the concentrations indicated below: (Lane **A**3) 0.02 μM (Lane **A**4) 0.2 μM (Lane **A**5) 2 μM **B**) Cleavage products generated by oligonucleotideEDTA•Fe **11** on duplex **B**. (Lane **B**1) Products of an adenine specific cleavage reaction. (Lane **B**2) Intact 3′ end labeled fragment incubated under the reaction conditions in the absence of **11**. (Lane **B**3-**B**5) DNA affinity cleavage products produced by **11**•Fe(II) at the concentrations indicated below: (Lane **B**3) 0.02 μM (Lane **B**4) 0.2 μM (Lane **B**5) 2 μM **C**) Cleavage products generated by oligonucleotideEDTA•Fe **12** on duplex **C**. (Lane **C**1) Products of an adenine specific cleavage reaction. (Lane **C**2) Intact 3′ end labeled fragment incubated under the reaction conditions in the absence of **12**. (Lane **C**3-**C**5) DNA affinity cleavage products produced by **12**•Fe(II) at the concentrations indicated below: (Lane **C**3) 0.02 μM (Lane **C**4) 0.2 μM (Lane **C**5) 2 μM.


A
Oligo 10
1 2 3 4 5
B
Oligo 11
1 2 3 4 5
C
Oligo 12
1 2 3 4 5
639 bp
Triple Helix Site
3' 32P

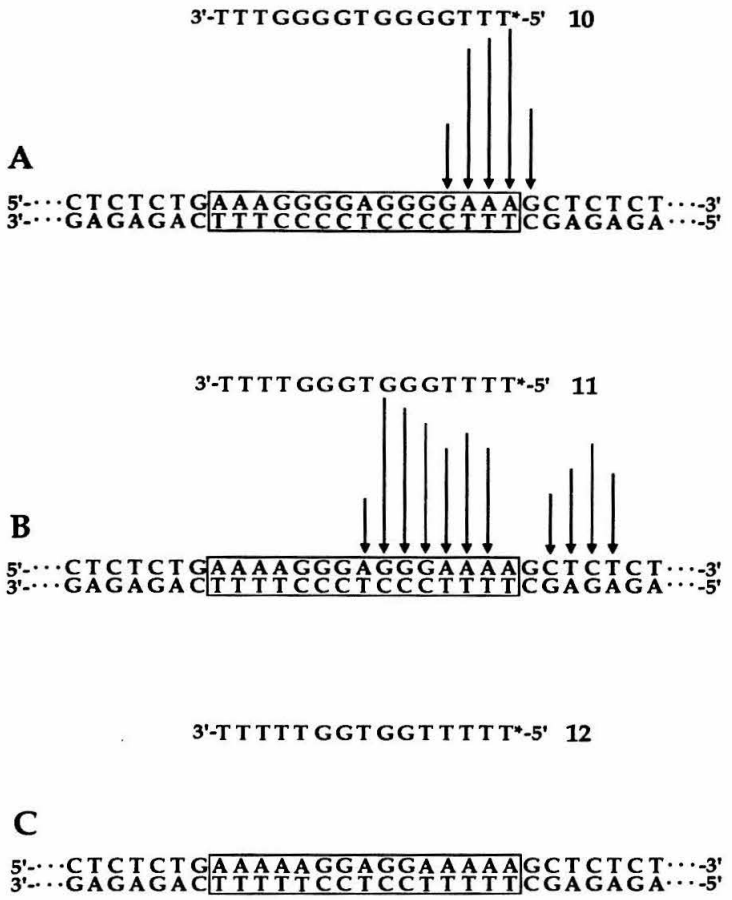


Figure 2.37 Histograms of DNA cleavage patterns generated by **10, 11** and **12** derived from densitometry of the autoradiogram in Figure 2.36.

At 1 mM spermine, 10 mM NaCl, pH=7, **10** bound to duplex A antiparallel to the purine strand of the duplex at oligonucleotide concentrations of ≥ 0.2 μM (Figures 2.36 and 2.37). Oligonucleotide **11** bound to duplex B in an antiparallel orientation at oligonucleotide concentrations of ≥ 2 μM (Figures 2.36 and 2.37). Interestingly, the cleavage pattern for **11** is more diffuse than that for **10**, with cleavage extending over 12 nucleotides. Under the same conditions, no binding of oligonucleotide **12** could be detected on duplex C.

These results indicated that the stability of purine motif triple helices could be correlated to the number of G•GC base triplets. Decreasing the number from eight to six destabilized the triple helix such that binding could only be detected at concentrations ≥ 2 μM. In addition, the diffuse cleavage pattern found for this triple helix suggested that the end of the oligonucleotide containing the T* may not have been as tightly bound to the duplex as seen for the combination of **10** and duplex **A**. A further decrease in the number of G•GC base triplets to four eliminated binding. These studies indicated that for a 15 mer triple helix, >27% of the base triplets (four of 15) had to be G•GC to support purine motif triple helix formation at micromolar concentrations of oligonucleotide. Although the T•AT and A•AT triplets stabilize a purine motif triple helix when flanked on either side by G•GC triplets,[27] nearest neighbor interactions could attenuate the stabilizing effects of these triplets, such that only triple helices with a minimum number of G•GC triplets will be stable.

Purine motif triple helix formation at a purine tract from the HIV proviral DNA. The sequence 5′-$A_4GA_4G_6A$-3′ is present in the long terminal repeat (LTR) of the HIV proviral DNA.[30] To determine if this sequence could support a purine motif triple helix, the plasmid pPBHIV was prepared with this sequence located on a *HindIII/SspI* restriction fragment. Oligonucleotides **13** and **14** were designed to bind antiparallel to the purine strand of this duplex by the

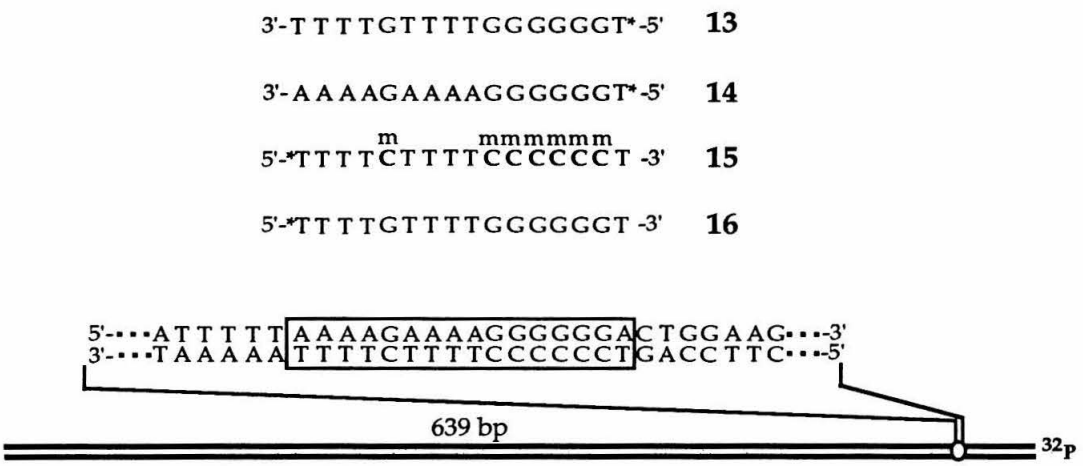


Figure 2.38 Sequences of oligonucleotidesEDTA **13-16** and the 16 bp purine tract found in the LTR of the HIV proviral DNA. This duplex sequence is present on the 639 bp *HindIII/SspI* fragment from pPBHIV.

formation of G•GC and T•AT(**13**) or G•GC and A•AT(**14**) triplets (Figure 2.38). In addition, oligonucleotide **15**, designed to bind parallel to the purine strand by the formation ^{Me}C+GC and T•AT triplets, was prepared and has been previously shown to bind this sequence at acidic pH.[31] Also, recently it was reported that oligonucleotides containing guanines and thymines could bind to purine tracts *parallel* to the purine strand of the duplex by the formation of Hoogsteen G•GC and T•AT triplets.[32] For this reason, oligonucleotide **16** was prepared and its binding was compared to that of **13-15**. ThymidineEDTA (T*) was incorporated at the 5′ end of each of these oligonucleotides and affinity cleaving was used to assess binding.

At 1 mM spermine, 10 mM NaCl, pH=7.0, **13-15** bind the target duplex in the orientation designed, as evidenced by the affinity cleavage pattern generated on the 3′ end ^{32}P labeled *HindIII/SspI* restriction fragment from pPBHIV. Only weak cleavage was detected for **16** (Figures 2.39 and 2.40). The quantitative affinity cleavage titration was used to measure the equilibrium binding constants for **13** and **14** binding to this target sequence. Oligonucleotide **13** bound with a binding constant of 2.9(±0.9)x10^6 M^{-1}, while **14** bound at 2.2(±0.2)x10^6 M^{-1} (Figure 2.41, Table II). Thus, the purine tract present in the HIV proviral DNA can support both pyrimidine and purine motif triple helices. When the purine motif triple helix is employed, oligonucleotides designed to bind by the formation of G•GC and T•AT *or* G•GC and A•AT triplets bind with nearly equal affinities. Since only weak cleavage could be detected for **16**, the antiparallel orientation is preferred for oligonucleotides containing guanosine and thymine for this sequence, as expected from our previous work.

The pH dependence on the recognition of GC-rich sequences through pyrimidine motif triple helix formation is well documented.[33] The requirement for protonation of cytosine in the third strand to form C+GC triplets limits the

Figure 2.39 Autoradiogram of an 8% denaturing polyacrylamide gel used to separate the cleavage products generated by oligonucleotidesEDTA•Fe **13-16** on 3′ end labeled *HindIII/SspI* restriction fragment from pPBHIV. The reactions were carried out by combining a mixture of oligonucleotide-EDTA and 2.5 equivalents of $Fe(NH_4)_2(SO_4)_2 \bullet 6H_2O$ with labeled restriction fragment [~ 100 nM in bp, ~ 20,000 cpm] in a solution of TrisOAc pH=7.4 (50 mM), NaCl (10 mM), spermine (1 mM) and calf thymus DNA (100 μM in bp) and then incubating at 24° C for 1 H. Cleavage reactions were initiated by the addition of dithiothreitol (DTT) (4 mM) and allowed to proceed for 12 H at 25°C. (Lane 1) Products of an adenine specific cleavage reaction.[15] (Lane 2) Intact 3′ end labeled fragment incubated under the reaction conditions in the absence of oligonucleotideEDTA. (Lanes 3-6) DNA affinity cleavage products generated at 1 μM oligonucleotideEDTA (Lane 3) **15** (Lane 4) **16** (Lane 5) **13** (Lane 6) **14**.

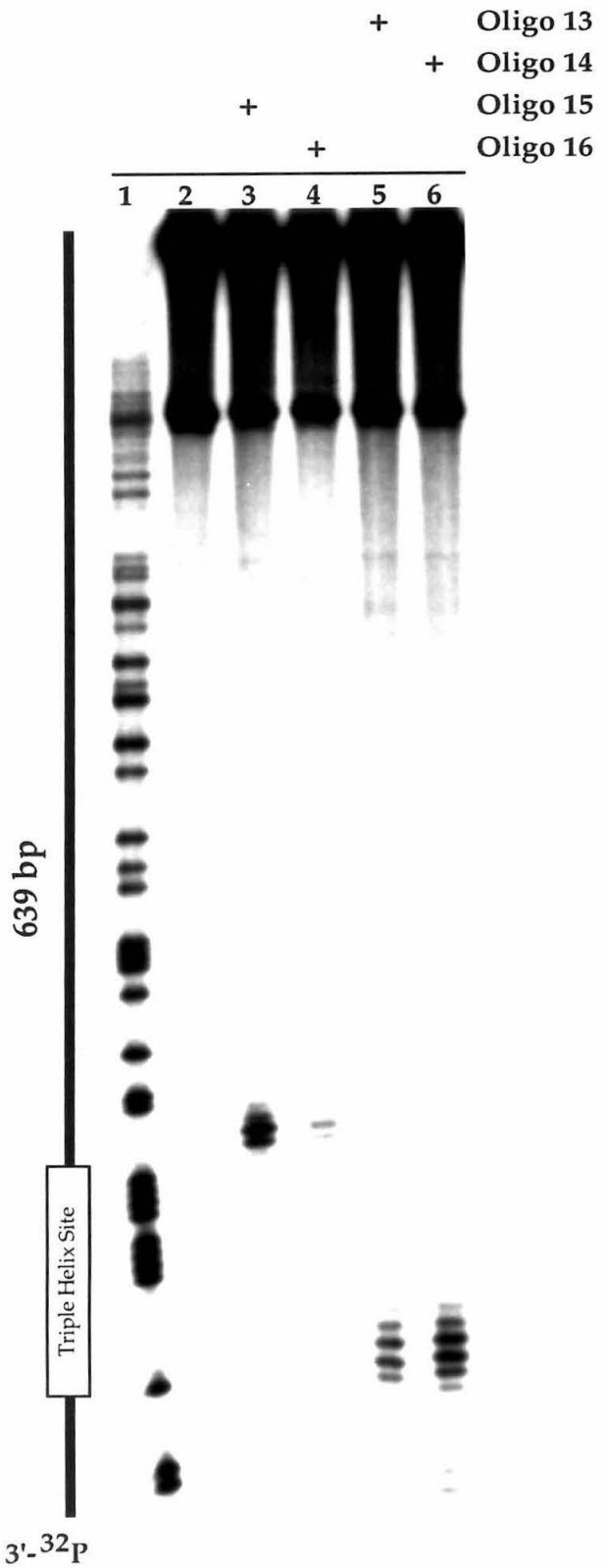

+ Oligo 13
+ Oligo 14
+ Oligo 15
+ Oligo 16
1 2 3 4 5 6
639 bp
Triple Helix Site
3'-32P

Figure 2.40 Histograms of DNA cleavage patterns generated by **13-16** on the *HindIII/SspI* fragment from pPBHIV as derived from densitometry of the autoradiogram in Fig. 2.39.

Figure 2.41 Data from quantitative affinity cleavage experiments involving oligonucleotidesEDTA **13** and **14** preequilibrated with labeled duplex DNA in association buffer at 24°C for 24 H followed by 6 H affinity cleavage reactions. The data points represent the average site specific cleavage signal intensities from three independent experiments. The sigmoidal curves show the titration binding isotherms plotted using the mean values of K_T for **13** and **14** (Table II) and equation 4. The data points were normalized using I_{sat} for each experiment and the binding curves were subsequently normalized using $I_{sat} = 1$ for equation 4.

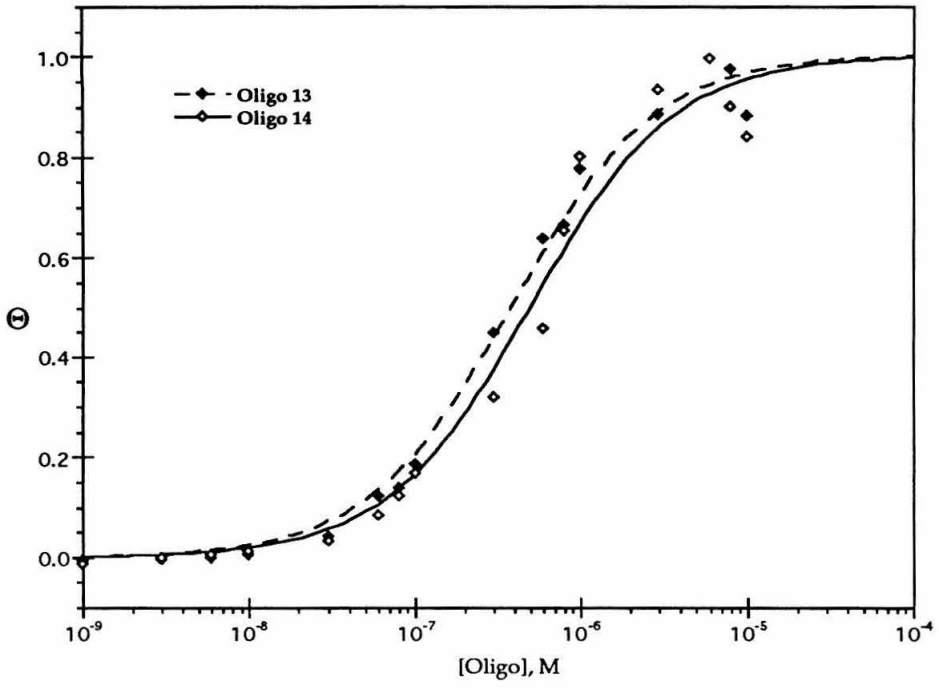


Table I: Association constants at pH = 7.0 and 25°C[a]

Oligo	K_T (M^{-1})	ΔG_T (kcal•mol^{-1})
13	2.5 (±0.4) x 10^6	-8.7 (±0.1)
14	2.0 (±0.7) x 10^6	-8.6 (±0.1)

[a]Values listed are mean values measured from three affinity cleavage titrations performed in association buffer (10 mM NaCl, 1.0 mM spermine, 0.1 mM bp calf thymus DNA, 50 mM TrisOAc, pH=7.0).

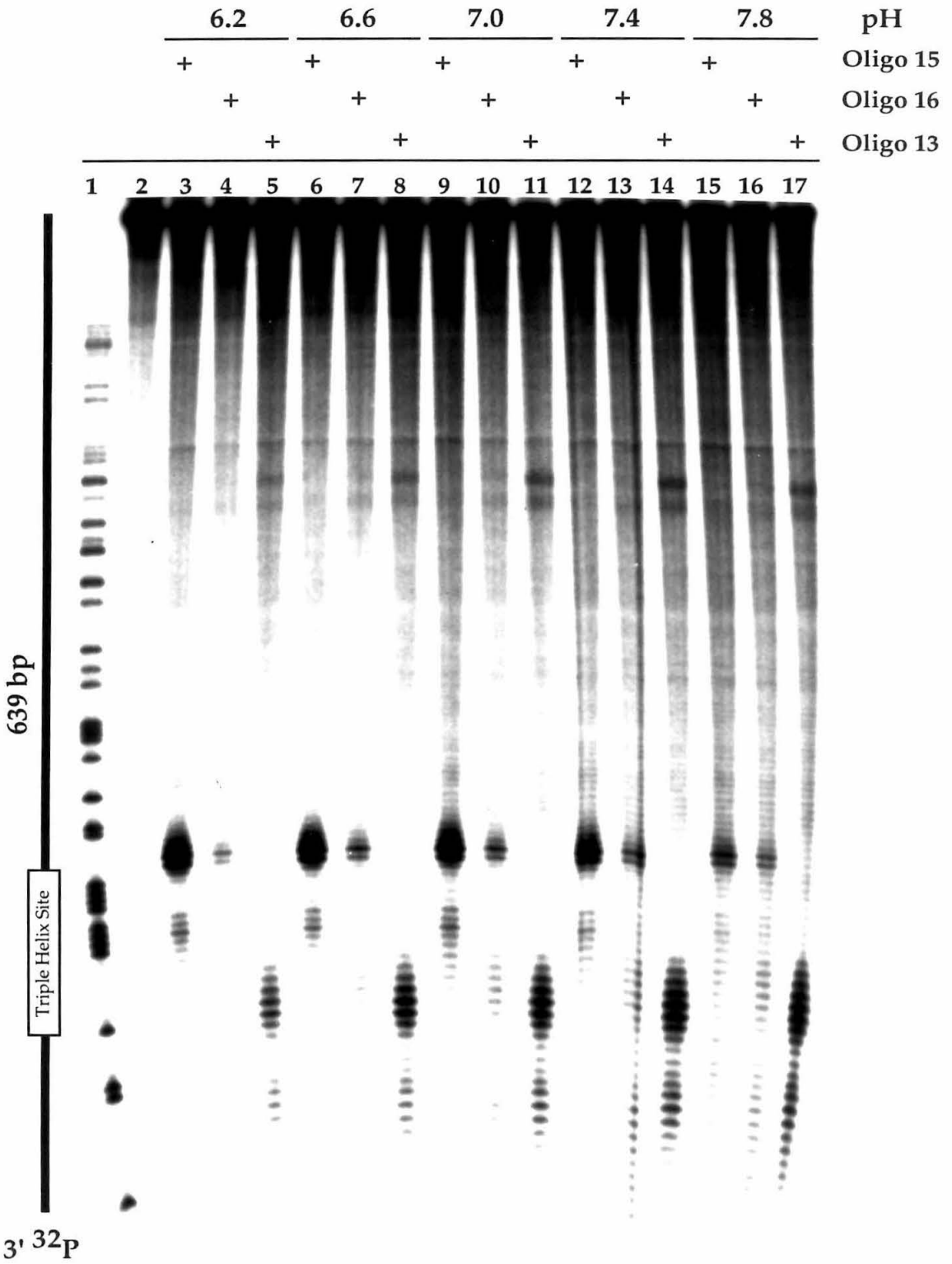
pH range (typically ≤ pH 7.0) within which this triple-helical structure can be formed. Koh and Dervan have shown that oligonucleotide **15** binds the HIV LTR purine tract poorly at alkaline pH due to the six contiguous ^{Me}C+GC triplets.[36] They described the design of a base analog, 3-methyl-5-amino-1*H*-pyrazolo[4,3-*d*]pyrimidin-7-one (P1), that can recognize GC base pairs at alkaline pH and allows for binding this sequence at pH≥7.0 with a pyrimidine motif oligonucleotide.

As expected, the binding of oligonucleotide **15** shows a strong dependence on pH through the range pH=7.0-7.8 with little cleavage detected at pH=7.8. However, the binding of the purine motif oligonucleotide **14** is relatively insensitive to pH from pH=6.6-7.8 (Figures 2.42 and 2.43). This is the expected result since no protonation is required to form the G•GC or T•AT triplets. Again, oligonucleotide **16** showed weak cleavage at all pH's. Thus, the formation of the purine motif triple helix is an alternative to the use of base analogs for triple helix formation at GC rich sites at pH≥7.0.

Conclusion and Prospectus. In conclusion, affinity cleaving has been used to characterize DNA triple helices with purine-rich third strands. These experiments have defined the groove location and orientation of the bound third strand. The common natural base triplets conferring the greatest stability have been identified and models for possible hydrogen bonding have been proposed. The energetic penalty for a single mismatched triplet in a 15 mer triple helix has been determined. A qualitative assessment of the effect of added ions, solution pH and sequence composition has also been made.

The power of using this structure as an approach to the sequence specific recognition of double helical DNA was first demonstrated by Hogan and co-workers with transcription inhibition studies.[5] Recently, the purine motif

Figure 2.42 Autoradiogram of an 8% denaturing polyacrylamide gel used to separate the cleavage products generated by oligonucleotidesEDTA•Fe **13, 15** and **16** on 3′ end labeled *HindIII/SspI* restriction fragment from pPBHIV as a function of pH. The reactions were carried out by combining a mixture of oligonucleotideEDTA and 2.5 equivalents of $Fe(NH_4)_2(SO_4)_2 \bullet 6H_2O$ with labeled restriction fragment [~ 100 nM in bp, ~ 20,000 cpm] in a solution of TrisOAc at the desired pH (50 mM), NaCl (10 mM), spermine (1 mM) and calf thymus DNA (100 μM in bp) and then incubating at 24° C for 1 H. Cleavage reactions were initiated by the addition of dithiothreitol (DTT) (4 mM) and allowed to proceed for 12 H at 25°C. (Lane 1) Products of an adenine specific cleavage reaction.[15] (Lane 2) Intact 3′ end labeled fragment incubated under the reaction conditions in the absence of oligonucleotideEDTA. (Lanes 3, 6, 9, 12 and 15) Cleavage products generated by 2.5 μM **15**. (Lane 4, 7, 10, 13 and 16) Cleavage products generated by 2.5 μM **16**. (Lane 5, 8, 11, 14 and 17) Cleavage products generated by 2.5 μM **13**.(Lanes 3-5) pH=6.2 (Lanes 6-8) pH=6.6 (Lanes 9-11) pH=7.0 (Lanes 12-14) pH=7.4 (Lanes 15-17) pH=7.8.


