Oligonucleotide-Directed Cleavage of Single- and Double-Stranded DNA by Double and Triple Helix Formation

Thesis by

Linda Chu-Li Griffin

In Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

1990

(Submitted August 1, 1989)

to my parents

and John

Acknowledgments

I would like to thank my research advisor, Peter Dervan, for his guidance, encouragement, and for providing the research opportunities that made this thesis possible. I also want to thank the members of the Dervan group, past and present, for providing a diverse and stimulating environment in which to grow and learn. Special thanks go to Prof. Michael Doyle, Al Sylwester, Laura Gilliom, Geoff Dreyer, Scott Youngquist, Jim Sluka, Dave Mendel, and Heinz Moser for help in the early years. Thanks to Scott Singleton for computer modeling and graphics of triple helical DNA, Tom Povsic for large scale preparation of plasmid HIV-CAT (provided by Prof. Alan Frankel), and Laura Kiesling and Martha Oakley for proofreading this thesis. Financial support from the Allied Corporation and the Department of Education is gratefully acknowledged. Finally, I would like to thank my family for their support and especially John for his love and encouragement.

Abstract

Part I

Oligonucleotide-Directed Cleavage of Single-Stranded DNA by Double Helix Formation

Chapter 1: Sequence-Specific Cleavage of Single-Stranded DNA with Oligonucleotide-EDTA•Fe(II): Study of Reaction Conditions

The machine synthesis of a DNA hybridization probe 19-nucleotides in length, equipped with the metal chelator EDTA **1** has the sequence 5'-GCAAGGCGAT*TAAGTTGGG-3', which is complementary to a 19-nucleotide sequence in M13mp7 (+) strand phage DNA. In the presence of Fe(II), O_2, and dithiothreitol, oligonucleotide **1** affords specific cleavage (0°C, pH 7.4, 20 hours) at its complementary sequence in the 7214 base M13mp7 (+) strand. Cleavage occurs over a range of 17 nucleotides at the site of hybridization of **1**. No other cleavage sites are observed in the 7214 base strand. Optimum cleavage conditions occur at concentrations of 0.25-1 μM T* oligonucleotide, 20 μM Fe(II), and 50-1000 mM NaCl. Optimum pH is 7.4 (25 mM tris-acetate buffer). The optimum cleavage time is between 16-24 hours at 0°C. The melting temperature (T_m) for T* oligonucleotide **1** and its 19-nucleotide complement under reaction conditions is 54.5°C, 3.5°C lower than the same duplex without T* (T_m=58°C).

Abstract

Part II

Oligonucleotide-Directed Cleavage of Double-Stranded DNA by Triple Helix Formation

Chapter 2: Recognition of Thymine•Adenine Base Pairs by Guanine in a Pyrimidine Triple Helix Motif

Oligonucleotide recognition offers a powerful chemical approach for the sequence-specific binding of double helical DNA. In the pyrimidine-Hoogsteen model, a binding site size of >15 homopurine base pairs affords >30 discrete sequence-specific hydrogen bonds to duplex DNA. Because pyrimidine oligonucleotides limit triple helix formation to homopurine tracts, it is desirable to determine whether oligonucleotides can be used to bind all four base pairs of DNA. A general solution would allow targeting of oligonucleotides (or their analogs) to any given sequence in the human genome. A study of 20 base triplets reveals that the triple-helix can be extended from homopurine to mixed sequences. Guanine contained within a pyrimidine oligonucleotide specifically recognizes thymine•adenine base pairs in duplex DNA. Such specificity allows binding at mixed sites in SV40 and HIV DNA.

Abstract

Part II

Oligonucleotide-Directed Cleavage of Double-Stranded DNA by Triple Helix Formation

Chapter 3: Recognition of All Four Base Pairs of Duplex DNA by Triple Helix Formation • Design of Pyrimidine Specific Bases

Oligonucleotide recognition offers a powerful chemical approach for the sequence-specific binding of double helical DNA. In the pyrimidine-Hoogsteen model, a binding site size of >15 homopurine base pairs affords >30 discrete sequence-specific hydrogen bonds to duplex DNA. Because pyrimidine oligonucleotides limit triple helix formation to homopurine tracts, it is desirable to determine whether oligonucleotides can be used to bind all four base pairs of DNA. A general solution would allow targeting of oligonucleotides (or their analogs) to any given sequence in the human genome. The novel base 4-(3-benzamido)phenylimidazole specifically recognizes pyrimidine•purine base pairs over purine•pyrimidine base pairs. Such specificity allows binding at an 18 base pairs site in SV40 DNA (pH 7.4, 40°C) containing all four base pairs.