6.2
6.6
7.0
7.4
7.8
pH
Oligo 15
Oligo 16
Oligo 13
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
639 bp
Triple Helix Site
3' 32P

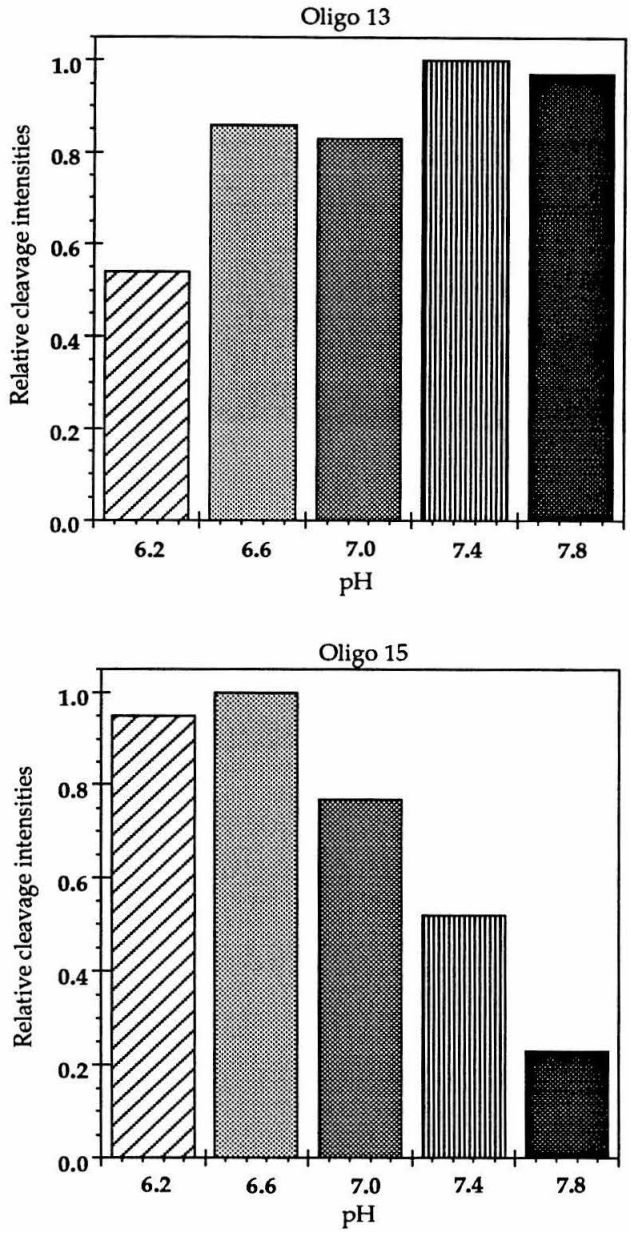


Figure 2.43 Bar graphs of relative cleavage intensities for **13** and **15** as a function of pH as derived from densitometry of the autoradiogram in Fig. 2.42.

triple helix has been employed to purify GC-rich duplexes by affinity capture,[34] induce mutations at a specific site in a phage genome,[35] and to bind double helical DNA containing all four base pairs by alternate strand triple helix formation.[36] However, the potential of this structure has likely not yet been fully realized. Structural analyses using NMR spectroscopy and X-ray crystallography will continue to identify the atomic positions and bond conformations in this triple helix, providing a detailed structural basis for the sequence specific recognition. Thermodynamic and kinetic studies on variables such as oligonucleotide sequence and length, ionic conditions and temperature will help define the scope and limitations to triple helix formation with the available bases and sugar-phosphate backbones. With this information, the purine motif triple helix might be used as a design template for the synthesis of new molecules containing nonnatural bases and backbones, capable of recognizing a wide variety of sequences under a number of different conditions.

Experimental Section

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. Deoxynucleoside triphosphates were purchased from Pharmacia as 100 mM solutions. 5'-(α-^{32}P) dATP, dGTP, dCTP and TTP (>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 maintenance of cell lines were accomplished using standard procedures.[37] Calf thymus DNA was purchased from Pharmacia.

Polyacrylamide and agarose gel electrophoresis were performed using 1XTBE buffer. 5′ and 3′ end labeling was accomplished using standard procedures.[37] Cerenkov radioactivity was measured on a Beckman LS 2801 scintillation counter.

Synthesis and purification of oligonucleotides. Oligonucleotides were prepared on an Applied Biosystems Model 380B DNA synthesizer with β-cyanoethyl phosphoramidites. Thymidine-EDTA (T*) was prepared as described and incorporated at the 3′ end of oligonucleotides via the 5'-O-DMT-thymidine-EDTA-triethylester 3'-succinyl controlled pore glass[20] or via the 5'-O-DMT-thymidine-EDTA-triethylester 3'-O-(2-cyanoethyl,N,N-diisopropyl) phosphoramidite.[14] Deprotection of T* oligonucleotides was carried out in 0.1 N NaOH at 55° C for 24 H. Oligonucleotides not containing T* were treated with concentrated NH_4OH at 55°C for 24 H. Oligonucleotides were purified either by polyacrylamide gel electrophoresis or by reverse phase chromatography on a Pharmacia FPLC system. In the gel purification technique, the deprotection solutions were neutralized with concentrated acetic acid and desalted on a Sephadex G-10 spun column followed by lyophilization. The residue resulting from lyophilization was resuspended in formamide loading buffer and then loaded onto a 20% polyacrylamide gel (19:1 monomer to bis) containing 8 M urea. The major UV absorbing band was cut out, crushed and the DNA was eluted with 200 mM NaCl, 1 mM EDTA at 37° C for 24 H. This solution was extensively (3-4 days) dialyzed against water. Alternatively, oligonucleotides were purified by using an Pharmacia FPLC ProRPC 10/10 (C2-C8) column with a 0-40% acetonitrile gradient in 100 mM triethylammonium acetate, pH=7. Chromatography solvent was removed by lyophilization. The resulting residue was dissolved in water and lyophilized further. Oligonucleotides were

quantitated by determining the concentration in aqueous solution by UV spectroscopy 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*) $cm^{-1}M^{-1}$. The solutions were lyophilized and the resulting oligonucleotides were stored as dry solids in 4-5 nmole fractions at -20° C.

Generation of labeled 39 bp oligonucleotide duplexes. Each of the single stranded 43 mer oligonucleotides (50 pmoles) was 5' end labeled with T4 polynucleotide kinase, γ-^{32}P-ATP(5000Ci/mmole, 10 mCi/mL). The labeled oligonucleotides were fractionated on a Sephadex G-50 spin column and precipitated with ethanol. This material was redissolved in 10 µL of H_2O and was added to its cold complementary strand (50 pmoles). The solutions were heated to 90° C for 3-4 minutes and allowed to cool for 1 H. The labeled duplexes were purified on a nondenaturing 15% polyacrylamide gel. The gel bands were visualized by autoradiography, and the appropriate bands were cut out of the gel and crushed. The DNA was eluted into 200 mM NaCl in a 37° C shaker for 24 H. The suspension was filtered, extracted with phenol and washed with diethylether. The DNA was precipitated with ethanol. Radioactivity was measured by scintillation counting.

Construction of plasmid DNA. Using T4 DNA ligase, the plasmid pPB19AG was prepared by ligation of the duplex formed between oligonucleotides of the sequence 5'-AATTCTCTCTAAAAAGGGAGGGGAGGGGAGGGAAAAA CTCTCT-3' and 5'-CTAGAGAGAGTTTTTCCCTCCCCTCCCCTCCCTTTT TAGAGAG-3' into pUC19, which had been previously digested with Eco RI and Xba I. The plasmid pPB19GC was analogously prepared from insertion of the duplex formed between 5'-AATTCTCTCTAAAAAGGGGGGGGAGGGG

AGGGAAAAACTCTCT-3' and 5'-CTAGAGAGAGTTTTTCCCTCCCCTCCC CCCCCTTTTTAGAGAG-3'. The plasmid pPB19CG was prepared from insertion of the duplex formed between 5'-AATTCTCTCTAAAAAGGGCGGGGAGGGG AGGGAAAAACTCTCT-3' and 5'-CTAGAGAGAGTTTTTCCCTCCCCTCCC CGCCCTTTTTAGAGAG-3'. The plasmid pPBSYMI containing duplex sequence **A** was prepared from insertion of the duplex formed between 5'-AATTCTCTCTGAAAGGGGAGGGGAAAGCTCTCT-3' and 5'-CTAGAGA CTTTCCCCTCCCCTTTCAGAGAG-3'. Plasmids pPBSYMII (duplex sequence **B**) and pPBSYMIII (duplex sequence **C**) were similarly prepared with the appropriate synthetic oligonucleotide duplexes. The plasmid pPBHIV was prepared from insertion of the duplex formed between 5'-AATTATTTTAAAAGAAAAGGGGGGACTGGAAG-3' and 5'-CTAGC TTCCAGTCCCCCCTTTTCTTTTAAAAAT-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. 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 labeled plasmid restriction fragments. Labeled restriction fragments from plasmids were generated as follows. Plasmid DNA (5-20 μg) was linearized with *Hind III* or *EcoO109I* followed by 3′ end labeling by filling in using the Klenow fragment of DNA polymerase I or Sequenase v. 2.0 and α-^{32}P dATP, dCTP, dTTP or dGTP. The labeled plasmid DNA was digested with *Ssp I* and the resulting fragment was isolated by nondenaturing 5% PAGE. Gel bands were visualized by autoradiography, cut out and eluted with 200 mM NaCl at

37°C for 24H. The suspension was filtered to remove polyacrylamide and the DNA was purified by ethanol precipitation.

Affinity Cleaving. In a typical cleavage experiment, the labeled DNA was mixed with salts, buffer, spermine, and the oligonucleotide EDTA•Fe(II). This solution was incubated for one hour at 24° C before the cleavage reaction was initiated by the addition of DTT giving a final volume of 20 μL. For reactions on plasmid restriction fragments, the reactions were stopped by precipitation with ethanol. For reactions on oligonucleotide duplexes, the reactions were stopped by freezing and lyophilization. Electrophoretic separation of the cleavage products was achieved on denaturing polyacrylamide gels. Specific reaction conditions are stated in the figure legends.

Generation of autoradiograms. Agarose and 8% polyacrylamide gels were transferred to filter paper and dried with a BioRad Model 483 Slab drier for 1 H at 80° C, and the cleavage products were visualized by exposure to X-ray film (Kodak) at -70° C with an intensifying screen. Polyacrylamide gels of 20% were directly exposed to X-ray film without drying. For the generation of histograms, the autoradiograms were scanned by an LKB Ultroscan XL Laser Densitometer.

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.[38] A Molecular Dynamics 400S PhosphorImager was used to obtain all data from the storage screens. The data were typically 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. For the generation of histograms, signal intensity was quantitated along a line drawn down the lane of interest and a line graph was generated using ImageQuant.

Quantitation of relative cleavage intensities by laser densitometry. For determination of the relative cleavage intensities generated by oligonucleotidesEDTA, autoradiograms of the polyacrylamide gels used to separate cleavage products were scanned by an LKB Ultroscan XL Laser Densitometer. The relative intensities were calculated by normalizing the area integration for the largest peak in that lane to the area integration for the largest peak on the gel.

Quantitation of cleavage efficiencies by storage phosphor autoradiography. For analysis of the relative stabilities of the 16 common natural base triplets, absolute cleavage efficiencies were determined by calculating the ratio of site specific cleavage products generated for each triplet combination to the total amount of radiolabeled DNA in that lane on each gel. To account for background cleavage generated by unbound oligonucleotide-EDTA•Fe, the value for site specific cleavage was calculated by subtracting the integral value corresponding to a reference site in each lane from the integral value at the main cleavage site. Relative efficiencies were calculated by normalizing the absolute cleavage efficiency for each triplet combination to that of the combination yielding the most efficient cleavage in each experiment.

Quantitative affinity cleavage titrations. In a typical titration, a stock solution was prepared containing TrisOAc at the desired pH, spermine, NaCl and calf

thymus DNA and the ^{32}P labeled restriction fragment. The stock solution was allowed to equilibrate for 5 min at room temperature and distributed among the reaction tubes. OligonucleotideEDTA was dissolved in an aqueous solution of $Fe(NH_4)(SO_4)_2 \bullet 6H_2O$ to produce a 10X stock solution which was serially diluted. The 10X stock solutions of oligonucleotideEDTA were added to appropriate reaction tubes. The resulting solutions were allowed to equilibrate for 24-48 H at 24°C. The cleavage reactions were initiated by the addition of an aqueous DTT solution (4 mM final concentration) and incubated for 4-6 H at 24°C. The final concentrations of components are listed in figure legends and tables. The specific activity of labeled duplex was typically 10000-20000 cpm in each reaction tube with a concentration of less than 0.05 nM. Reactions were stopped by precipitation with ethanol. The precipitate was dissolved in water and lyophilized. The resulting pellet was dissolved in formamide TBE loading buffer and the resulting solutions were assayed for Cerenkov radioactivity by scintillation counting. The samples were heat denatured at 90°C for 4 min and loaded onto an 8% denaturing polyacrylamide gel. After electrophoresis in TBE buffer, the gel was dried on a slab gel drier and exposed to a storage phosphor screen at 24°C.

Quantitative affinity cleavage titration fitting procedure. The site specific cleavage intensities for all oligonucleotideEDTA concentrations were calculated using eq. 5, where I_{tot} and I_{ref} represent the cleavage intensities in the site and reference blocks, respectively. Intensities were determined from the integration of storage phosphor images as described.

$$I_{site} = I_{tot} - \lambda I_{ref} \qquad (5)$$

For each experiment, λ was calculated from the average of the I_{tot}/I_{ref} ratios for the three lowest oligonucleotideEDTA concentrations. A theoretical binding curve was fit to the experimental data using the apparent maximum cleavage (I_{sat}) and the K_T as adjustable parameters:

$$I_{fit} = I_{sat} \frac{K_T [O]_{tot}}{1 + K_T [O]_{tot}} \qquad (6)$$

The difference between I_{fit} and I_{site} for all data points was minimized using the nonlinear least-squares fitting procedure of KaleidaGraph (version 2.1; Abelbeck Software) running on a Macintosh IIfx computer. Fits were performed without weighting data points. All data points were included in the fitting procedure unless visual inspection of the computer image revealed a flaw at either the target site or reference blocks. Data from experiments for which fewer than 75% of the data lanes were usable were discarded. K_T values are reported as the mean of three or four titrations ± one standard deviation.

References and Notes

1. (a) Lipsett, M.N. Biochem. Biophys. Res. Comm. 1963, 11, 224. (b) Lipsett, M. N. *J. Biol. Chem.* **1964**, 239, 1256.
2. Mark, C.; Thiele, D.; *Nucleic Acids Res.* **1978**, 5, 1017.
3. Broitman, S. L.; Im, D. D.; Fresco, J. Y. *Proc. Natl. Acad. Sci. USA* **1988**, 85, 3781.
4. Kohwi, Y.; Kohwi-Shigematsu, T. *Proc. Natl. Acad. Sci. USA* **1988**, 85, 3781.
5. Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, J.; Hogan, M. E. *Science* **1988**, 241, 456.

6. (a) Hoogsteen, K.; *Acta Cryst.* **1959**, 12, 822. (b) Durland, R.H.; Kessler, D.J.; Gunnell, S.; Duvic, M.; Pettitt, B.M.; Hogan, M.E.; *Biochemistry* **1991**, 30, 9246.

7. (a) Kang, C-H.; Zhang, X.; Ratliff, R.; Mayzis, R.; Rich, A.; *Nature* **1992**, 356, 126. (b) Smith, F.W.; Feigon, J.; *Nature* **1992**, 356, 164.

8. Bugg, C.E.; Thewalt, U.T.; Marsh, R.E.; *Biochem. Biophys. Res. Comm.* **1968**, 33, 436.

9. Pranata, J.; Wierschke, S.G.; Jorgensen, W.L.; *J. Am. Chem. Soc.* **1991**, 113, 2810.

10. Williams, N. G.; Williams, L. D.; Shaw, B. R.; *J. Am. Chem. Soc.* **1989**, 111, 7205.

11. Klug, A.; Ladner, J.; Robertus, J.D.; *J. Mol. Biol.* **1974**, 89, 511.

12. Maher, L. J.; Wold, B.; Dervan, P. B. *Science* **1989**, 245, 725.

13. Griffin, L. C.; Dervan, P. B. *Science* **1989**, 245, 967.

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

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

16. Moser, H. E.; Dervan, P.B. *Science* **1987**, 238, 645.

17. (a) Chen, F-M.; *Biochemistry* **1991**, 30, 4472. (b) Pilch, D.S.; Levenson, C.; Shafer, R.H.; *Biochemistry* **1991**, 30, 6081. (c) Radhakrishnan, I.; de los Santos, C.; Patel, D.J.; *J. Mol. Biol.* **1991**, 221, 1403. (d) Milligan, J.F.; Krawczyk, S.H.; Wadwani, S.; Matteucci, M.D.; *Nucleic Acids Res.* **1993**, 21, 327. (e) Roy, C.; *Nucleic Acids Res.* **1993**, 21, 2845. (f) Radhakrishnan, I.; Patel, D.J.; *J. Am. Chem. Soc.* **1993**, 114, 1615. (g) Radhakrishnan, I.; Patel, D.J.; *Structure* **1993**, 1, 135.

18. Saenger, W.; *Principles of Nucleic Acid Structure* (Springer-Verlag, NewYork, 1984).

19. Beal, P.A.; Dervan, P.B.; *Science* **1991**, 251, 1360.

20. Strobel, S.A.; Ph.D. Dissertation, California Institute of Technology, 1992.

21. Strobel, S.A.; Doucette-Stamm, L.A.; Riba, L.; Housman, D.E.; Dervan, P.B.; *Science* **1991**, 254, 1639.

22. Griffin, J.H.; Ph.D. Dissertation, California Institute of Technology, 1989.

23. (a) Campos, J.L.; Subirana, J.A.; *J. Biomolec. Struct. Dynam.* **1991**, 8, 793. (b) Beltran, R.; Martinez-Balbas, A.; Bernues, J.; Bowater, R.; Azorin, F.; *J. Mol. Biol.* **1993**, 230, 966.

24. Maher, L.J.; Dervan, P.B.; Wold, B.J.; *Biochemistry* **1990**, 29, 8820.

25. (a) Williamson, J.R.; Raghuraman, M.K.; Cech, T.R.; *Cell* **1989**, 59, 871. (b) Raghuraman, M.K.; Cech, T.R.; *Nucleic Acids Res.* **1990**, 18, 4543. (c) Lee, J.; *Nucleic Acids Res.* **1990**, 18, 6057. (d) Hardin, C.C.; Henderson, E.; Watson, T.; Prosser, J.K.; *Biochemistry* **1991**, 30, 4460. (e) Hardin, C.C.; Watson, T.; Corregan, M.; Bailey, C.; *Biochemistry* **1992**, 31, 833.

26. Singleton, S.F.; Dervan, P.B.; *J. Am. Chem. Soc.* **1992**, 114, 6956.

27. Beal, P.A.; Dervan, P.B.; *Nucleic Acids Res.* **1992**, 20, 2773.

28. Kiessling, L.L.; Griffin, L.C.; Dervan, P.B.; *Biochemistry* **1992**, 31, 2829.

29. Stilz, H.U.; Dervan, P.B.; *Biochemistry* **1993**, 32, 2177.

30. Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.; *Cell* **1985**, 40, 9.

31. Koh, J.S.; Dervan, P.B.; *J. Am. Chem. Soc.* **1992**, 114, 1470.

32. Sun, J.S.; de Bizemont, T.; Montenay,-Garestier, T.; Helene, C.; C.R. *Acad. Sci. Paris Ser. 3* **1991**, 313, 585.

33. Singleton, S.F.; Dervan, P.B.; *Biochemistry* **1992**, 31, 10995 and references therein.

34. Takabatake, T.; Asada, K.; Uchimura, Y.; Ohdate, M.; Kusukawa, N.; *Nucleic Acids Res.* **1992**, 20, 5853.

35. Havre, P.A.; Gunther, E.J.; Gasparro, F.P.; Glazer, P.M.; *Proc. Natl. Acad. Sci. USA* **1993**, 90, 7879.

36. (a) Beal, P.A.; Dervan, P.B.; *J. Am. Chem. Soc.* **1992**, 114, 4976. (b) Jayasena, S.D.; Johnston, B.H.; *Nucleic Acids Res.* **1992**, 20, 5279.

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

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

Chapter 3

Recognition of Double Helical DNA by Alternate Strand Triple Helix Formation

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 been used to mediate single site specific cleavage of human chromosomal DNA[1k] as well as to block transcription *in vitro*.[1m,3a] The ability 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 specifically to purine tracts of double helical DNA in the major groove parallel to the Watson-Crick (WC) purine strand (pyrimidine motif). 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 to purine tracts of double helical DNA in the major groove antiparallel to the WC purine strand (purine motif).[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 consecutive base pairs without large 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 dicussed above, the third strands bind in opposite orientations with respect to the purine strand of the duplex. Therefore, it seemed 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 3.1). The base triplets of such a structure are shown in Figure 3.2, where thymidine and 5-methyldeoxycytidine are bound to one strand of the duplex and deoxyadenosine and deoxyguanosine to the other.

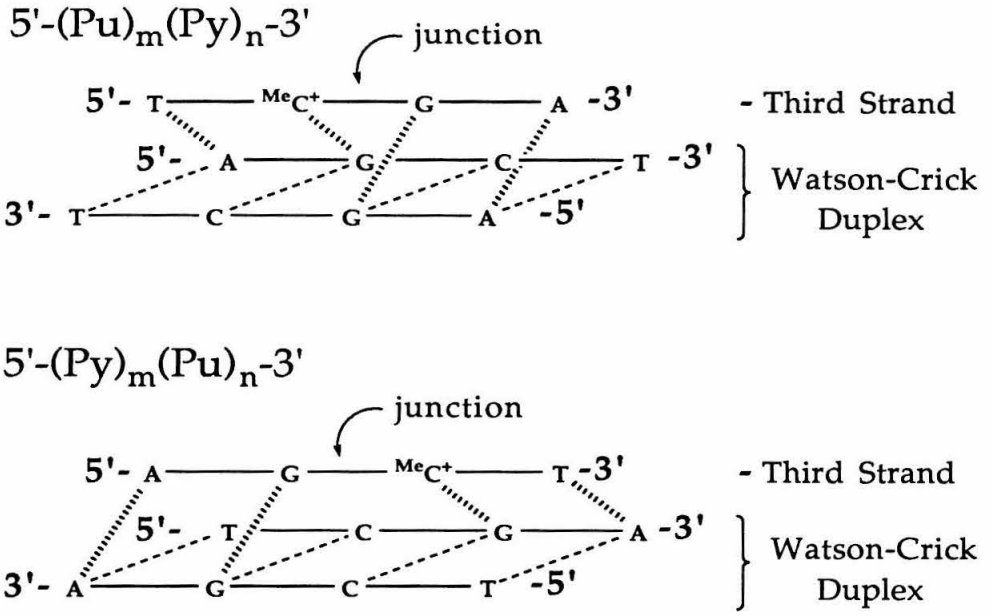


Figure 3.1 Models representing oligonucleotides binding in the major groove and simultaneously recognizing both strands of the duplex by forming hydrogen bonds to the purine in a WC base pair. Base triplet specificities and required strand orientations are maintained by alternating between triple helix motifs at junctions between purines and pyrimidines on one strand of the duplex. (Top) Triple helix formed at a 5′-$(Pu)_m(Py)_n$-3′ target site. (Bottom) Triple helix formed at a 5′-$(Py)_m(Pu)_n$-3′ target site.

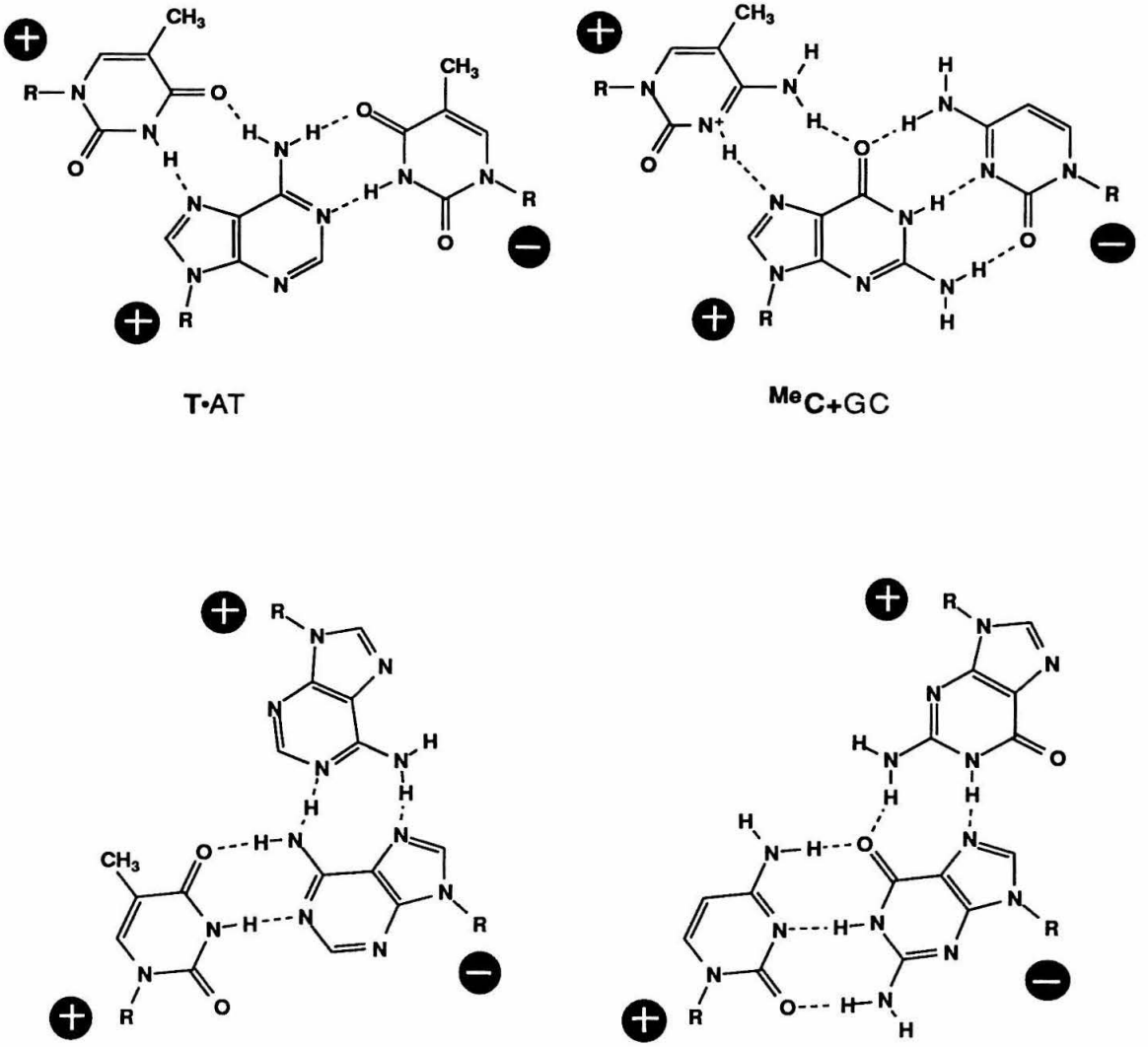


Figure 3.2 Models for base triplets formed when an oligonucleotide binds in the major groove and simultaneously recognizes both strands. All bases are in the *anti* conformation. Plus and minus indicate relative strand orientations.

To test this hypothesis, two duplex sequences were prepared, each containing 18 bp binding sites composed of two adjacent nine bp half sites on alternate strands capable of being bound by either pyrimidine or purine oligonucleotides. The sequence content is similar 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. Modeling of these two sequences suggested that the adjacent triple helix binding sites are in different relative positions in space due to the right-handed helical nature of the double helix. Therefore, designing triple helix forming oligonucleotides to bind to these two sequence types would likely require different structures at the junction where the transition from one motif to the other takes place.

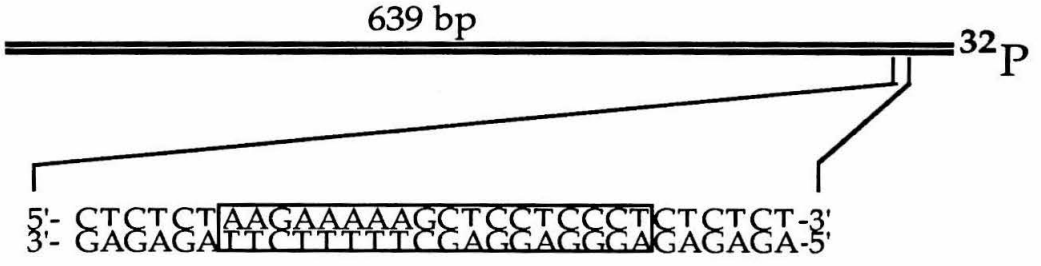
The 5'-$(Pu)_9(Py)_9$-3 sequence. The potential binding site, 5'-AAGAAAAAGCTCCTCCCT-3', shown in Figure 3.3, consists of two adjacent nine bp purine tracts on alternate strands with a 5'-GC-3' junction. Oligonucleotides **1-5** were designed to bind this sequence by the formation of a pyrimidine motif triple helix at the 5' half site and/or a purine motif 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 3.4). Interestingly, a 16 mer oligonucleotide lacking the center two nucleotides cleaves the target equally as efficiently as **1** (data not shown).

DNase I footprinting was used to confirm the binding site size for **5**, identical in sequence to **1** and **2**, lacking thymidine-EDTA (Figure 3.5). As expected from the affinity cleaving results, **5** protects both strands of the duplex from cleavage by 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 $K_a = 5.7 \pm 1.4 \times 10^6\ M^{-1}$ (pH=7.0, 1.5 mM spermine, 5 mM NaCl, 10 mM $MgCl_2$, 10 mM $CaCl_2$) corresponding to a binding free energy $\Delta G_b = -9.2 \pm 0.1$ kcal/mol (Figure 3.6).[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 ^{Me}C+GC and CG•G base triplets. (See Figure 3.2.) 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 3.7).

The results of these experiments are summarized in Figure 3.8. These observations 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.



1 5'-*TTCTTTTTCGAGGAGGGT-3'

2 5'- TTCTTTTTCGAGGAGGGT*-3'

3 5'-*TTCTTTTTC -3'

4 5'- GAGGAGGGT*-3'

5 5'- TTCTTTTTCGAGGAGGGT-3'

Figure 3.3 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 3.4 (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 $Fe(NH_4)_2(SO_4)_2 \bullet 6H_2O$ (500 nM) with the 3' end ^{32}P labeled 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.[17] (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.

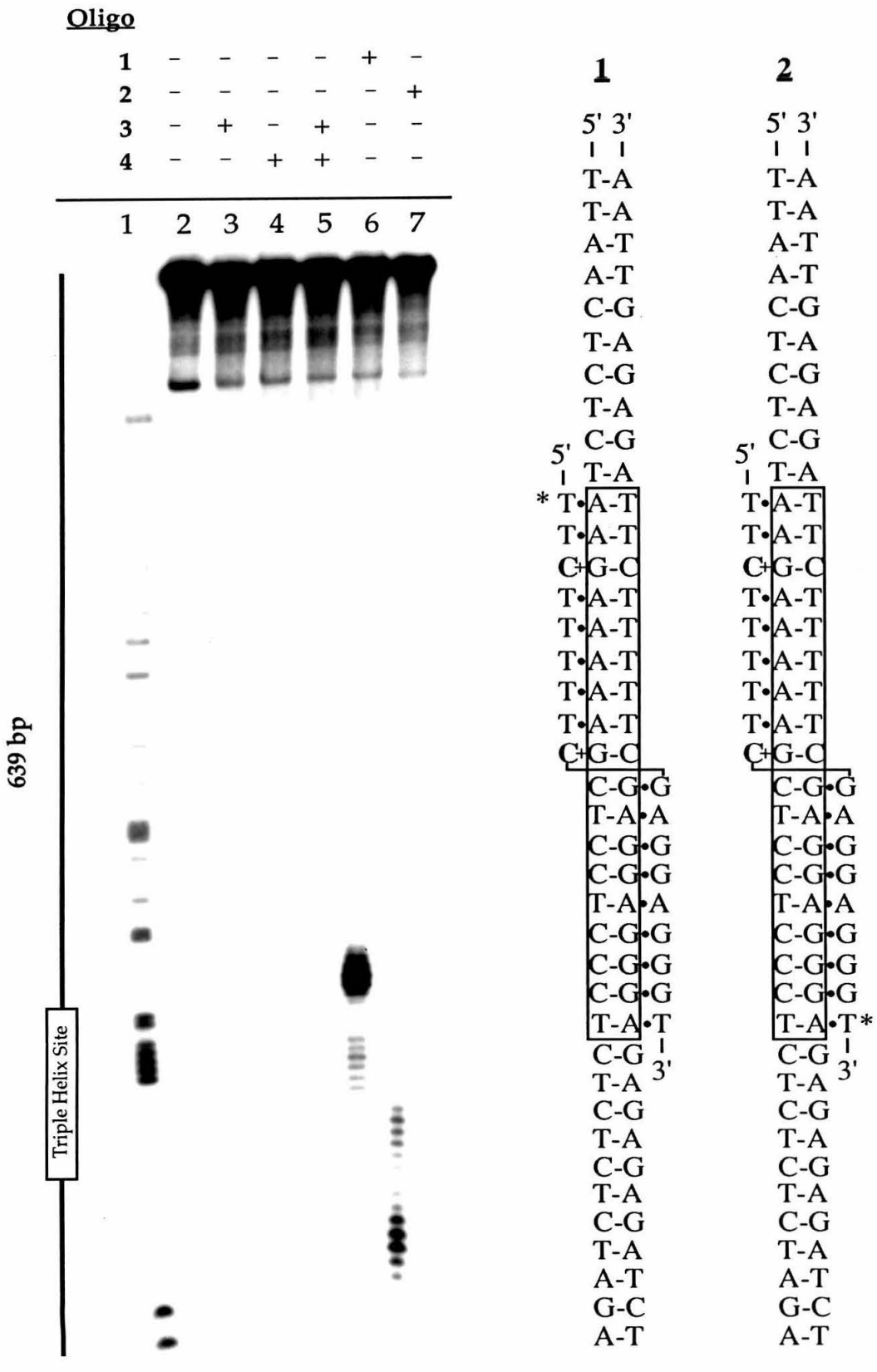
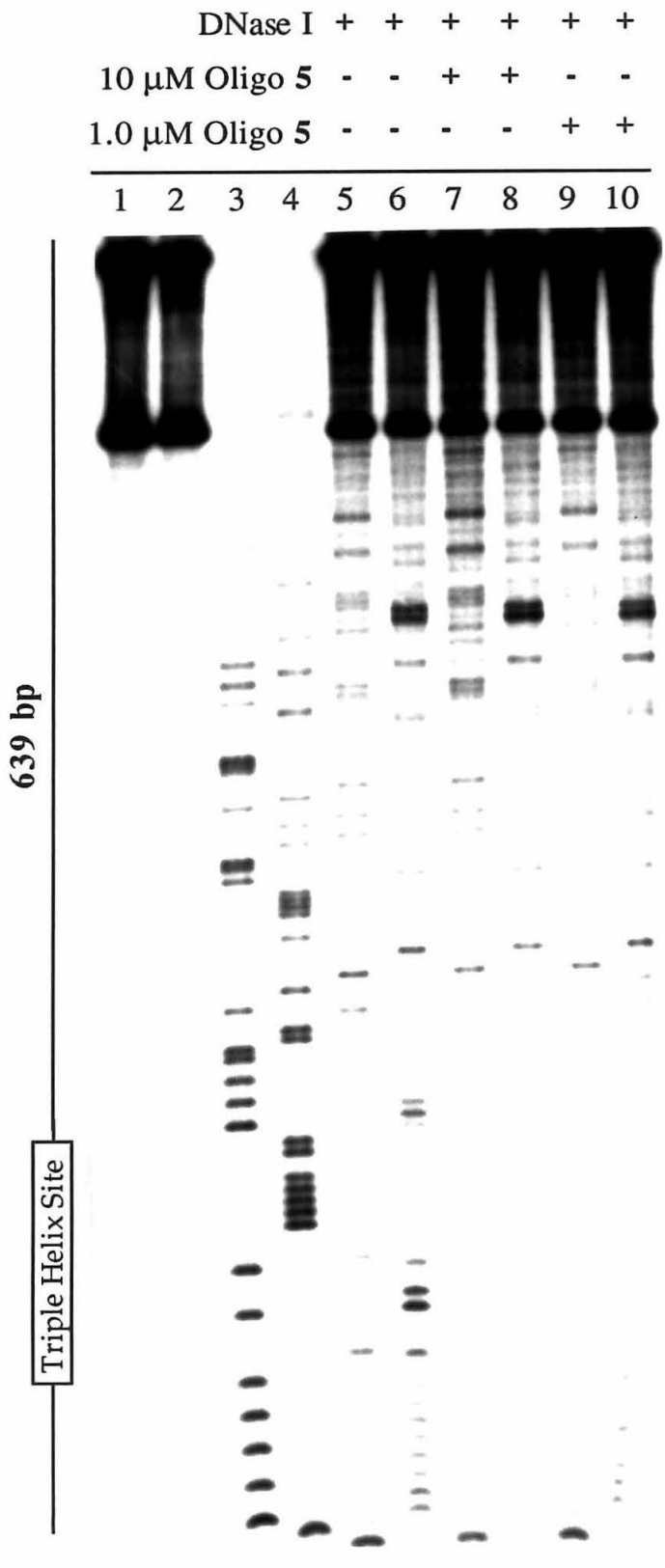

Oligo
1 - - - - + -
2 - - - - - +
3 - + - + - -
4 - - + + - -
1 2 3 4 5 6 7
639 bp
Triple Helix Site
1
2
5' 3'
T-A
T-A
A-T
A-T
C-G
T-A
C-G
T-A
C-G
T-A
5'
*
T•A-T
T•A-T
C+G-C
T•A-T
T•A-T
T•A-T
T•A-T
T•A-T
C+G-C
C-G•G
T-A•A
C-G•G
C-G•G
T-A•A
C-G•G
C-G•G
C-G•G
T-A•T
3'
C-G
T-A
C-G
T-A
C-G
T-A
C-G
T-A
A-T
G-C
A-T
T-A•T*

Figure 3.5 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), $MgCl_2$ (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 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.[17] (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**.


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

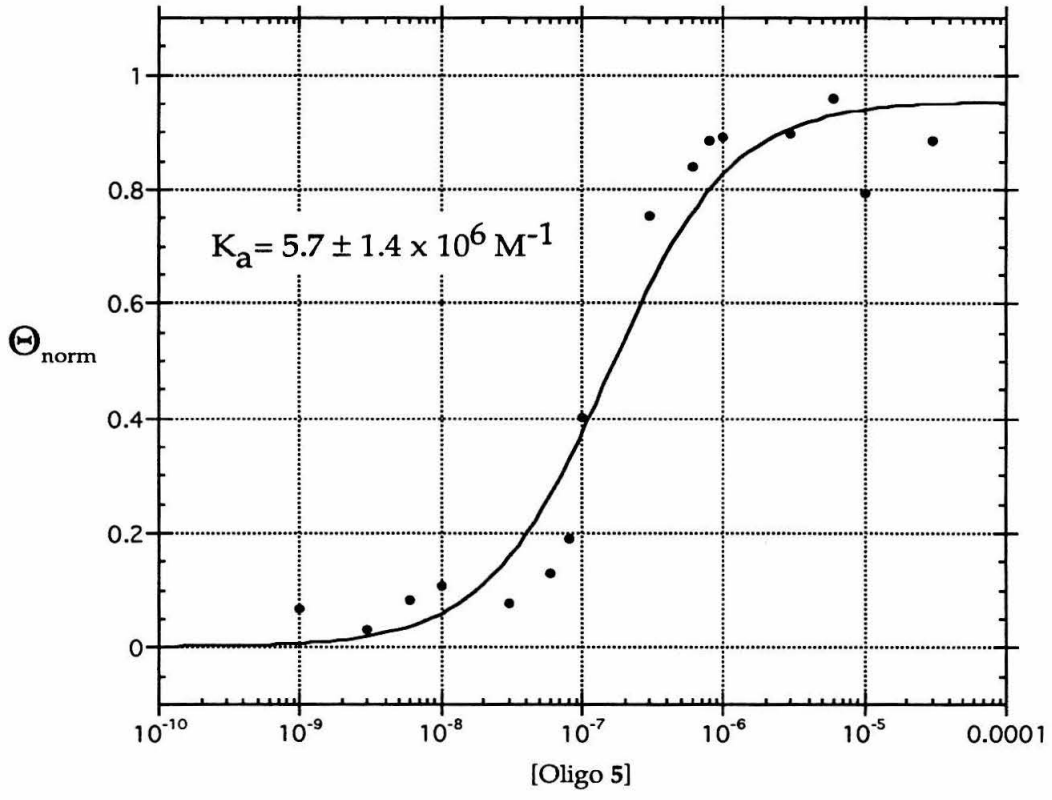


Figure 3.6 Data for DNase I footprint titrations involving **5**. The Θ_{norm} values were determined as described in the experimental section. The line indicates the best-fit binding isotherm from a nonlinear least squares analysis using eq. 2.

Figure 3.7 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.[17] (Lanes 5 and 6) Cleavage products generated in the 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**.

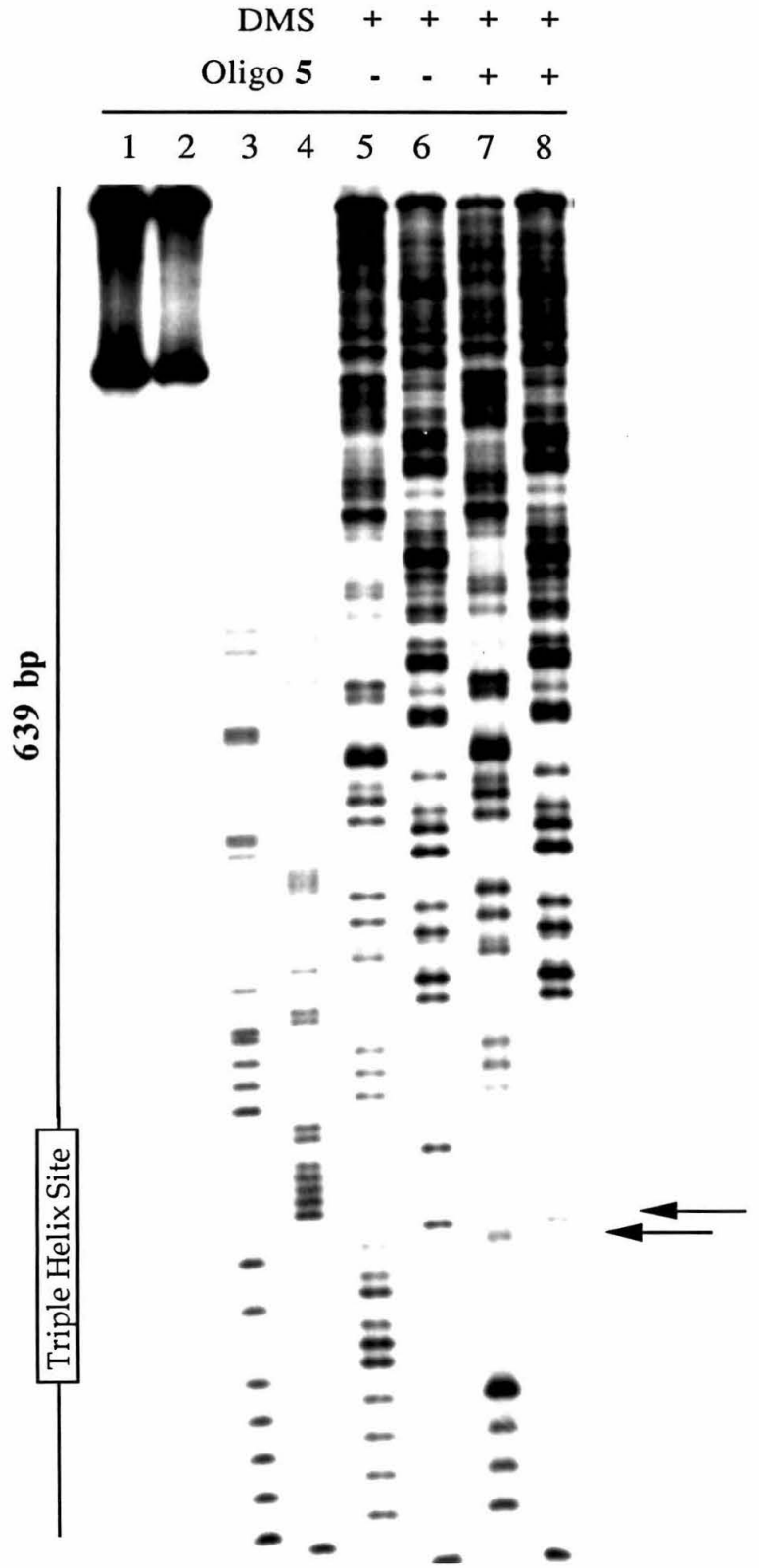
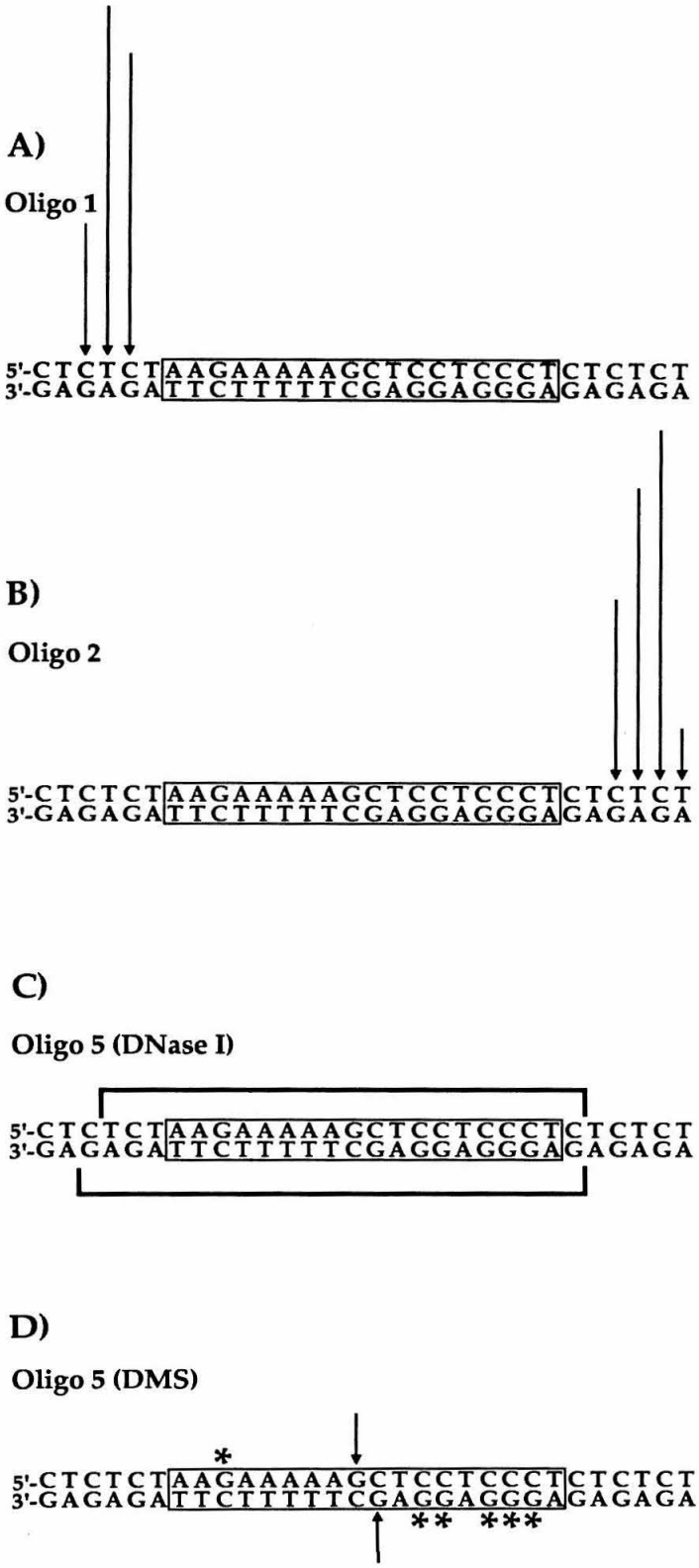

DMS + + + +
Oligo 5 - - + +
1 2 3 4 5 6 7 8
639 bp
Triple Helix Site

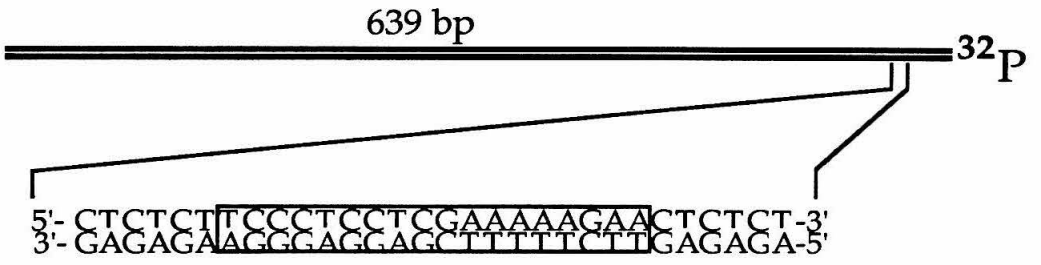
Figure 3.8 **(A)** Cleavage pattern generated by oligonucleotide **1**. The box indicates the double stranded sequence bound. Positions of the arrows show the sites of cleavage and heights indicate extent of cleavage at that site. **(B)** Cleavage pattern generated by oligonucleotide **2**. **(C)** DNase 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.


A)
Oligo 1
5'-CTCTCTAAGAAAAAGCTCCTCCCTCTCTCT
3'-GAGAGATTCTTTTTTCGAGGAGGGAGAGAGA
B)
Oligo 2
5'-CTCTCTAAGAAAAAGCTCCTCCCTCTCTCT
3'-GAGAGATTCTTTTTTCGAGGAGGGAGAGAGA
C)
Oligo 5 (DNase I)
5'-CTCTCTAAGAAAAAGCTCCTCCCTCTCTCT
3'-GAGAGATTCTTTTTTCGAGGAGGGAGAGAGA
D)
Oligo 5 (DMS)
5'-CTCTCTAAGAAAAAGCTCCTCCCTCTCTCT
3'-GAGAGATTCTTTTTTCGAGGAGGGAGAGAGA

The 5'-(Py)9(Pu)9-3' sequence. The target site, 5'-TCCCTCCTCGAAAAAGAA-3', shown in Figure 3.9, is similar in sequence content to that shown in Figure 3.3, 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 purine motif triple helix at the 5' half site and/or a pyrimidine motif 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 3.10, 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 (Figure 3.10, lanes 8 and 9).[9] One thymidine is not sufficient, as shown by the low cleavage efficiency of **7** (Figure 3.10, 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 3.11). 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 (Figure 3.12). DMS footprinting studies on this structure have shown that all the guanines in the binding site are protected equally from methylation.

The intercalating 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 3.13). In addition to protection from cleavage within the binding site, a region of



6 5'-*TGGGAGGAGCTTTTTCTT-3'

7 5'-*TGGGAGGAGTCTTTTTCTT-3'

8 5'-*TGGGAGGAGTTCTTTTTCTT-3'

9 5'- TGGGAGGAGTTCTTTTTCTT*-3'

10 5'-*TGGGAGGAG-3'

11 5'- CTTTTTCTT*-3'

12 5'- TGGGAGGAGTTCTTTTTCTT-3'

Figure 3.9 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 3.10 (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 $Fe(NH_4)_2(SO_4)_2 \bullet 6H_2O$ (500 nM) with the 5' end ^{32}P labelled Hind III/Ssp I 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.[17] (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.

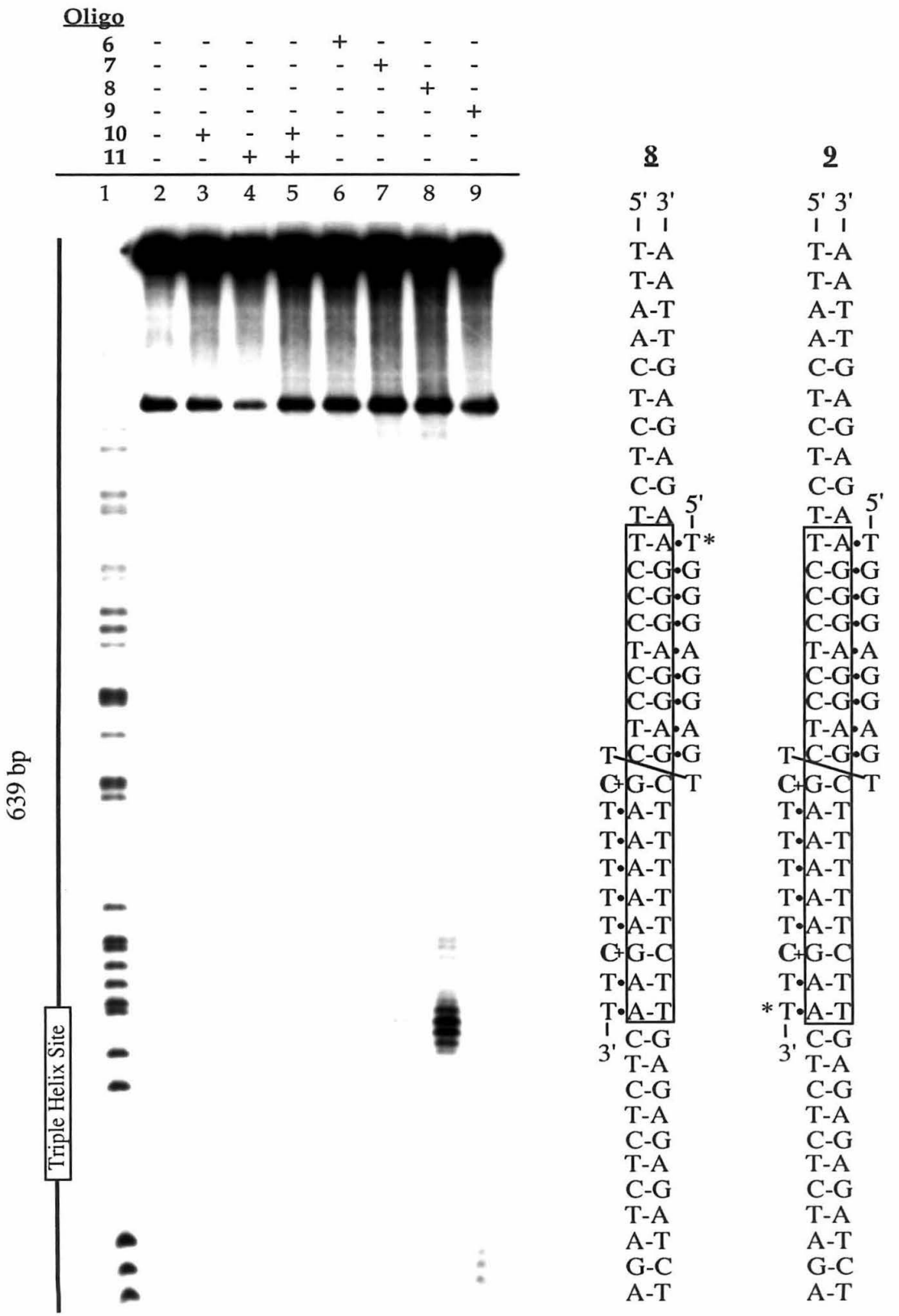
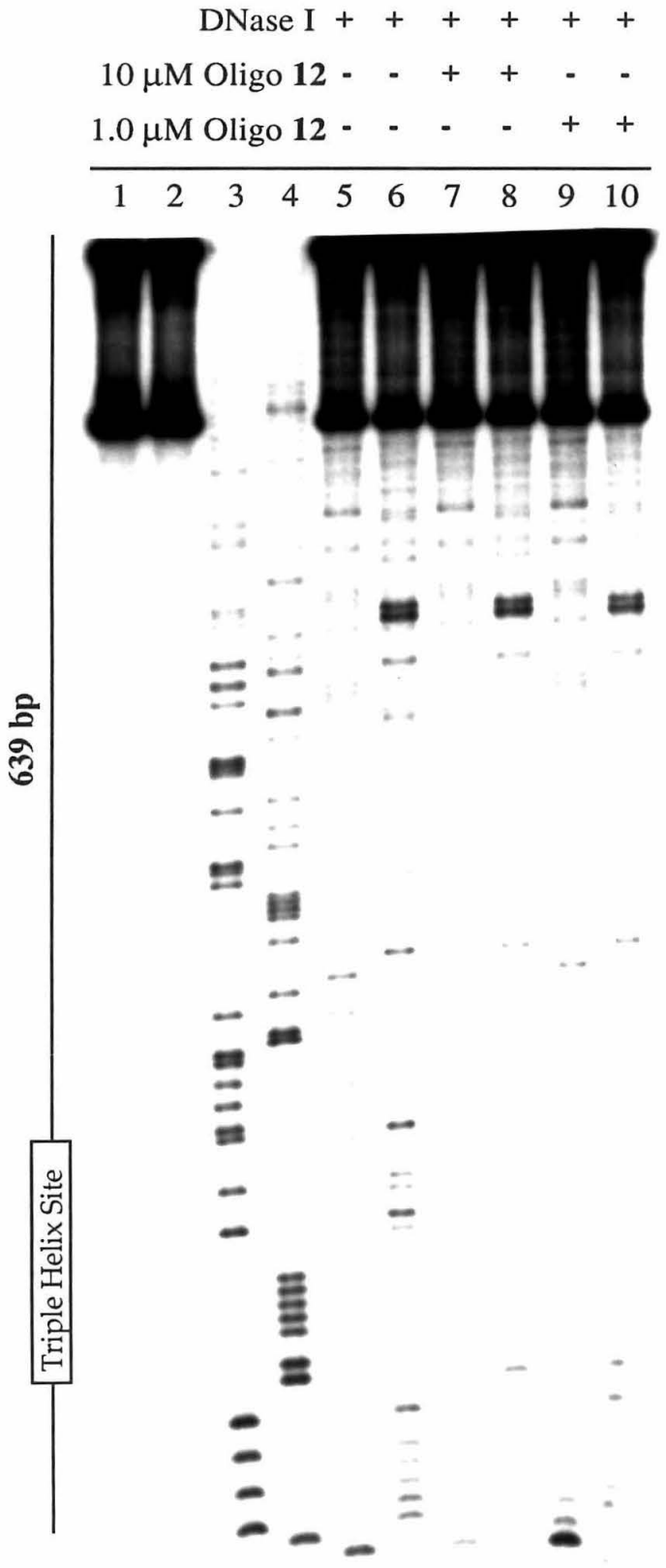

Oligo
6 - - - - + - - -
7 - - - - - + - -
8 - - - - - - + -
9 - - - - - - - +
10 - + - + - - - -
11 - - + + - - - -
1 2 3 4 5 6 7 8 9
639 bp
Triple Helix Site
8
9
5' 3'
T-A
T-A
A-T
A-T
C-G
T-A
C-G
T-A
C-G
T-A
5'
T-A•T*
C-G•G
C-G•G
C-G•G
T-A•A
C-G•G
C-G•G
T-A•A
C-C-G•G
C+G-C T
T•A-T
T•A-T
T•A-T
T•A-T
T•A-T
C+G-C
T•A-T
T•A-T
3'
C-G
T-A
C-G
T-A
C-G
T-A
C-G
T-A
A-T
G-C
A-T
5'
T-A•T
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
C+G-C
T•A-T
*T•A-T
3'
C-G
T-A
C-G
T-A
C-G
T-A
C-G
T-A
A-T
G-C
A-T

Figure 3.11 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 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.[17] (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**.


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

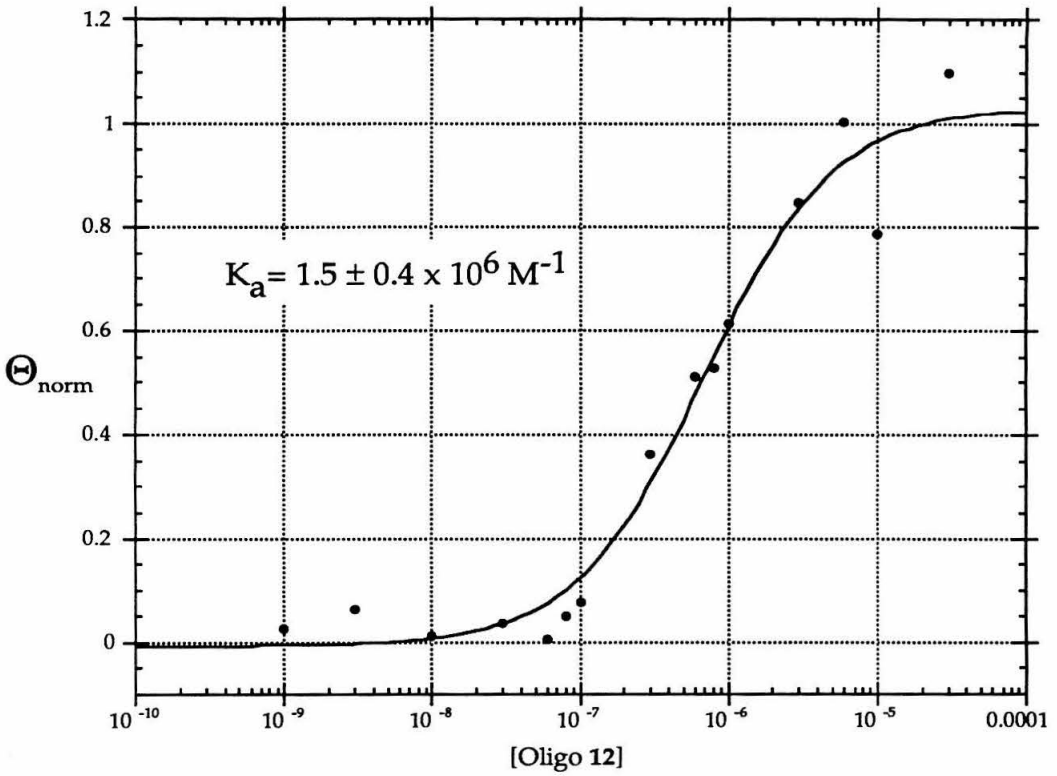
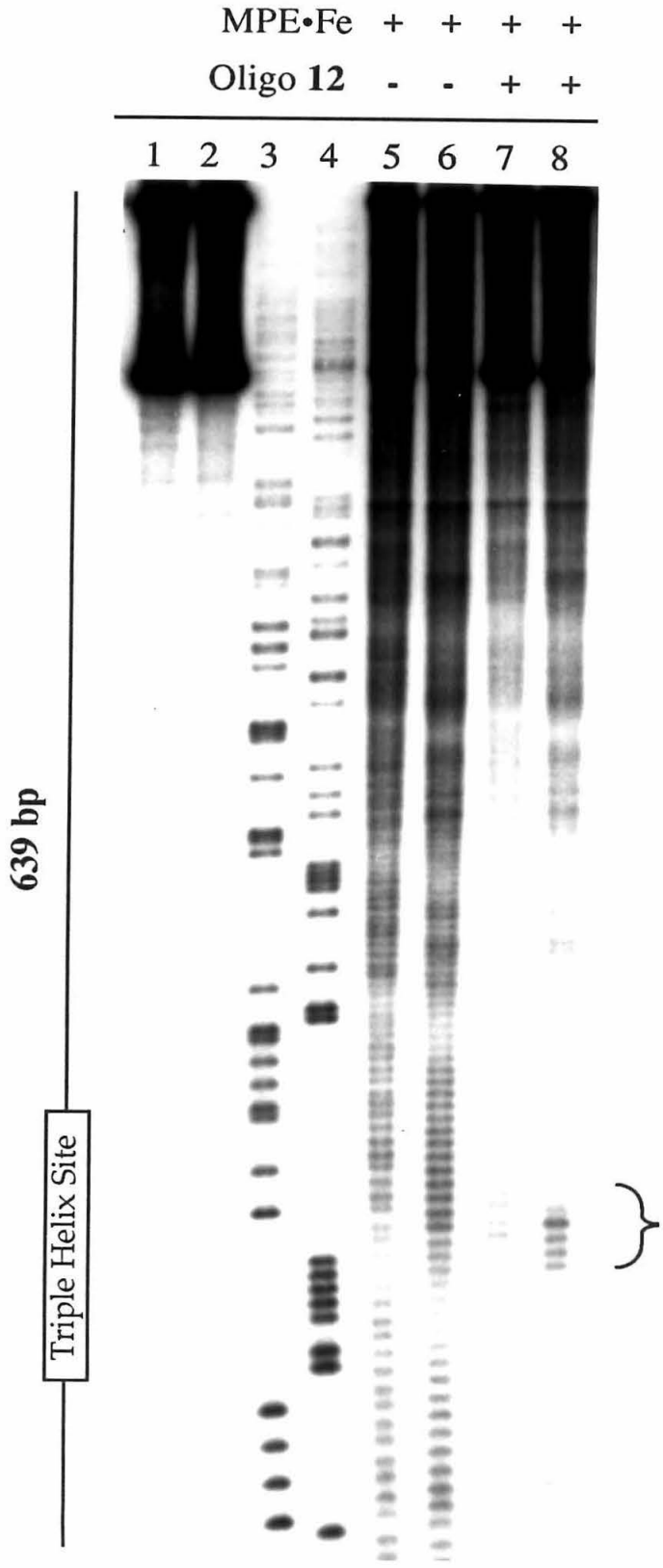


Figure 3.12 Data for DNase I footprint titrations involving **12**. The Θ_{norm} values were determined as described in the experimental section. The line indicates the best-fit binding isotherm from a nonlinear least squares analysis using eq. 2.

Figure 3.13 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), $MgCl_2$ (10 mM), $CaCl_2$ (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.[17] (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.


MPE•Fe + + + +
Oligo 12 - - + +
1 2 3 4 5 6 7 8
639 bp
Triple Helix Site

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 3.14. These results show that targeting the 5'-(Py)9(Pu)9-3' sequence requires at least a two nucleotide linker between the two binding domains. This is in contrast to the 5'-(Pu)9(Py)9-3' 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)9(Pu)9-3' site is less stable, with an association constant lower by a factor of four. Interestingly, a recent report on alternate strand triple helix formation has also shown that triple helices formed at duplex sequences of the type 5′-$(Py)_m(Pu)_m$-3′ are less stable than those at the 5′-$(Pu)_m(Py)_n$-3′ sequence.[5d] In addition, the MPE•Fe reactivity pattern indicates that the base pairs near the site of crossover between strands provide an unusually stable binding site for an intercalator. In a related study, 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]

It is clear that the stucture and stability of the alternate strand triple helix depends on whether the duplex target is 5'-$(Pu)_m(Py)_n$-3' or 5'-$(Py)_m(Pu)_n$-3'. Analysis of models of these sequences with oligonucleotides bound to adjacent binding sites may explain some of these observations. For the sequence 5'-$(Pu)_m(Py)_n$-3', the site of crossover between strands is located where adjacent binding sites overlap (Figure 3.15, 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

Figure 3.14 (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. 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]

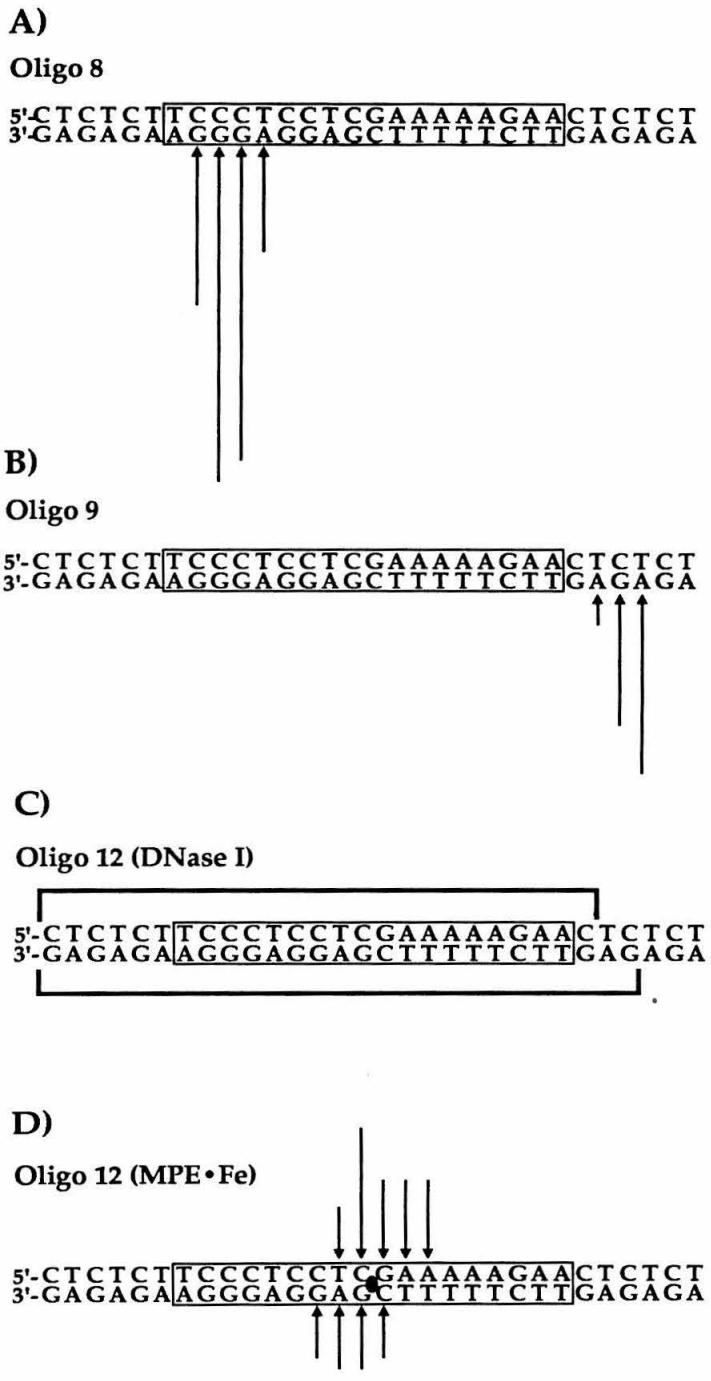
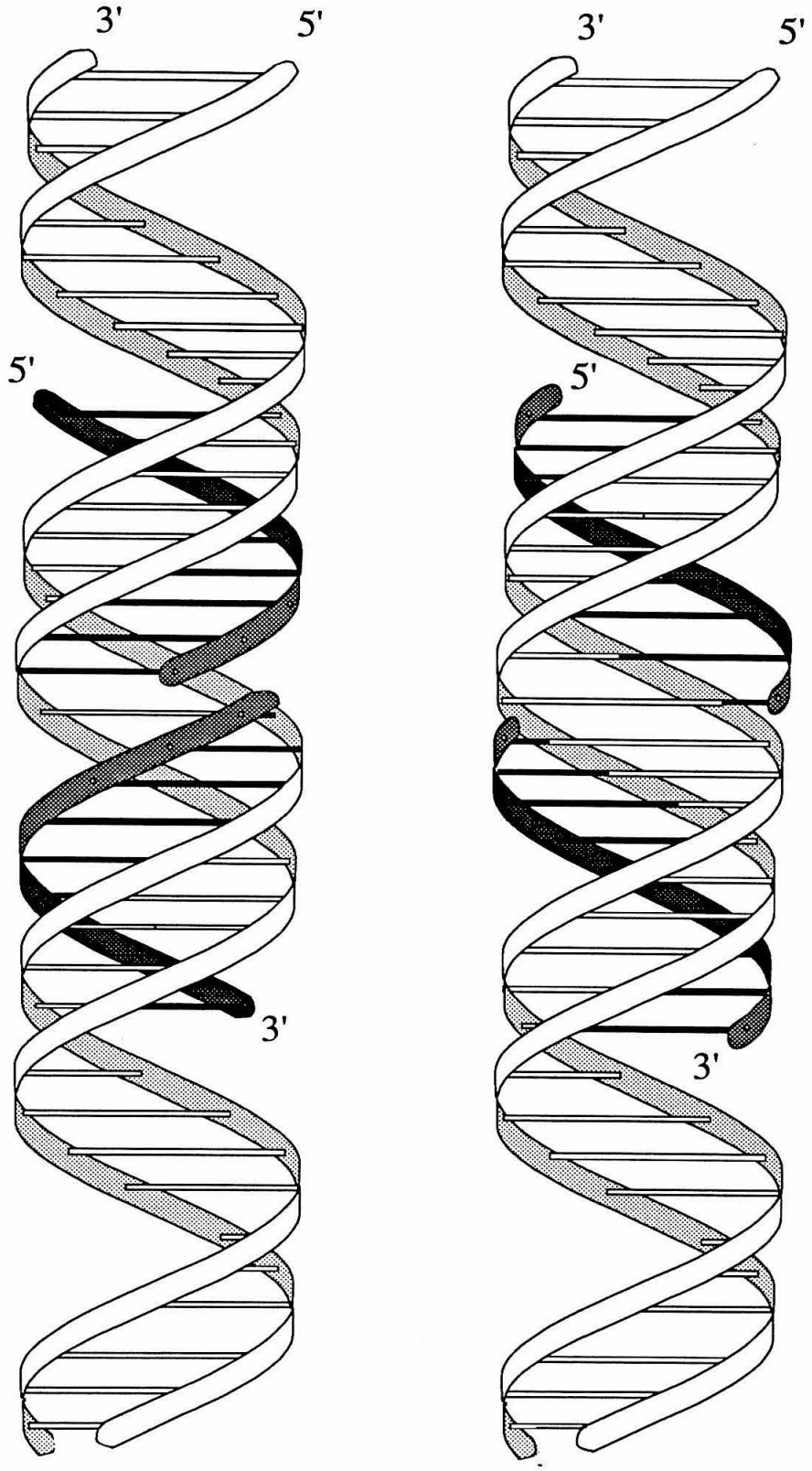

A)
Oligo 8
5'-CTCTCTTCCCTCCTCGAAAAAGAACTCTCT
3'-GAGAGAAGGGAGGAGCTTTTTCTTGAGAGA
B)
Oligo 9
5'-CTCTCTTCCCTCCTCGAAAAAGAACTCTCT
3'-GAGAGAAGGGAGGAGCTTTTTCTTGAGAGA
C)
Oligo 12 (DNase I)
5'-CTCTCTTCCCTCCTCGAAAAAGAACTCTCT
3'-GAGAGAAGGGAGGAGCTTTTTCTTGAGAGA
D)
Oligo 12 (MPE•Fe)
5'-CTCTCTTCCCTCCTCGAAAAAGAACTCTCT
3'-GAGAGAAGGGAGGAGCTTTTTCTTGAGAGA

Figure 3.15 **(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.


3'
5'
5'
3'
3'
5'
5'
3'

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'-$(Py)_m(Pu)_n$-3', the adjacent binding sites do not overlap (Figure 3.15, 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[12] at the center of this alternate strand triple helical structure, creating a site for intercalation.

Multiple crossovers in the major groove. To determine if an 18 bp binding site comprised of three consecutive 6 bp purine tracts on alternate strands could support an alternate strand triple helix, two additional duplex sequences of the type 5′-$(Pu)_l(Py)_m(Pu)_n$-3′ and 5′-$(Py)_l(Pu)_m(Py)_n$-3′ were prepared. These sequences are examples of potential binding sites containing both the 5′-PuPy-3′ and 5′-PyPu-3′junctions. For oligonucleotides to bind to these sequences by alternate strand triple helix formation, they must bind a 6 bp purine tract via base triplets, cross the major groove and bind the adjacent purine tract, then cross back to bind to the third purine tract. Oligonucleotides were designed to bind to these duplex sequences according to the guidelines for the necessary linker structures at the junctions discussed in the previous sections. These oligonucleotides were comprised of three binding domains of six nucleotides capable of forming base triplets with a two thymidine linker to cross the 5′-PyPu-3′ junction and a 3′-5′ phosphodiester to cross the 5′-PuPy-3′ junction.

The 5′-(Pu)$_6$(Py)$_6$(Pu)$_6$-3′ sequence. The potential binding site 5′-AGAAAGCTCCCCGAAAGA-3′ consists of three consecutive purine tracts on alternate strands with 5′-GC-3′ and 5′-CG-3′ junctions. Oligonucleotides **13-15** were designed to bind this site by the formation of pyrimidine motif triple helices at the 5′ and 3′ segments and a purine motif triple helix at the center segment (Figure 3.16). At pH =7.0 and 25°C, no specific cleavage could be detected for these oligonucleotides, indicating that alternate strand triple helix formation at this site is much less facile than at the single crossover sequences discussed in the previous sections. In fact, only weak cleavage could be detected for **14** and **15** at acidic pH's with no apparent difference in stability.

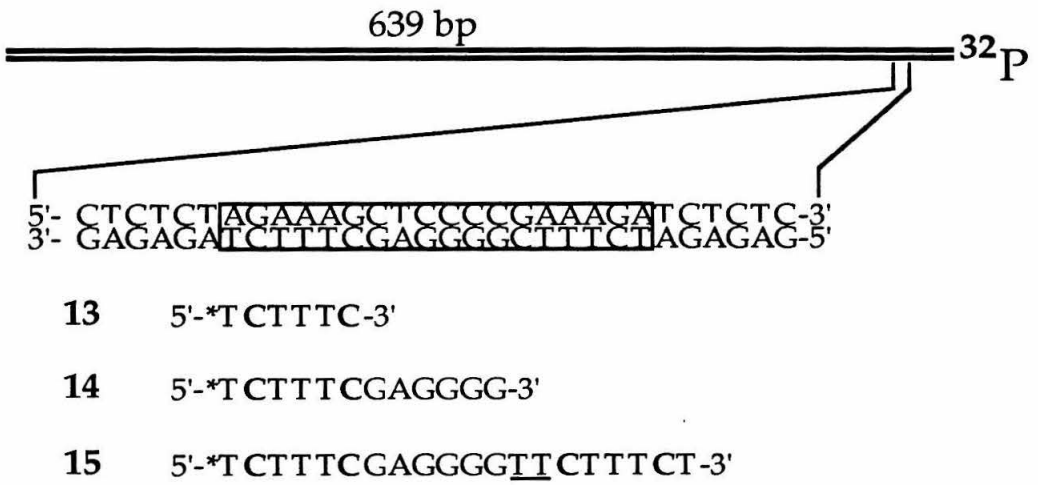
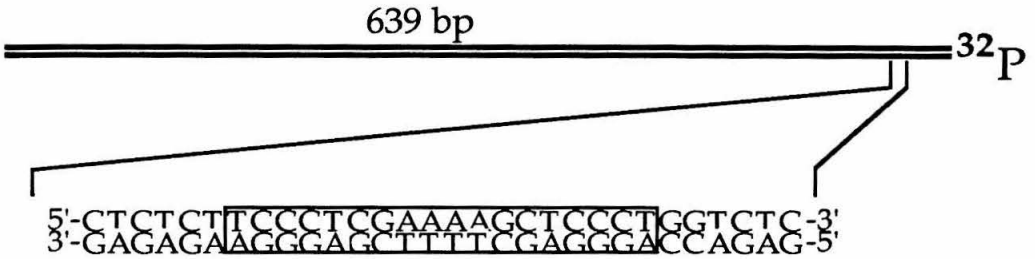


Figure 3.16 The duplex target site present on the 639 bp Hind III/Ssp I restriction fragment from plasmid pPBCRIII containing three six-mer purine tracts on alternate strands joined at 5'-PuPy-3' and 5′-PyPu-3′ junctions. The double stranded target sequence region is boxed. The sequences of oligonucleotide-EDTA's **13-15** are shown, where T* indicates the position of thymidine-EDTA, bold type **C** indicates 5-methyldeoxycytidine and the underlined thymidines indicate the linker domain.

The 5′-(Py)$_6$(Pu)$_6$(Py)$_6$-3′ sequence. The potential binding site 5′-TCCCTCGAAAAGCTCCCT-3′ consists of consecutive purine tracts on alternate strands with 5′-CG-3′ and 5′-GC-3′ junctions. Oligonucleotide **16** was designed to bind this site by the formation of purine motif triple helices at the 5′ and 3′ segments and a pyrimidine motif triple helix at the center segment (Figure 3.17). However, very little cleavage could be detected for **16** under all conditions tested.



16 5'-*TGGGAGTTCTTTTCGAGGGT-3'

Figure 3.17 The duplex target site present on the 639 bp Hind III/Ssp I restriction fragment from plasmid pPBCRIV containing three six-mer purine tracts on alternate strands joined at 5'-PyPu-3' and 5′-PuPy-3′ junctions. The double stranded target sequence region is boxed. The sequence of oligonucleotide-EDTA **16** is shown, where T* indicates the position of thymidine-EDTA, bold type **C** indicates 5-methyldeoxycytidine and the underlined thymidines indicate the linker domain.

These unsuccessful attempts to target 18 bp sequences by alternate strand triple helix formation with two crossovers in the major groove point out a limitation to this approach to binding mixed sequences of double helical DNA. For successful alternate strand triple helix formation, the minimum length of purine tracts appears to be between 6 and 9 bp. The exact length will likely be dependent on sequence and structure of the linker. Optimization of the linker structure, in particular the design of linkers which allow for

continuous π stacking through the junctions, would likely confer greater stability to the alternate strand triple helices.

Conclusions. 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′-5′ phosphodiester linkages which bind mixed sequences of duplex DNA by alternate strand triple helix formation. The duplex sequence type 5'-$(Pu)_m(Py)_n$-3' was targeted by linking two binding domains with a phosphodiester. However, to bind the duplex sequence type 5'-$(Py)_m(Pu)_n$-3', the third strand required two linker nucleotides at the site of crossover in the major groove. The design of oligonucleotides containing both pyrimidines and purines and all natural phosphodiester linkages capable of binding a variety of sequence types by crossovers in the major groove may now be possible.

Experimental Section

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 maintenance of cell lines were accomplished using standard procedures.[14] Calf thymus DNA was purchased from Pharmacia. Polyacrylamide gel electrophoresis was performed in 1XTBE buffer.[14] 5′ and 3′ end labelling was accomplished using standard procedures.[14] 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 b-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 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'-CTAGAGAGAGAGGGAGGAGCTTTTTCTTAGAGAG-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'-AATTCTCTCTTCCCTCCTCGAAAAAGAACTCTCT-3' and 5'-CTAGAGAGAGTTCTTTTTCGAGGAGGGAAGAGAG-3'. The plasmids pPBCRIII and pPBCRIV were prepared from insertion of the duplexes formed between 5'-AATTCTCTCTAGAAAGCTCCCCGAAAGATCTCTC-3' with 5'-CTAGAGAGAGTCTTTCGGGGAGCTTTCTAGAGAG-3' and 5'-AATTCTCTCTTCCCTCGAAAGCTCCCTTCTCTC-3' with 5'-CTAGAGAGAGAGGGAGCTTTTCGAGGGAAGAGAG-3', respectively. 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.[14] 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 Labeled 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.[15] 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. For DNase footprint titrations, the apparent site saturation was determined as described by Ackers[7] using the following equation:

$$\Theta_{app} = 1 - \frac{I_{tot}/I_{ref}}{I_{tot}^{\circ}/I_{ref}^{\circ}} \qquad (1)$$

where I_{tot} and I_{ref} are the integrated volumes at the target site and reference sites, respectively, and I_{tot}° and I_{ref}° correspond to those values for a control lane with no oligonucleotide added. The data were fit to a binding isotherm by minimizing the difference between Θ_{app} and Θ_{fit}:

$$\Theta_{fit} = \Theta_{min} + (\Theta_{max} - \Theta_{min}) \frac{K_a[L]}{1 + K_a[L]} \qquad (2)$$

where [L] corresponds to the total oligonucleotide concentration, K_a corresponds to the apparent association constant, and Θ_{min} and Θ_{max} represent the experimentally determined site saturation values when the site

is unoccupied or saturated, respectively. Data were normalized using the following equation:

$$\Theta_{norm} = \frac{(\Theta_{app} - \Theta_{min})}{(\Theta_{max} - \Theta_{min})} \qquad (3)$$

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.[16] 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 and Notes

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. (d) Jayasena, S.D.; Johnston, B.H.; *Nucleic Acids Res.* **1992**, 20, 5279.

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. Circular permutation experiments[13] were carried out in attempt to determine the extent of bending in the triple helix formed by **12**. However, the results were inconclusive due to the instability of the structure under the conditions necessary for the assay.

13. Zwieb, C.; Kim, J.; Adhya, S.; *Biomethods* **1991**, 5, 245.

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

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

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

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

Chapter 4

Investigations Directed Toward the Sequence-specific Transesterification of RNA

Introduction

Reagents capable of the sequence specific cleavage of nucleic acids are indispensible research tools for molecular biology and show promise as theraputics in the fight against disease.[1,2] Therefore, the design and synthesis of new sequence specific cleaving molecules is currently an active area of research in academics and industry.[3] The generation of these reagents requires the combination of the two functions of sequence specific recognition and cleavage of the sugar-phosphodiester backbone. One approach is to synthesize DNA/RNA binding molecules with reactive functionalities capable of oxidative degradation of the (deoxy)ribose, electrophilic modification of the bases or hydrolysis of the phosphodiester backbone (Figure 4.1). Examples of designed molecules with the ability to cleave DNA

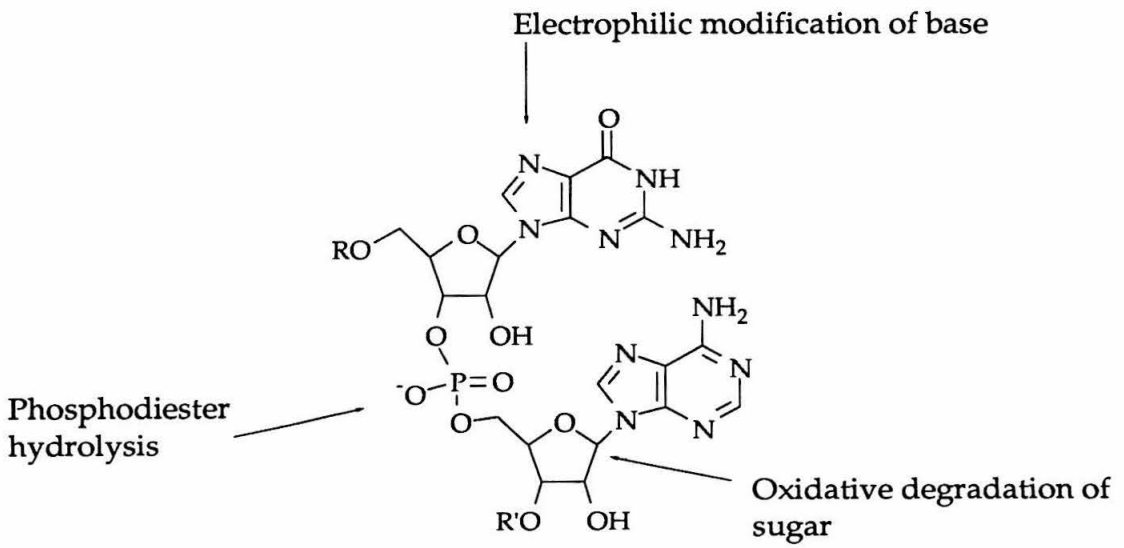


Figure 4.1 Reactions of nucleic acids which can lead to strand scission

at specific sites through oxidative degradation of the deoxyribose are plentiful.[4] These molecules are believed to generate lesions in the strand by a

reaction cascade initiated through hydrogen atom abstraction from the deoxyribose, often by a diffusible oxidant capable of reacting at several nucleotide positions. More specific cleavage has been achieved by tethering electrophilic functional groups to sequence specific DNA binding molecules, which react at nucleophilic sites on the purine bases leading to alkaline labile positions in the DNA strand.[3a-f]

The paucity of reports on designed molecules capable of specific scission of the nucleic acid strand through hydrolysis of the phosphodiester suggests this approach may be more challenging.[5] However, hydrolytic cleavage has several advantages over electrophilic modification or oxidative cleavage. Hydrolysis is not limited to specific base positions, generates termini which are compatible with subsequent enzymatic manipulation and cleaves the phosphodiester backbone without the use of highly reactive oxidants or subsequent base workup. Therefore, a major reseach goal is the development of reagents which promote sequence specific *hydrolytic* cleavage of nucleic acids.

The stability of the phosphodiesters in nucleic acids poses a significant obstacle to the development of these reagents. For instance, the phosphodiester bond in RNA has been estimated to have a $t_{1/2}$ for hydrolysis near 2000 years at 25°C, pH=7.[6] However, this obstacle is overcome by both protein enzymes and ribozymes, which carry out hydrolytic cleavage of RNA in minutes at 37°C (rate accelerations of 10^9-10^{12}).[7] This high level of catalysis is usually achieved at an active site containing specifically bound metal ions. The role of the metal ion is not always known in detail, and several mechanisms by which metal ions might facilitate these reactions have been proposed.[8] These include a) electrophilic activation of the phosphorus, b) charge neutralization of the phosphate, c) delivery of a nucleophile

(coordinated ^{-}OH) at physiological pH, d) delivery of a general acid (coordinated H_2O) or general base (coordinated ^{-}OH), at physiological pH, e) orientation of reacting groups through metal ion coordination, and f) activation of the phosphodiester in a strained chelate.

Certain metal ions and metal complexes have been shown to promote cleavage of RNA phophodiesters through a transesterification reaction resulting in the formation of a 2',3'cyclic phosphodiester (Figure 4.2).[9] The

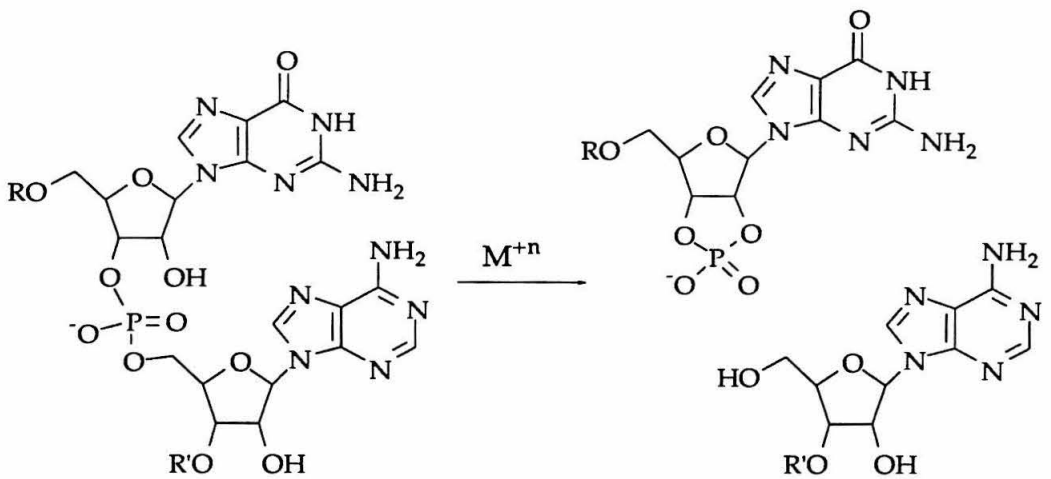
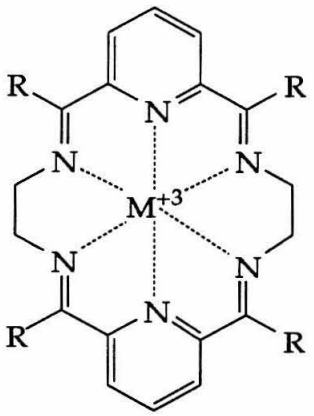


Figure 4.2 Some metal ions can promote the transesterification of RNA leading to a product with a 2',3' cyclic phosphodiester.

lanthanide(III) ions are particularly competent, with thulium(III) accelerating the transesterification of ApA by a factor of 10^7 over background at 30°C, pH=8.[9e] The reactivities of these ions are maintained in complexes with the hexadentate Schiff-base macrocyclic ligand (**L**), which are relatively inert to metal release (Figure 4.3).[9f] Interestingly, the RNA transesterifcation reactions promoted by these ions and complexes are much more facile with phosphodiesters in extrahelical nucleotides (i.e., phosphodiesters in single

stranded regions, loops and bulges) than for phosphodiesters in helical secondary structures.[10] It is believed that this is a result of the conformational rigidity of the helical regions which prevent the phosphodiester and 2' hydroxyl from adopting the proper relative orientation for reaction.

The ability of these complexes to promote RNA transesterifcation, their stability toward metal release and the possibility of attachment to RNA binding molecules through functionalization of the macrocycle lead us to



$L \bullet M^{+3}$

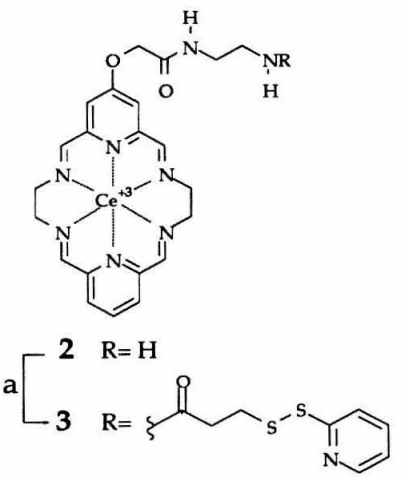
(1): R=H, $M^{+3}=Ce^{+3}$

Figure 4.3 Hexadentate Schiff-base macrocyclic lanthanide(III) complex

investigate methods for their attachment to oligonucleotides and to assess the reactivities of the resulting compounds. It seemed reasonable that the recognition properties of oligonucleotides and the reactivity of these metal complexes might be combined to generate reagents capable of sequence specific recognition of RNA through specific hydrogen bonding and induction of localized strand breaks by lanthanide(III) complex promoted transesterification reactions.

Results and Discussion

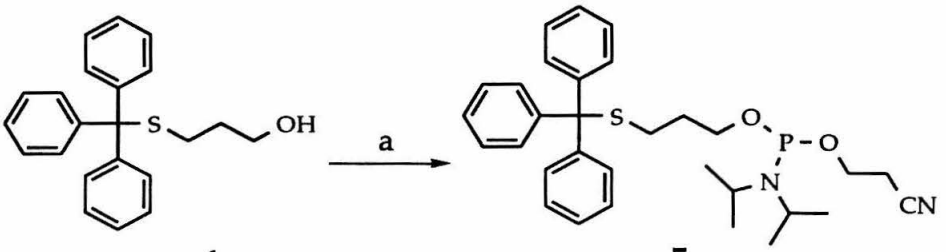
Preparation of modified oligdeoxyribonucleotides. Disulfide bond formation was chosen as a mild and specific method for the covalent attachment of the lanthanide(III) complex to oligonucleotides. For this purpose, oligonucleotides containing free thiol groups and a metal complex with an activated disulfide were prepared. For the generation of the suitably functionalized metal complex, cerium(III) complex (**2**) was prepared via the eight step synthesis of Carreira and Dervan.[11] Further elaboration of the primary amine by acylation with 3-(2-pyridyldithio)propionic acid *N*-hydroxysuccinimide ester gave complex (**3**) functionalized through an 11 atom linker to an activated disulfide (Scheme I). Thiol containing oligodeoxyribonucleotides were generated by automated synthesis using



Scheme I a) 3-(2-pyridyldithio)propionic acid *N*-hydroxysuccinimide ester, TEAA, CH_3CN.

standard solid phase methods with a protected thiol phosphoramidite used in the last coupling step. The synthesis of this phosphoramidite is shown in

Scheme II. S-Trityl-3-mercaptopropanol (**4**), prepared in a single step from 3-bromopropanol and tritylmercaptan, was activated by reaction with 2-cyanoethyl *N*,*N*-diisopropylphosphoramidous chloride giving the S-trityl protected phosphoramidite (**5**) in 85% yield. Oligonucleotides prepared with this phosphoramidite were deprotected at the base and phosphodiester



4 **5**

Scheme II a) 2-cyanoethyl *N*,*N*-diisopropylphosphoramidous chloride, Hunig's base, DCM.

positions with concentrated NH_4OH and purified by reverse phase chromatography. Removal of the thiol protecting group was achieved by treatment with silver nitrate followed by removal of excess silver by complexation with dithiothreitol (DTT), as described by Connolly and Rider.[12]

Disulfide exchange reactions between thiol containing oligodeoxyribonucleotides and (**3**) were followed by reverse phase HPLC (Figures 4.4 and 4.5). For instance, S-trityl protected oligonucleotide (**6**) was treated with silver nitrate followed by DTT generating a new oligonucleotide (**7**) with a retention time shorter by 11.6 minutes (Figure 4.5, **A** and **B**). This oligonucleotide was treated with 3.3 equivalents of (**3**) (Figure 4.5, **C**) in an aqueous solution of triethylammonium acetate, pH=7.0. The HPLC chromatogram of the reaction mixture after one hour contained three peaks, corresponding to the pyridine-2-thione byproduct of the disulfide

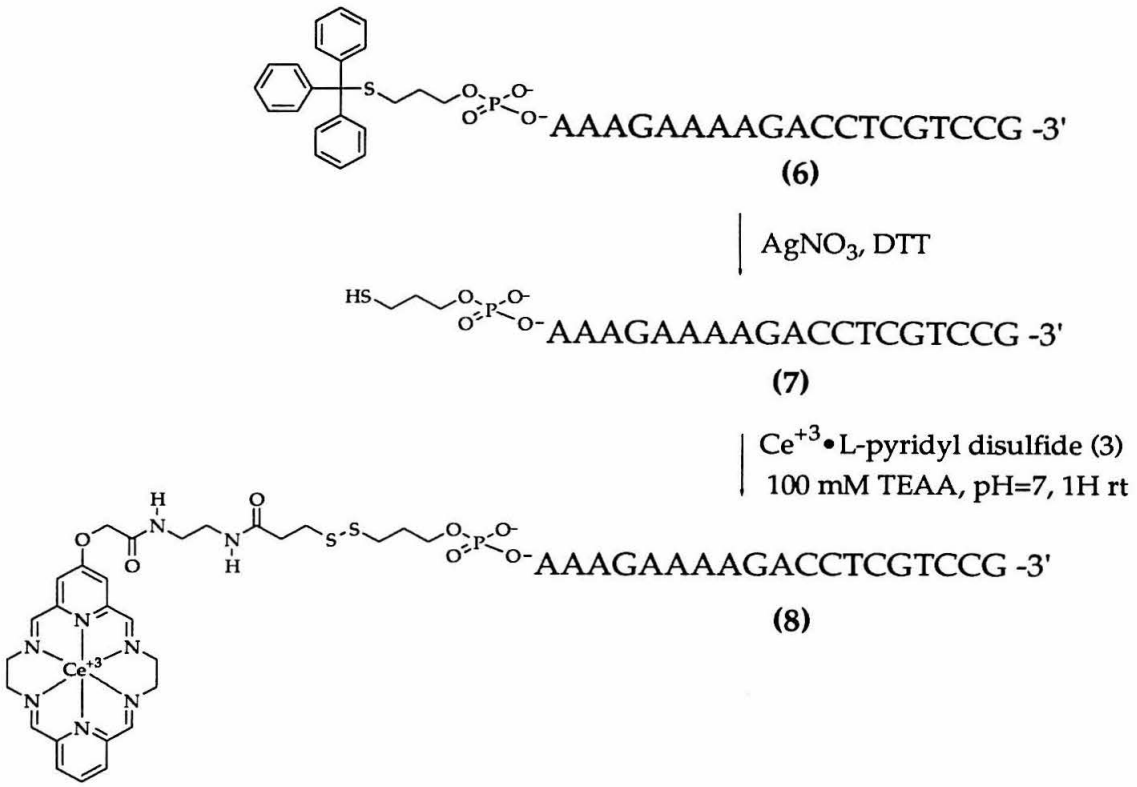
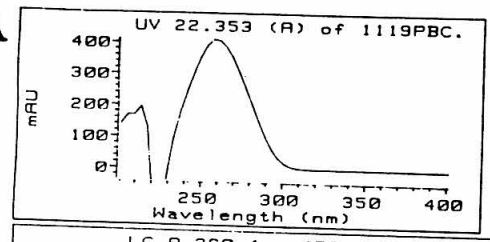
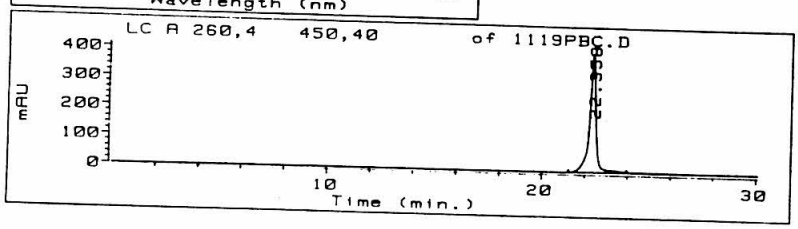


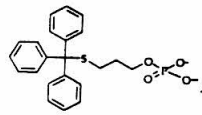
Figure 4.4 Generation of modified oligodeoxyribonucletides via disulfide exchange reactions between thiol containing oligonucleotides and (**3**).

Figure 4.5 HPLC analysis of reaction components in the generation of modified oligodeoxyribonucleotide (**8**). HPLC analyses were carried out as described in the experimental section. **A**) S-Trityl oligodeoxyribonucleotide (**6**) **B**) Thiol containing oligodeoxyribonucleotide (**7**) obtained by treatment of (**6**) with $AgNO_3$ followed by DTT **C**) Activated cerium(III) complex (**3**) **D**) Products of the reaction of (**7**) and (**3**) after 1H at room temperature **E**) Coinjection of (**7**) with the (**7**) + (**3**) reaction products **F**) Modified oligodeoxyribonucleotide (**8**) purified by ethanol precipitation.

A



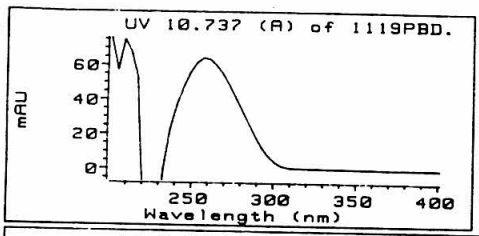


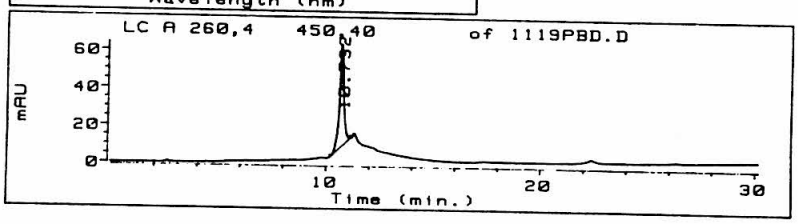


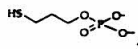
AAAGAAAAGACCTCGTCCG -3'

6

B



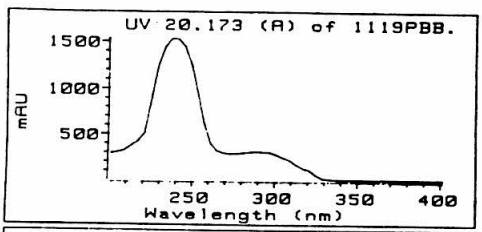


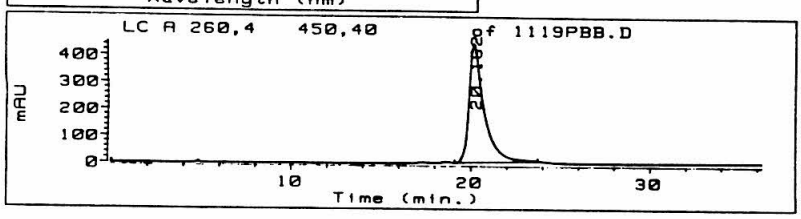


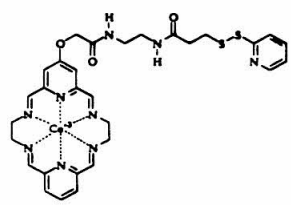
AAAGAAAAGACCTCGTCCG -3'

7

C

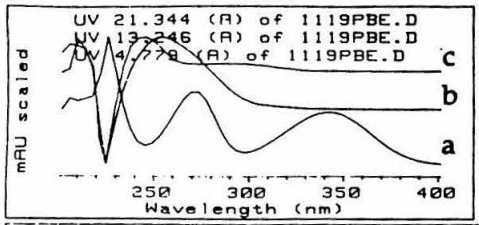


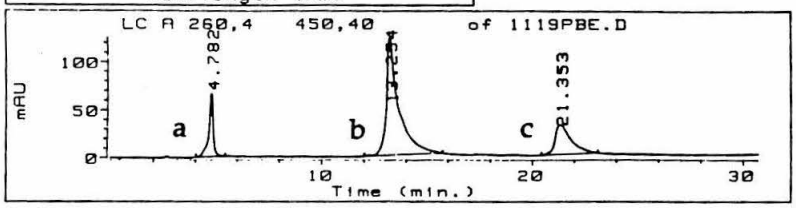




3

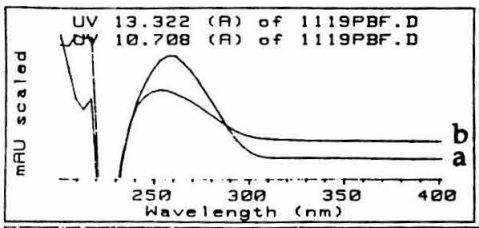
D

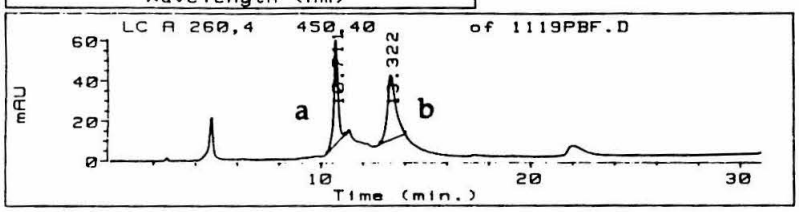




7 + 3 $\xrightarrow[\text{TEAA, pH=7}]{\text{1H, rt}}$

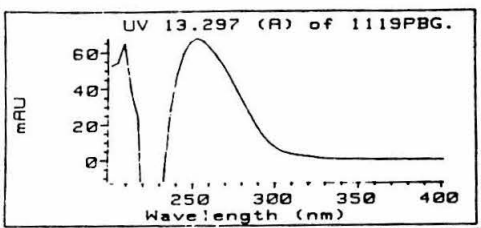
E

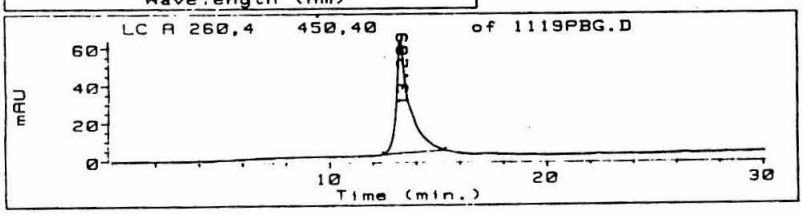


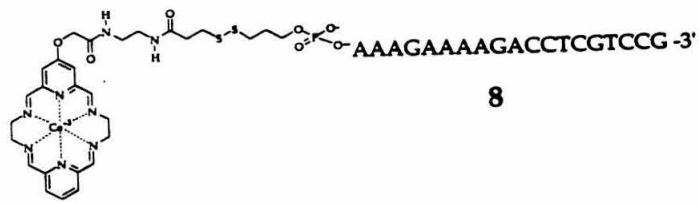


Coinjection of reaction mixture with 7.

F







8

exchange reaction, a new oligonucleotide, and unreacted (**3**) (Peaks a, b and c respectively in Figure 4.5, **D**). Coinjection of this reaction mixture with (**7**) shows conclusively that a new oligonucleotide was formed in the reaction with a retention time 2.5 min. longer than (**7**) (Figure 4.5, **E**, peaks a and b). Precipitation from the reaction mixture with ethanol and sodium acetate gave (**8**) as a single compound by HPLC analysis.

Analysis of the transesterification potential of modified oligonucleotides. The ability of the 19 nt oligodeoxyribonucleotide (**8**) to promote transesterification reactions in a 5'-^{32}P labeled 37 nt oligoribonucleotide was assessed by gel electrophoresis. Formation of a hybrid duplex between this oligoribonucleotide and (**8**) would place a region of single stranded RNA near the 5' end of (**8**) (Figure 4.6). However, under a variety of salt conditions and temperatures, no reproducible cleavage of the oligoribonucleotide above background could be detected (Figure 4.7).

5'-R-AAAGAAAAGACCTCGTCCG-3'

3'-cagcgagguguuuucuuuuuucuuuucuggagcaggc-5' ^{32}P

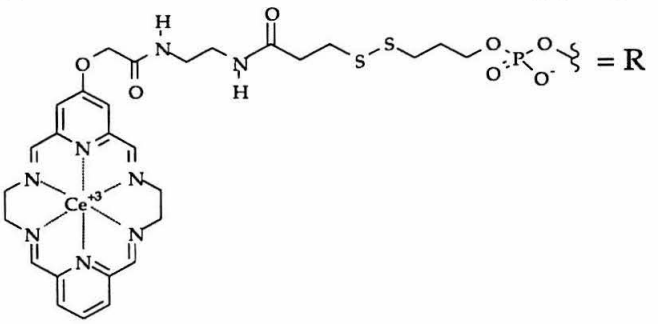
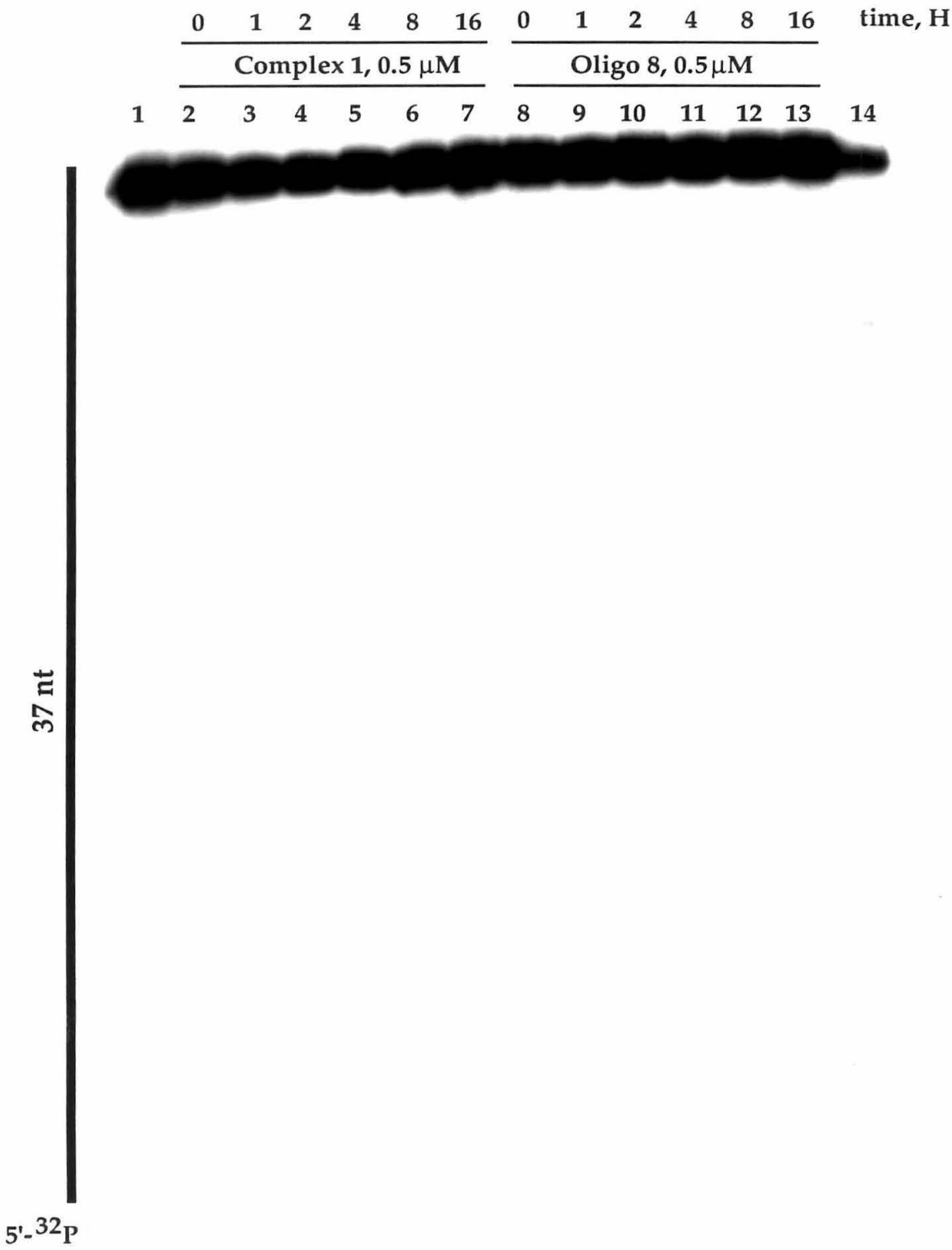


Figure 4.6 Sequences of modified oligodeoxyribonucleotide (**8**) and its RNA target (lower case letters indicates RNA).

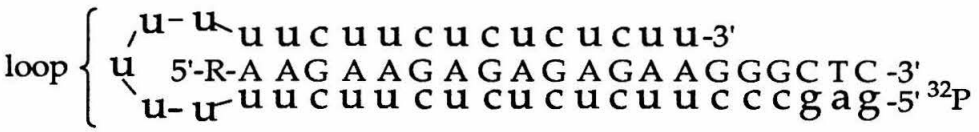
To determine if different sequences or configurations of RNA and modified DNA might lead to specific transesterification, the modified oligodeoxyribonucleotides and target oligoribonucleotides in Figure 4.8 were

Figure 4.7 Autoradiogram of the 20% denaturing polyacrylamide gel used to analyze transesterification reactions. The reactions were carried out by combining ^{32}P labeled RNA (50 nM, $2x10^6$ cpm) in a solution of 50 mM Hepes, pH=7.0 and 10 mM NaCl with either cerium(III) complex (**1**) (lanes 2-7) or modified oligodeoxyribonucleotide (**8**) (lanes 8-13) in a total volume of 200 μL and incubating at 37°C. At various times, 20 μL of the reaction mixture was removed and the reaction was stopped by ethanol precipitation and the products were analyzed by gel electrophoresis. (Lane 1) Untreated labeled 37 nt oligoribonucleotide (Lanes 2-7) Complex (**1**) at 0.5 μM (Lanes 8-13) Modified oligodeoxyribonucleotide (**8**) at 0.5 μM (Lanes 2 and 8) O hours (Lanes 3 and 9) 1 hour (Lanes 4 and 10) 2 hours (Lanes 5 and 11) 4 hours (Lanes 6 and 12) 8 hours (Lanes 7 and 13) 16 hours (Lane 14) Products of a reaction with the G specific ribonuclease T_1.


0 1 2 4 8 16 0 1 2 4 8 16 time, H
Complex 1, 0.5 μM
Oligo 8, 0.5 μM
1 2 3 4 5 6 7 8 9 10 11 12 13 14
37 nt
5'-32P

A

B



C

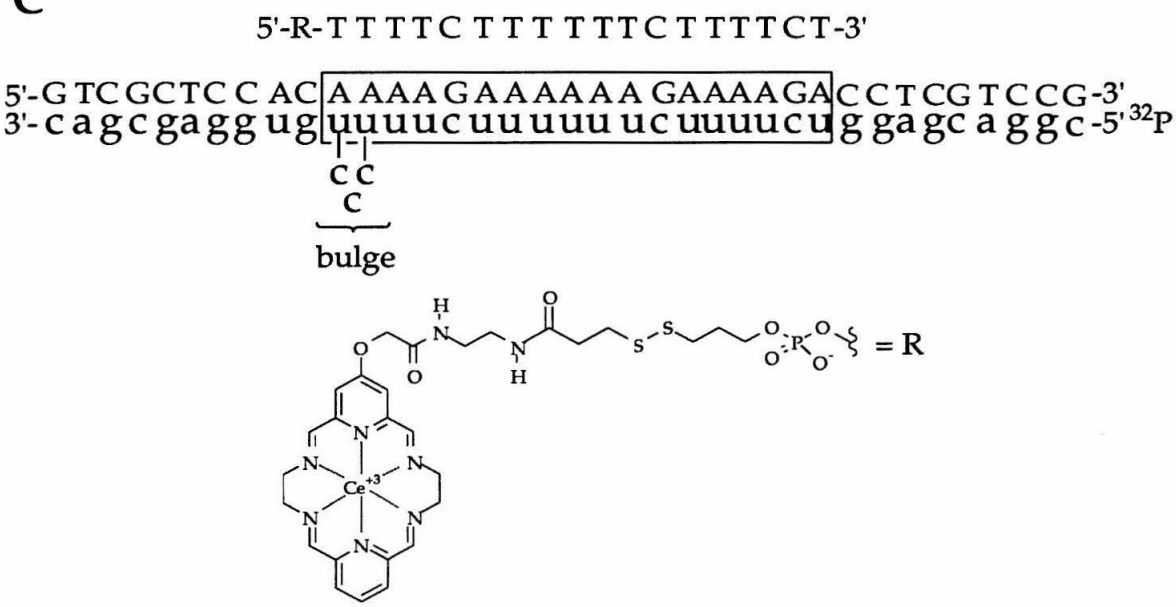


Figure 4.8 Double and triple helical structures prepared to assess the transesterfication potential of modified oligodeoxyribonucleotides. (RNA is designated by lower case letters.)

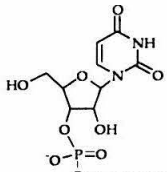
prepared. These oligonucleotides were designed to form hybrid secondary structures placing reactive nucleotide positions in the RNA near the modified 5' ends of the oligodeoxyribonucleotides. The hybrid duplex in Figure 4.8 A contains a different duplex sequence than that discussed in the previous section. The triple helix shown in Figure 4.8 B was designed to place the attached metal complex near loop nucleotide positions in the RNA. The triple helix in Figure 4.8 C was designed to place the metal complex near bulged RNA nucleotides. Conjugation of the macrocyclic cerium(III) complex to the oligodeoxyribonucleotides was carried out via disulfide exchange reactions with thiol containing oligonucleotides and (**3**) as described in the previous section. However, upon incubation under a variety of conditions, no reproducible cleavage above background could be detected for these structures, even though gel shift analyses and footprinting experiments showed that the structures in Figure 4.8 A and B were forming.

The lack of detailed mechanistic studies on lanthanide ion/complex mediated transesterification reactions made the interpretation of these unexpected negative results difficult. For instance, the specific role of the metal ion in catalysis is not known. However, it has been proposed that these ions facilitate reaction by electrophilic catalysis through specific coordination to the phophodiester oxygens.[9d] This mechanism is supported by an X-ray structural analysis of crystals containing complex (**1**) bound to dimethylphosphate, which showed each of the nonbridging oxygens of dimethylphosphate coordinated to a metal complex.[13] Given this information, it seemed plausible that no transesterification was seen because the metal complex was intramolecularly coordinated to DNA phosphodiesters in the modified oligodeoxyribonucleotide, which might lock the complex in a nonproductive binding mode.

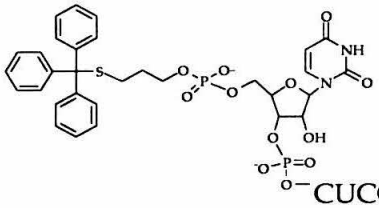
This hypothesis suggested that a self cleavage reaction might be detected if the metal complex were covalently attached to a labeled oligoribonucleotide and intramolecular coordination were to occur. To test this idea and to determine if the conjugation procedure could be used to generate modified RNA's, oligoribonucleotides (**9-12**) were prepared. Oligoribonucleotide (**10**) was prepared by automated RNA synthesis using standard procedures with phosphoramidite (**5**) employed in the last coupling step. Deprotection of the RNA was carried out in ammonia saturated methanol followed by tetrabutylammonium fluoride in tetrahydrofuran. The 5' S-trityl-oligoribonucleotide (**10**) was purified by polyacrylamide gel electrophoresis. (Oligoribonucleotide (**9**) was prepared as a size marker for electrophoresis analyses). Oligonucleotides (**9**) and (**10**) were 3' end ^{32}P labeled by established procedures.[14]

Disulfide exchange reactions between thiol containing ^{32}P labeled oligoribonucleotides and (**3**) were followed by 20% polyacrylamide gel electrophoresis (Figure 4.9). S-Trityl protected oligoribonucleotide (**10**) was treated with silver nitrate followed by dithiothreitol (DTT) generating a single new oligonucleotide (**11**) which migrated through the gel more rapidly, consistent with a lower molecular weight (Figure 4.9, lanes 2 and 3). This oligonucleotide was treated with the activated cerium(III) complex (**3**) in an aqueous solution of triethylammonium acetate, pH=6.0. After one hour at room temperature, the reaction mixture contained a single new oligonucleotide (**12**) which migrated through the gel more slowly than either (**10**) or (**11**), consistent with the addition of molecular weight and positive charge (Figure 4.9, lane 4). Degradation of the RNA by excess cerium complex

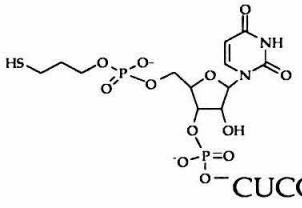
Figure 4.9 (Top) Sequences and structures of oligoribonucleotides **9-12** with the synthesis scheme for the generation of (**12**), modified at its 5′ end with the cerium(III) macrocyclic complex. (Bottom) Storage phosphor autoradiogram of the 20% polyacrylamide gel used to analyze the relative mobilities of **9-12**. (Lane 1) Oligoribonucleotide (**9**) (Lane 2) Oligoribonucleotide (**10**) (Lane 3) Oligoribonucleotide (**11**) (Lane 4) Oligoribonucleotide (**12**).



CUCCCUUCUCUCUp*Cp -3' **9**

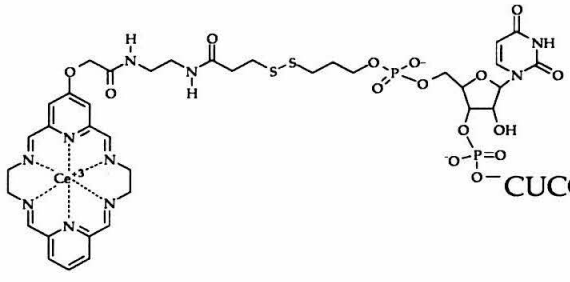


CUCCCUUCUCUCUp*Cp -3' **10**



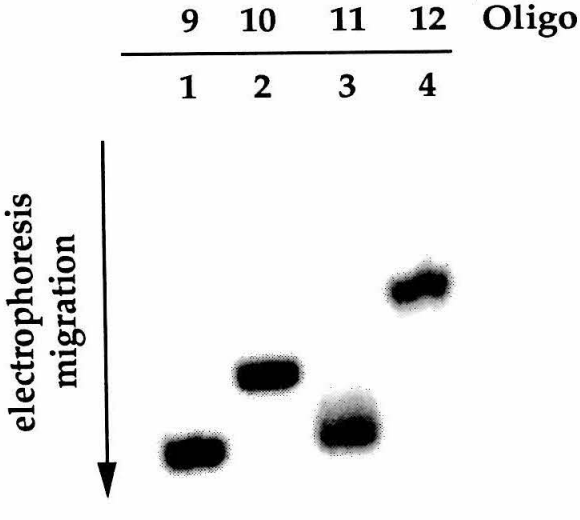
$AgNO_3$, DTT

CUCCCUUCUCUCUp*Cp -3' **11**



Ce^{+3}•L-pyridyl disulfide
100 mM TEAA, pH=6, 1H rt

CUCCCUUCUCUCUp*Cp -3' **12**



(**3**) was minimized by running the reaction at room temperature, pH=6 for only one hour. Precipitation from the reaction mixture with ethanol and sodium acetate gave purified (**12**).

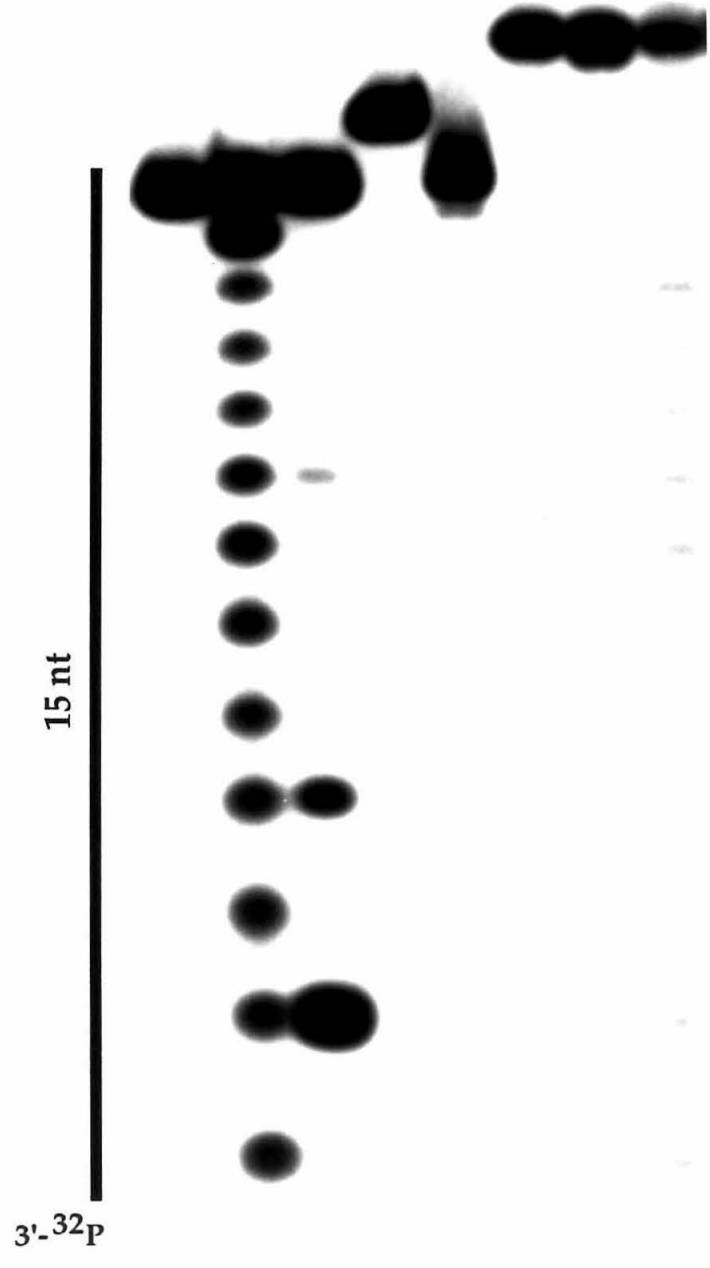
When modified oligoribonucleotide (**12**) was incubated at 37°C for 8 hours, no strand scission could be detected arising from transesterification promoted through intramolecular phosphodiester coordination by the attached metal complex (Figure 4.10, lane 7). When (**12**) was treated with 0.2 mM complex (**1**) under the same conditions, cleavage at all nucleotide positions was detected, indicating that these phosphodiesters were reactive and these solution conditions supported the transesterification reaction. An RNA conjugate prepared with the lanthanum(III) complex gave identical results.[10b]

Modeling of the linker suggested that intramolecular coordination might occur at the 5' phosphodiester linking the oligonucleotide to the tethered complex. This phosphodiester does not have an adjacent 2' OH and therefore cannot undergo the transesterfication reaction. If coordination were to occur here, no cleavage would be detected.

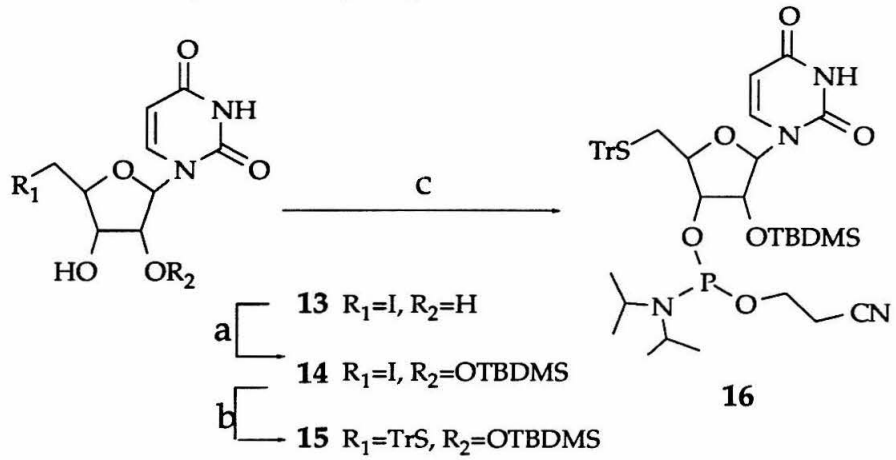
To determine if a different linker structure would direct the transesterification reaction to internal RNA phosphodiesters, a shorter linker without the 5' terminal phosphodiester was designed. For this purpose, a 5'-deoxy-5'-S-trityl-uridine phosphoramidite was prepared. The synthesis was carried out based on the published synthesis for the 5'-deoxy-5'-S-trityl-thymidine phosphoramidite, with an added step for the protection of the uridine 2' OH (Scheme III).[15] Thus, 5'-deoxy-5'-iodouridine was protected at the 2'OH by reaction with *t*-butyldimethysilyl chloride in the presence of silver nitrate as described by Ogilvie for uridine analogs.[16] Displacement of

Figure 4.10 Autoradiogram of the 20% denaturing polyacrylamide gel used to analyze transesterification reactions. The reactions were carried out by incubating ^{32}P labeled (**12**) (7 μM, 9000 cpm) in a solution of 50 mM Hepes, pH=7.5 for 14 H, 37°C in the presence or absence of 0.2 mM (**1**). (Lane 1-3) Oligoribonucleotide **9** (Lane 2) Alkaline hydrolysis products (Lane 3) Products of a reaction with C specific ribonuclease CL_3 (Lane 4) Oligoribonucleotide (**10**) (Lane 5) Oligoribonucleotide (**11**) (Lanes 6-8) Oligoribonucleotide (**12**) (Lane 6) No incubation at 37°C (Lane 7) Incubation under the reaction conditions for 14H (Lane 8) Incubation under the reaction conditions in the presence of 0.2 mM (**1**).

	1	2	3	4	5	6	7	8
Complex 1, 0.2 mM								+
Oligo 12						+	+	+
Oligo 11					+			
Oligo 10				+				
G reaction			+					
alkaline hydrolysis		+						
Oligo 9	+	+	+					



iodide by the sodium salt of trityl mercaptan followed by phosphitylation gave the necessary protected phosphoramidite (**16**).



Scheme III a) TBDMSCl, $AgNO_3$, THF, pyridine; b) TrSNa, DMF; c) 2-cyanoethyl *N*, *N*-diisopropyphosphoramidous chloride, DMAP, THF.

Oligoribonucleotides (**17-20**) were prepared to test the effect of changing the structure of the linker connecting the cerium(III) complex to the RNA. Oligoribonucleotide (**18**) was generated by automated RNA synthesis using standard procedures with phosphoramidite (**16**) employed in the last coupling step. Deprotection and purification of the RNA was carried out as before. Oligoribonucleotide (**17**) was prepared as a size marker for electrophoresis analyses. Oligoribonucleotides (**17**) and (**18**) were 3' end ^{32}P labeled by established procedures.[14] Detritylation and conjugation proceeded as described for (**12**) to generate modified oligoribonucleotide (**20**) (Figure 4.11).

Again, upon incubation of (**20**) at 37°C, no cleavage corresponding to transesterfication could be detected. An RNA conjugate prepared with the lanthanum(III) complex gave identical results.[10b] Interestingly, even though all the phosphodiesters in the molecule had adjacent 2' OH nucleophiles, no

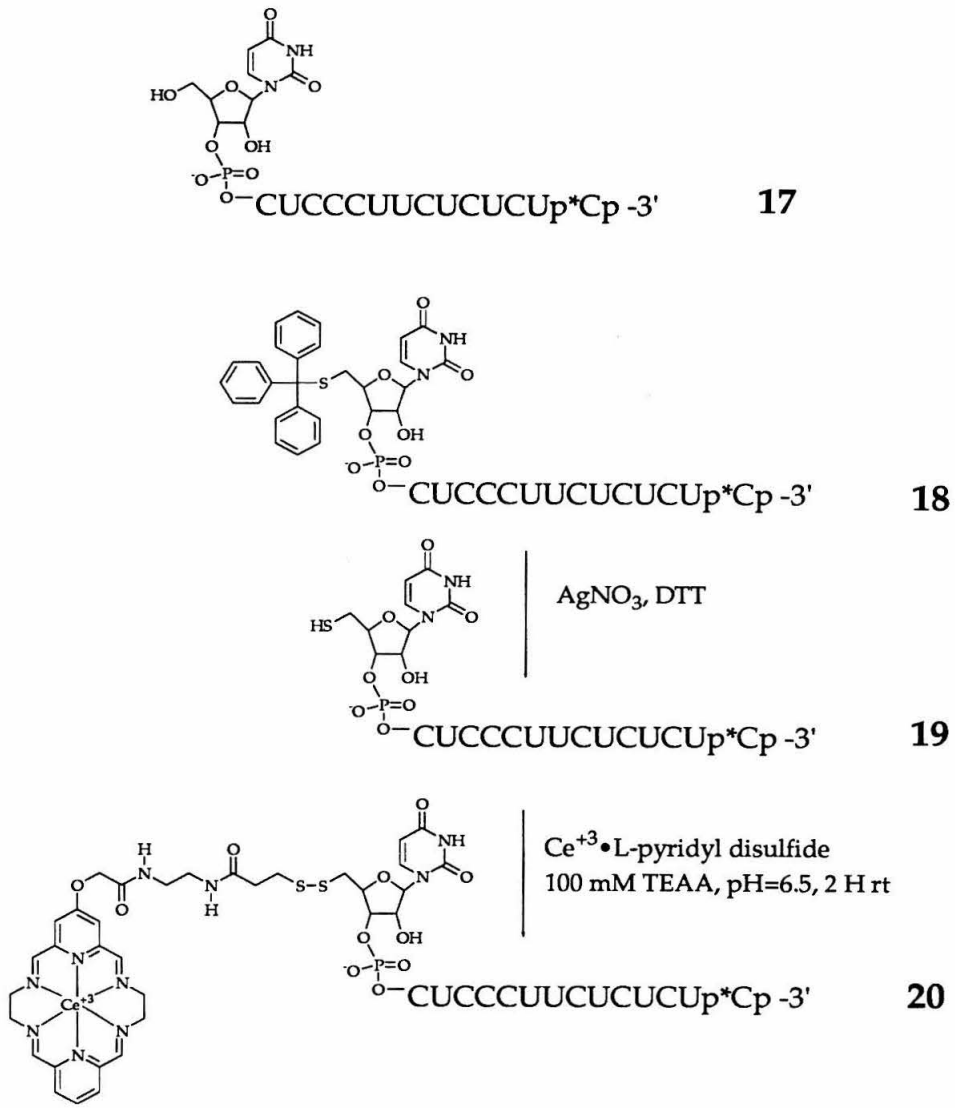


Figure 4.11 Sequences and structures of oligoribonucleotides **17-20** with the reaction scheme for the generation of (**20**) modified at its 5' end with the macrocyclic cerium(III) complex.

self-cleavage was apparent. Therefore, if intramolecular coordination were occurring, it was in a mode which did not lead to transesterfication.

Over a 48 H incubation at 37°C, a slow decomposition reaction to a material which migrated through the gel more rapidly than (**20**) was detected. However, this product was not the result cleavage of one of the internucleotide phosphodiesters (Figure 4.12). The presence of EDTA at 1 mM concentration accelerated the rate of this decompostion reaction by approximately a factor of 2.[17] The gel mobility of the decomposition product and the rate acceleration by EDTA is consistent with demetallation of the cerium(III) complex. This correlates with the observations of Morrow and co-workers that the polycarboxylate chelator EGTA accelerates the decomposition of lanthanide complexes of the type $L \bullet Ln^{+3}$, where L is the macrocylic ligand in Figure 4.4 with R= methyl.[6f]

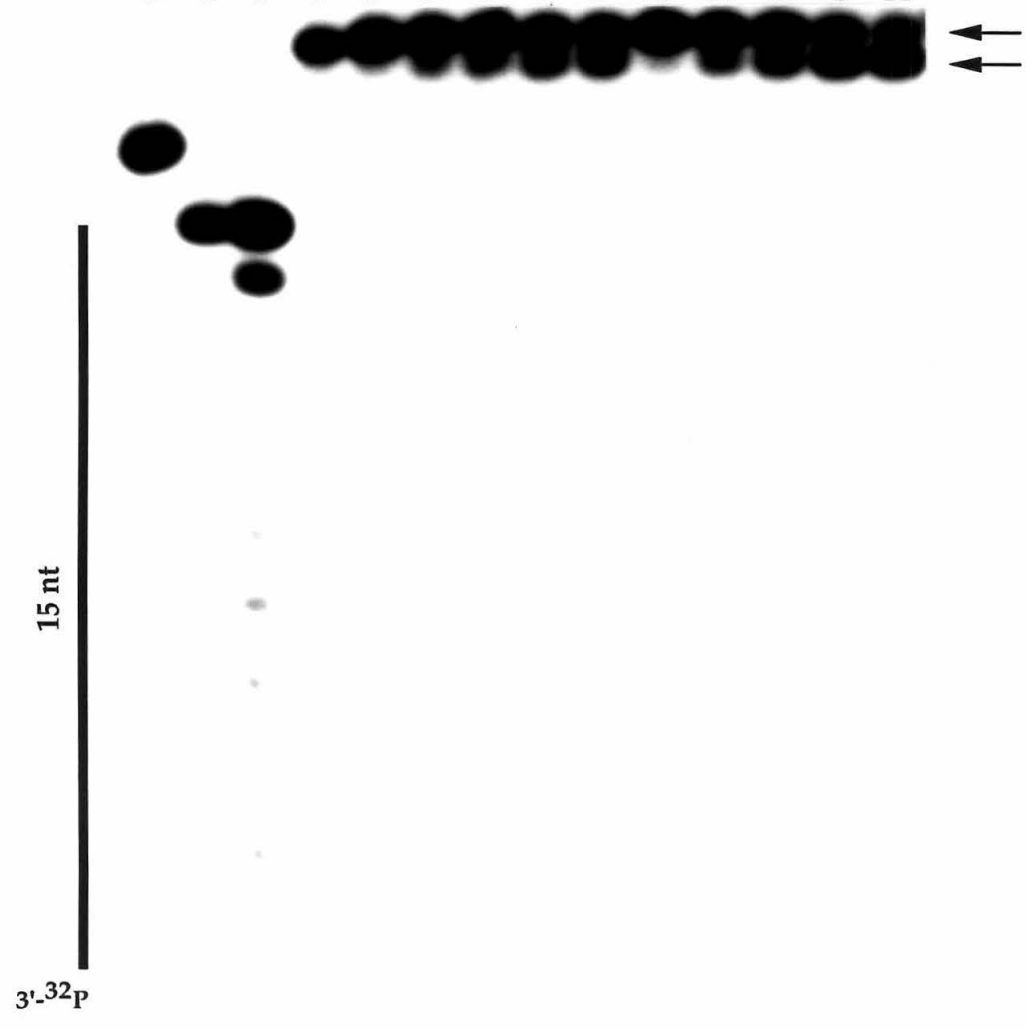
Conclusions and Prospectus

Methods for covalent attachment of the hexadentate Schiff-base macrocyclic cerium(III) complex (**1**) to oligodeoxyribo- and oligoribonucleotides have been developed. However, no success has been made in directing the transesterification reaction promoted by (**1**) using these modified oligonucleotides. Interestingly, RNA binding peptides modified with a hexadentate Schiff-base macrocyclic lanthanum(III) complex have been shown to direct the transesterification reaction of specific phosphodiesters in a bound RNA.[10b] Therefore, it appears that the structure of the RNA binding molecule may be a key factor in determining the success of a directed reaction.

The precise reason for our inability to direct the reaction of (**1**) by covalent attachment to RNA binding oligodeoxyribonucleotides or to the

Figure 4.12 Autoradiogram of the 20% denaturing polyacrylamide gel used to analyze transesterification reactions. The reactions were carried out by incubating ^{32}P labeled (**20**) (0.1 μM, 120 kcpm) in a solution of 10 mM TrisHCl pH=7.2, 70 mM NaCl at 37°C with or without 1 mM EDTA in a total volume of 160 μL. At various times, 30 μL of the reaction mixture was removed and the reaction was stopped by ethanol precipitation and products were analyzed by gel electrophoresis. (Lane 1) Oligoribonucleotide (**18**) (Lanes 2 and 3) Oligoribonucleotide (**17**) (Lane 2) Untreated (**17**) (Lane 3) Alkaline hydrolysis products (Lanes 4-14) Oligoribonucleotide (**20**) (Lanes 10-14) EDTA, 1 mM (Lane 4) Untreated (**20**) (Lanes 5 and 10) 0 hours (Lanes 6 and 11) 8 hours (Lanes 7 and 12) 16 hours (Lanes 8 and 13) 32 hours (Lanes 9 and 14) 48 hours Arrows at right indicate the mobility of oligoribonucleotide **20** (top) and the degradation product (bottom).

time, H					0	8	16	32	48	0	8	16	32	48
EDTA, 1 mM										+	+	+	+	+
Oligo 20				+	+	+	+	+	+	+	+	+	+	+
alkaline hydrolysis			+											
Oligo 17		+	+											
Oligo 18	+													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14



oligoribonucleotide itself is not clear at this time. Studies designed to define the role of the lanthanide ion/complex in the reaction mechanism would aid further experiments. Also, given the crystal structure of (**1**) bound to dimethylphosphate discussed previously, it is possible that more than one metal ion per phosphodiester is required to promote the reaction. Therefore, the development of RNA binding molecules containing more than one metal complex may be necessary.

Experimental Section

General. ^{1}H NMR spectra were recorded on a General Electric-QE 300 NMR spectrometer. High resolution mass spectra were recorded using fast atom bombardment (FAB) at the Mass Spectrometry Facility at the University of California, Riverside. Preparative reverse phase chromatography was performed using a Pharmacia FPLC system with UV detection at 254 nm. E. Merck silica gel 60F (0.25mm) and Whatman KC18F reverse phase plates were used for analytical thin layer chromatographic monitoring of reactions. Flash chromatography was executed with E. Merck silica gel (230-400 mesh). Ultraviolet spectra were obtained with a Hewlett Packard 8452A Diode Array Spectrometer. All reagents and solvents were purchased from Aldrich and used without further purification unless otherwise noted. Distilled, deionized water was used for all aqueous reactions and dilutions. Macrocyclic cerium(III) complex (**1**) was prepared as decribed by Abid and Fenton.[18] The substituted macrocyclic cerium(III) complex (**2**) was prepared via the eight step synthesis of Carriera and Dervan from cheladamic acid.[11] The 3-(2-pyridyldithio)propionic acid *N*-hydroxysuccinimide ester was prepared as

described by Carlsson et al. in two steps from 3-mercaptopropionic acid.[19] S-Trityl-3-mercaptopropanol was prepared via the one step procedure of Connolly and Rider from 3-bromopropan(1)ol.[12] 5'-Deoxy-5'-iodouridine was prepared as described by Verheyden and Moffatt in two steps from 2',3'-O-isopropylideneuridine.[20] Tetrahydrofuran (THF) was distilled under nitrogen from sodium/benzophenone ketyl. Dichloromethane (DCM), triethylamine (TEA), diisopropylethylamine (Hunig's base) and pyridine were distilled under argon from calcium hydride. Dry DMF was purchased from Fluka. 2'-3'-O-Isopropylideneuridine was purchased from Spectrum. DNA phosphoramidites were purchased from Applied Biosystems. RNA phosphoramidites were purchased from BioGenex Laboratories (San Ramon, CA). Polyacrylamide gel electrophoresis was performed using 1XTBE buffer unless otherwise noted. 5'-(γ-^{32}P)ATP (>5000 Ci/mmol) and 5'-(^{32}P-pCp)-3'(3000 Ci/mmol) were obtained from Amersham. 5' and 3' end labeling was executed using standard procedures. NonIdet P-40 detergent was purchased from Sigma. Cerenkov radioactivity was measured on a Beckman LS 2801 scintillation counter.

<u>Hexadentate Shiff-base macrocyclic cerium(III) complex (3).</u> Amine (**2**) (20 mg, 0.025 mmol) was dissolved in 0.163 mL 100 mM triethylammonium acetate, pH = 7.6. To this solution was added a solution of 3-(2-pyridyldithio)propionic acid *N*-hydroxysuccinimide ester (23 mg, 0.075 mmol) in 0.081 mL acetonitrile. The reaction was shaken at room temperature for 40 min. The resulting suspension was diluted into 250 mM ammonium acetate, pH = 4.7, and filtered. The product was purified by reverse phase chromatography using a Pharmacia FPLC system with a C8 ProRPC 10/10 column and a gradient of 0-65% acetonitrile over 40 minutes in 250 mM

ammonium acetate, pH = 4.7. The solvent was removed by lyophilization to give 6 mg of (**2**), 24%. ^{1}H NMR (D_2O, 300 MHz) 14.8 (s, 2H), 14.4 (s, 2H), 10.9 (m, 3H), 10.5 (s, 2H), 7.9 (d, 1H, J= 4.5 Hz), 7.7(d, 1H, J= 7.8 Hz), 7.5 (t, 1H, J= 7.8 Hz) 6.9 (m, 1H), 6.0 (s, 2H), 3.8 (m, 2H), 3.7(m, 2H), 3.1(t, 2H, J= 6.3 Hz), 3.0 (br s, 4H) 2.8(t, 2H, J= 6.3 Hz), 2.6 (br s, 4H), 1.1 (s, 9H). FABMS $[M^{+3}\text{-}2Ac^-]^+$ calcd for $C_{34}H_{39}N_9O_7S_2Ce$: 889.147, found 889.143.

S-Trityl-O-*N,N*-diisopropyl-β-cyanoethylphosphoramidite-3-mercaptopropan (1)ol (5). To a solution of S-Trityl-3-mercaptopropanol (0.36 gm, 1.08 mmol) and TEA (0.263 mL, 1.98 mmol) in 10 mL freshly distilled DCM was added 2-cyanoethyl *N,N*-diisopropylphosphoramidous chloride (0.20 mL, 0.9 mmol via syringe. The solution was stirred at 0° C under argon for 30 min. The reaction mixture was diluted with DCM and washed with saturated $NaHCO_{3(aq)}$. The aqueous fraction was collected and washed with additional DCM. The organic fractions were combined, dried(Na_2SO_4), filtered and concentrated. The product was purified by flash chromatography on silica gel with an 4:1 solution of hexane and diethylether containing 1% TEA. The solvent was removed *in vacuo* to give 0.412 gm of (**3**) as a yellow oil, 85%. ^{1}H NMR ($CDCl_3$, 300 MHz) δ 7.4-7.2 (m, 15H), 3.8 (m, 2H), 3.6 (m, 4H), 2.6 (t, 2H, J= 6.3 Hz), 2.2 (t, 2H, J= 7.4 Hz), 1.6 (m, 2H), 1.1-1.2(m, 12 H) FABMS $[M\text{-}H]^+$ calcd for $C_{31}H_{38}O_2N_2SP$: 533.239, found 533.241.

Synthesis and purification of S-Trityl-oligodeoxyribonucleotide (6). S-Trityl-oligodeoxyribonucleotide (**6**) was prepared on an Applied Biosystems Model 380B DNA synthesizer with β-cyanoethyl phosphoramidites. The protected thiol functionality was incorporated into the oligonucleotide via the coupling of phosphoramidite (**5**) as the last step in automated synthesis. Base labile

protecting groups were removed by treatment of the oligonucleotide with concentrated NH_4OH at 55°C for 24 H. Purification was carried out by reverse phase chromatography using a Pharmacia FPLC ProRPC 10/10 (C2-C8) column with a 0-40% CH_3CN gradient in 100 mM triethylammonium acetate, pH=7. Phosphoramidite (**5**) was estimated to have coupled with an efficiency of >90% based upon the amount of the n-1 byproduct seen during purification. Chromatography solvent was removed by lyophilization. The resulting residue was dissolved in water and lyophilized further. The oligonucleotide was quantitated by determining the concentration in aqueous solution by UV spectroscopy at 260 nm, using the following molar extinction coefficients for each base: 15400 (A), 11700 (G), 7300 (C) 8800 (T) $cm^{-1}M^{-1}$. The solutions were lyophilized and the resulting oligonucleotide was stored as dry solid in 4-5 nmole fractions at -20° C.

Generation of the oligodeoxyribonucleotide-hexadentate Schiff-base macrocyclic cerium(III) complex conjugate (8). To remove the S-trityl protecting group, (**6**) (0.03 µmol in 30 µL 100 mM TEAA, pH=6.5) was treated with a solution of silver nitrate (5 µmol in 5 µL H_2O). This reaction was shaken at room temperature for 30 min. Dithiothreitol (DTT) (6.4 µmol in 6.4 µL H_2O) was added and the reaction was shaken an additional 5 min. at room temperature, at which time the precipitated Ag^+ salt was removed by centrifugation. The precipitate was washed with 60 µL of 100 mM TEAA, pH=7.6 and centrifuged again. The supernatants containing oligonucleotide-thiol (**7**) and excess DTT were combined. HPLC analysis of this solution indicated complete deprotection of the oligonucleotide. DTT was removed by extraction six times with an equal volume of H_2O saturated ethyl acetate. The resulting solution of (**7**) in 100 mM triethylammonium acetate was adjusted

to pH=7.0 with a 0.5 M solution of triethylammonium acetate, pH=7.6. To 90 μL of this solution containing 0.027 μmol (**7**) was added 50 μL of an aqueous solution containing 0.090 μmol of pyridyldisulfide (**3**). The reaction was allowed to proceed for 1 H at room temperature with periodic shaking. HPLC analysis of the reaction solution after 1 H indicated complete conversion of (**7**) to a single new oligodeoxyribonucleotide (**8**). This compound was purified by precipitation twice from an aqueous solution containing 85 mM sodium acetate and 70% ethanol at -20°C for 15 min. The precipitate was washed with 70% ethanol and briefly dried *in vacuo*. This compound was stored as a dry pellet in 5 nmole aliquots at -20°C. Periodic HPLC analysis showed that conjugates prepared by this technique are unchanged for weeks at -20°C. However, incubation under reducing conditions (DTT) will regenerate the thiol containing oligonucleotide starting materials.

HPLC Analysis. Analytical HPLC was performed with a Hewlett-Packard 1090 liquid chromatographer fitted with a diode array detecter using a Brownlee Labs Aquapore OD-300 4.6 mm x 22 cm, 7-μm C18 column. A gradient of 0-65% acetonitrile in 100 mM ammonium acetate, pH=4.7 over 40 min. was employed. Typically 1-5 nmoles of oligonucleotides and 0.05-0.1 μmoles of (**3**) was used for each chromatogram.

Synthesis and purification of oligoribonucleotides. RNA synthesis was carried out on an Applied Biosystems 394 DNA and RNA synthesizer using 2'OTBDMS-β-cyanonethyl phosphoramidites protected at the base with phenoxyacetyl(PAC) for rA and rG and benzyl (Bz) for rC. Base labile protecting groups were removed by treatment with ammonia saturated methanol for 24 H at room temperature. Silyl protecting groups were

removed by treatment with 0.1 M tetrabutylammonium fluoride in THF for 24 H at room temperature. The solvent was removed by lyophilization and the resulting material was redissolved in H_2O and lyophilized further. Excess salts were removed by dialysis of an aqueous solution of this material against 10 mM TrisHCl, pH=7.3 for 48 H followed by an 8 H dialysis against H_2O. The resulting solution was filtered and lyophilized. The oligoribonucleotide was dissolved in formamide TBE loading buffer and loaded onto a 15% polyacrylamide gel for purification by electrophoresis.[1] The gel bands were visualized by UV shadowing and excised. RNA was extracted from the gel slice in an aqueous 0.05% Nonidet P-40 solution at room temperature for 1-4 H and the resulting suspension was filtered. Ethanol precipitation from 85 mM sodium acetate followed by washing with 70% ethanol and drying in vacuo provided the purified oligoribonucleotide. The RNA was quantitated by determining the concentration in aqueous solution by UV spectroscopy at 260 nm, using the following molar extinction coefficients for each base: 15400 (A), 11700 (G), 7300 (C) 9900 (U) $cm^{-1}M^{-1}$. The solutions were lyophilized and the resulting oligonucleotide was stored as dry solid in 2 nmole fractions at -20° C.

Transesterification reactions. In a typical set of transesterfication reactions, labeled RNA was added to a solution of buffer (Tris or Hepes) and salt (NaCl or $NaNO_3$) and this solution was allowed to equilibrate briefly. The hexadentate Shiff-base macrocyclic cerium (III) complex (**1**), modified oligodeoxyribonucleotide or unmodified oligodeoxyribonucleotide at various concentrations was added and the reaction was allowed to incubate at various temperatures for various time intervals. At the end of the incubation period, the RNA was precipitated by the addition of ethanol and sodium acetate and

chilling to -20°C for 15 min. The precipitate was isolated by centrifugation and washed with 70% ethanol and dried briefly *in vacuo*. The resulting pellets were dissolved in formamide TBE loading buffer and the RNA was fractionated by 20% polyacrylamide gel electrophoresis. Gel bands were visualized by autoradiography by exposure to X-ray film (Kodak) at -70° C with an intensifying screen or by phosphor storage autoradiography.

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.[21] A Molecular Dynamics 400S PhosphorImager was used to obtain all data from the storage screens.

Synthesis and purification of S-Trityl-oligoribonucleotide (10). Modified oligoribonucleotide (**10**) was prepared on an Applied Biosystems Model 394 DNA and RNA synthesizer with 2'-OTBDMS-β-cyanoethyl phosphoramidites. The protected thiol functionality was incorporated into the oligonucleotide via the coupling of phosphoramidite (**5**) as the last step in automated synthesis. This RNA oligonucleotide was deprotected and purified on a 20% polyacrylamide gel as described above for synthetic oligoribonucleotides.

Generation of the ^{32}P labeled oligoribonucleotide-hexadentate Schiff-base macrocyclic cerium(III) complex conjugate (12). The S-trityl-oligoribonucleotide (**10**) was 3' end labeled with RNA ligase and 5'-(^{32}pCp)-3' according to published procedures. To remove the S-trityl protecting group, 3' end ^{32}P labeled **10** (4.5 nmol, 210 kcpm in 24 μL H_2O) was treated with a

solution of silver nitrate (5 μmol in 5 μL H_2O). This reaction was shaken at room temperature for 30 min. Dithiothreitol (DTT) (6.4 μmol in 6.4 μL H_2O) was added and the reaction was shaken an additional 5 min. at room temperature, at which time the precipitated Ag^+ salt was removed by centrifugation. The precipitate was washed with 30 μL of 100 mM TEAA, pH=7.6 and centrifuged again. The supernatants containing oligonucleotide-thiol (**11**) and excess DTT were combined. Analysis of the ^{32}P labeled components of this solution by electrophoresis on a 20% polyacrylamide gel indicated complete deprotection of the oligonucleotide. DTT was removed by extraction six times with an equal volume of H_2O saturated ethyl acetate. The resulting solution of (**11**) in 50 mM triethylammonium acetate was adjusted to pH=6.0 with a 0.5 M solution of triethylammonium acetate, pH=7.6. To 36 μL of this solution containing 3.2 nmol, 148 kcpm (**11**) was added 20 μL of an aqueous solution containing 0.036 μmol of pyridyldisulfide (**3**). The reaction was allowed to proceed for 1 H at room temperature with periodic shaking. Electrophoretic analysis of the reaction solution after 1 H indicated complete conversion of (**11**) to a single new slower migrating labeled species (**12**). This compound was purified by precipitation twice from an aqueous solution containing 85 mM sodium acetate and 70% ethanol at -20°C for 15 min. The precipitate was washed with 70% ethanol and briefly dried *in vacuo*. This compound was stored as a frozen aqueous solution at 40 μM, 2,500 cpm/μL at -20°C.

Self cleavage experiments with ^{32}P labeled oligoribonucleotide-hexadentate Schiff-base macrocyclic cerium(III) complex conjugates. In a typical self cleavage experiment, labeled RNA conjugate was removed for storage at -20°C and added to a solution of buffer (Hepes or Tris) and NaCl. The RNA

conjugate was incubated in this solution in the presence and absence of the cerium(III) complex (**1**) at various temperatures for various time intervals. At the end of the incubation period, the RNA was precipitated by the addition of ethanol and sodium acetate and chilling to -20°C for 15 min. The precipitate was isolated by centrifugation and washed with 70% ethanol and dried briefly *in vacuo*. The resulting pellets were dissolved in formamide TB loading buffer (no EDTA) and the RNA was fractionated by 20% polyacrylamide gel electrophoresis in 1XTB buffer (no EDTA). Gel bands were visualized by autoradiography by exposure to X-ray film (Kodak) at -70° C with an intensifying screen or by phosphor storage autoradiography. Specific reaction conditions are given in the figure legends.

5'-Deoxy-5'-iodo-2'-O-tertbutyldimethylsilyl-uridine (14). 5'-Iodouridine (**13**) (1.78 gm, 5 mmol) was dissolved in 50 mL freshly distilled THF and 2.2 mL dry pyridine was added. Silver nitrate (1.7 gm, 10 mmol,) was added and the suspension was stirred until most of the silver nitrate had dissolved. *t*-Butyldimethylsilyl chloride (1.36 gm, 9 mmol) was added and the resulting suspension was stirred at room temperature under argon for 5.5 H. The reaction was filtered and the filtrate was diluted into DCM. This solution was washed with 5% $NaHCO_{3(aq)}$. The fractions were partisioned and the aqueous fraction was further washed with DCM. The organic fractions were combined, dried (Na_2SO_4), filtered and concentrated. The product was purified by flash chromatography on silica gel with a 1:1 solution of ethyl acetate and petroleum ether. The solvent was removed *in vacuo* to give 1.77 gm of (**7**) as a white foam, 75%. ^1H NMR ($CDCl_3$, 300 MHz) δ 8.6 (br s, 1H), 7.6 (d, 1H, J= 8.1 Hz), 5.9-5.8 (m, 2H), 4.4 (m, 1H), 4.0-3.9 (m, 2H), 3.6-3.5 (m, 2H),

0.95 (s, 9H), 0.15 (s, 6H) FABMS MH^+ calcd for $C_{15}H_{26}O_5N_2ISi$ 469.066, found 469.064.

S-Trityl-5'-deoxy-5'-thio-2'-O-tertbutyldimethylsilyl-uridine (15). A 60% dispersion of NaH in mineral oil (136 mg, 3.4 mmol) was suspended in 5 mL dry DMF under argon. The suspension was cooled to 0° C and a 25 mL DMF solution of triphenylmethylmercaptan (1.56 gm, 5.66 mmol) was added. This solution was stirred at room temperature for 10 min., at which time it was cooled on an ice bath and a 20 mL DMF solution of (**14**) (1.77gm, 3.77 mmol) was added. The solution was stirred at 0°C for 30 min. The reaction was diluted into a mixture of 5% $NaHCO_{3(aq)}$ and DCM and extracted. The aqueous fraction was washed with DCM and the organic fractions were combined, dried (Na_2SO_4), filtered and concentrated. The product was purified by flash chromatography on silica gel with a 3:2 solution of hexane and ethyl acetate. The solvent was removed *in vacuo* to give 1.19 gm of (**15**) as a white foam, 51%. ^1H NMR ($CDCl_3$, 300 MHz) δ 8.3 (br s, 1H), 7.5-7.2 (m, 16H), 5.7 (m, 2H), 4.1 (m, 1H), 3.7 (m, 1H), 3.8 (m, 1H), 2.7 (m, 1H), 2.5 (m, 1H), 0.91 (s, 9H), 0.09 (s, 3H), 0.11(s, 3H) FABMS $[M-H]^+$ calcd for $C_{35}H_{39}O_5N_2SSi$: 615.235, found 615.235.

S-Trityl-5'-deoxy-5'-thio-3'-O-(2-cyanoethyl-*N,N*-diisopropylphosphor amidite)-2'-O-tertbutyldimethylsilyl-uridine (16). To a stirred solution of Hunig's base (0.17 mL, 0.972 mmol), 2-cyanoethyl *N,N*-diisopropylphosphoramidous chloride (0.071 mL, 0.316 mmol) and DMAP (6 mg, 0.049 mmol) in 1 mL freshly distilled THF (Na, benzophenone) was added dropwise a solution of (**15**)(0.15 gm, 0.243 mmol) in 1 mL THF. The solution was stirred under argon at room temperature for 3.5 H, at which

time the reaction was complete as indicated by analysis using C18 reverse phase TLC with an 85% acetonitrile in H_2O mobile phase. The reaction was diluted into ethyl acetate and this solution was extacted four times with saturated $NaCl_{(aq)}$. The organic fraction was dried (Na_2SO_4), filtered and concentrated. The product was purified by flash chromatography on silica gel using a solution of 3:2 to 1:1 petroleum ether and ethyl acetate with 1% triethylamine. The solvent was removed *in vacuo* to yield 0.109 gm of (**16**) as a white foam, 55%. ^1H NMR ($CDCl_3$, 300 MHz) δ 8.5 (br s, 1H), 7.4-7.2 (m, 17H), 5.7 (m, 2H), 4.2 (m, 2H), 4.9-3.5 (complex m, 5H), 2.7-2.4(complex m, 4H), 1.2-1.1(m, 13H), 0.9 (s, 9H), 0.8 (m, 6H) FABMS MH^+ calcd. for $C_{43}H_{58}O_6N_4SSiP$: 817.358, found 817.354.

Synthesis and purification of S-Trityl-oligoribonucleotide (18). Modified oligoribonucleotide (**18**) was prepared on an Applied Biosystems Model 394 DNA and RNA synthesizer with 2'-OTBDMS-β-cyanoethyl phosphoramidites. The protected thiol functionality was incorporated into the RNA via the coupling of phosphoramidite (**16**) as the last step in automated synthesis. This RNA oligonucleotide was deprotected and purified on a 20% polyacrylamide gel as described above for synthetic oligoribonucleotides. Phosphoramidite (**16**) was estimated to have coupled with an efficiency of ~70% based upon the amount of the n-1 by product seen during purification.

Generation of the ^{32}P labeled oligoribonucleotide-hexadentate Schiff-base macrocyclic cerium(III) complex conjugate (20). The S-trityl-oligoribonucloetide (**18**) was labeled, detritylated and conjugated to the macrocylic metal complex as described for (**12**) with slight modification. The

conjugation reaction for (**20**) was carried out at pH=6.5 for 2 H. The final disulfide product was purified on a 20% polyacrylamide gel. Isolation from the polyacrylamide gel was carried out as described for the synthesis and purification of ribonucleotides.

References and Notes

1. Sambrook, J.; Fritsch, E.F.; Maniatis, T.; *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989).
2. Sugiyama, H.; Ehrenfeld, G.M.; Shipley, J.B.; Kilkuskie, R.E.; Chang, L.H.; Hecht, S.M.; *J. Natural Products* **1985**, 48, 869.
3. (a) Baker, B.F.; Dervan, P.B.; *J. Am. Chem. Soc.* **1985**, 107, 8266. (b) Povsic, T.J.; Dervan, *J. Am. Chem. Soc.* **1990**, 112, 9428. (c) Fedorova, O.S.; Knorre, D.G.; Podust, L.M.; Zarytova, V.F.; *FEBS Lett.* **1988**, 228, 273. (d) Shaw, J.P.; Milligan, J.F.; Krawcyzk, S.H.; Matteucci, M.D.; *J. Am. Chem. Soc.* **1991**, 113, 7765. (e) Meyer, R.B.; Tabone, J.C.; Hurst, G.D.; Smith, T.M.; Gamper, H.; *J. Am. Chem. Soc.* **1989**, 111, 8517. (f) Kutyavin, I.V.; Gamper, H.B.; Gall, A.A.; Meyer, R.B.; *J. Am. Chem. Soc.* **1993**, 115, 9303. (g) Lee, M.D.; Ellestad, G.A.; Borders, D.B.; *Acc. Chem. Res.* **1991**, 24, 235.
4. (a) Hertzberg, R. P.; Dervan, P.B.; *J. Am. Chem. Soc.* **1982**, 104, 313. (b) Schultz, P. G.; Taylor, J. S.; Dervan, P. B.; *J. Am. Chem. Soc.* **1982**, 104, 6861. (c) Sluka, J.; Bruist, M.; Horvath, S. J.; Simon, M. I.; Dervan, P. B.; *Science* **1987**, 238, 1129. (d) Moser, H. E.; Dervan, P.B.; *Science* **1987**, 238, 645. (e) Chen, C-H.B.; Gorin, M.B.; Sigman, D.; *Proc. Natl. Acad. Sci. USA* **1993**, 90, 4206. (f) Bruice, T.W.; Wise, J.G.; Rosser, D.S.E.; Sigman, D.S.; *J. Am. Chem. Soc.* **1991**, 113, 5446. (g) Pyle, A.M.; Long, E.C.;

Barton, J.K.; *J. Am. Chem. Soc.* **1989**, 111, 4520. (h) Mack, D.P.; Dervan, P.B.; *J. Am. Chem. Soc.* **1990**, 112, 4604.

5. Basile, L.A.; Raphael, A.L.; Barton, J.K.; *J. Am. Chem. Soc.* **1987**, 109, 7550.

6. (a) Chin, J.; Banaszcyzk, M.; Jubian, V.; Zou, X.; *J. Am. Chem. Soc.* **1989**, 111, 186. (b) Chin, J.; *Acc. Chem. Res.* **1991**, 145. (c) Matsumoto, Y.; Komiyama, M.; *J. Chem. Soc. Chem. Comm.* **1990**, 1050.

7. (a) Lesser, D.; Kurpiewski, M.; Jen-Jacobson, L.; *Science* **1990**, 250, 776. (b) Cech, T.; *Science* **1987**, 236, 1532.

8 Hendry, P.; Sargeson, A.M.; *Prog. Inorg. Chem. Bioinorg. Chem.* **1990**, 38, 201.

9. (a) Stern, M.K.; Bashkin, J.K.; Sall, E.D.; *J. Am. Chem. Soc.* **1990**, 112, 5357. (b) Shelton, V.M.; Morrow, J.R.; *Inorg. Chem.* **1991**, 30, 4295. (c) Breslow, R.; Huang, D-L.; *Proc. Natl. Acad. Sci. USA* **1991**, 88, 4080. (d) Morrow, J.R.; Buttrey, L.A.; Berback, K.A.; *Inorg. Chem.* **1992**, 31, 16. (e) Komiyama, M.; Matsumura, K.; Matsumoto, Y.; *J. Chem. Soc. Chem. Comm.* **1992**, 640. (f) Morrow, J.R.; Buttrey, L.A.; Shelton, V.M.; Berback, K.A.; *J. Am. Chem. Soc.* **1992**, 114, 1903.

10. (a) Kolasa, K.; Morrow, J.; Sharma, A.P.; *Inorg. Chem.* **1993**, 32, 3983. (b) Kelly, J.J.; Carreira, E.M.; Dervan, P.B. (in preparation).

11. Carreira, E.M.; Dervan, P.B.; unpublished synthetic methods.

12. Conolly, B.A.; Rider, P.; *Nucleic Acids Res.* **1985**, 13, 4485.

13. Takasaki, B.; Hynes, R.C.; Chin, J.; Carreira, E.; Dervan, P.B.; unpublished data.

14. Uhlenbeck, O.C.; Cameron, V.; *Nucleic Acids Res.* **1977**, 4, 85.

15. Sproat, B.S.; Beijer, B.; Rider, P.; Neuner, P.; *Nucleic Acids Res.* **1987**, 15, 4837.

16. Hakimelahi, G.H.; Proba, Z.A.; Ogilvie, K.K.; *Can. J. Chem.* **1982**, 60, 1106.

17. Plots of the -ln(intact **20**/total labeled RNA) vs. time in the absence and presence of EDTA were linear and yielded apparent reaction half lifes of 69 and 34 hours, respectively.

18. Abid, K.K.; Fenton, D.E.; *Inorg. Chem. Acta* **1985**, 95, 119.

19. Carlsson, J.; Drevin, H.; Axen, R.; *Biochem. J.* **1978**, 173, 723.

20. Verheyden, J.P.H.; Moffatt, J.G.; *J. Org. Chem.* **1970**, 35, 2319.

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