Table of Contents

List of Figures and Tables

Part I

Chapter 1:

Part II

Overview:

Chapter 2:

Appendix Chapter 2:

Chapter 3:

Appendix Chapter 3:

Part I

Oligonucleotide-Directed Cleavage of Single-Stranded DNA by Double Helix Formation

Chapter 1

Sequence-Specific Cleavage of Single-Stranded DNA with Oligonucleotide-EDTA•Fe(II): Study of Reaction Conditions

Introduction:

Sequence-specific cleavage of DNA by restriction endonucleases has found may applications, such as DNA sequence determinations, chromosome analyses, gene isolation, and recombinant DNA manipulations, but is limited by the specificities and natural availability of restriction endonucleases (Smith, 1979; Modrich, 1982; Roberts, 1983). It has been shown previously that attachment of EDTA•Fe(II) to a DNA binding molecule creates a DNA cleaving molecule (Hertzberg & Dervan, 1982; Schultz *et al.*, 1982; Taylor *et al.*, 1984; Schultz & Dervan, 1983; Schultz & Dervan, 1984; Hertzberg & Dervan, 1984). For example, oligopeptide-EDTA•Fe(II) molecules distamycin-EDTA•Fe(II) and penta-*N*-methylpyrolecarboxamide-EDTA•Fe(II) cleave double-stranded DNA (25°C, pH 7.4) adjacent to specific A+T-rich regions 5 and 7 base pairs in length, respectively (Taylor *et al.*, 1984; Schultz & Dervan, 1983; Schultz & Dervan, 1984; Youngquist & Dervan, 1984). For cleavage of single-stranded DNA it has been shown that attachment of EDTA•Fe(II) to an oligonucleotide also creates a sequence-specific DNA cleaving molecule (Dreyer & Dervan, 1985; Chu & Orgel, 1985; Boutorin *et al.*, 1984; Boidot-Forget *et al.*, 1986). The use of an oligonucleotide as the sequence-specific binding unit gives flexibility with regard to target sequence and specificity limited only by the length of the oligomer (Grineva & Karpova, 1973; Grineva *et al.*, 1977; Martin & Haseltine, 1981; Knorre & Vlassov, 1985; Szybalski, 1985; Podhajska & Sybalski, 1985; Kim *et al.*, 1988). Previous studies of DNA cleavage by oligonucleotide-EDTA have been conducted with short oligonucleotides (Chu & Orgel, 1985), 167-bp restriction fragment (Dreyer & Dervan, 1985), or poly A < 200 bases (Boutorin *et al.*, 1984, Boidot-Forget *et al.*, 1986). The specific cleavage of larger DNA by oligonucleotide-EDTA is now described.

We report the machine synthesis of a DNA hybridization probe 19 nucleotides in length, containing an EDTA-functionalized derivative of thymidine, T*. The T* oligonucleotide synthesized, 5'-GCAAGGCGAT*TAAGTTGGG-3', is complementary to a 19-nucleotide sequence in M13mp7 (+) strand phage DNA. In the presence of Fe(II), O_2, and dithiothreitol, the oligonucleotide-EDTA **1** affords cleavage (0°C, pH 7.4, 20 hours) at its 19-nucleotide complement in the 7214 nucleotide long M13mp7. The reaction conditions giving optimum cleavage efficiency of the DNA by the T* oligonucleotide are reported as well as the effect of T* on the melting temperature of an oligonucleotide duplex.

MATERIAL AND METHODS

Materials. M13mp7 (+) strand DNA was purchased from Pharmacia. All enzymes were from Boehringer Mannheim. Amersham supplied [γ-^{32}P]ATP.

Synthesis of T.* (Fig. 1) T* was prepared as previously described (Dreyer & Dervan, 1985) with the following modifications: **Nucleoside 2a.** After concentrating the crude reaction to dryness the remaining white solid was washed twice with a NaCl solution (100 g/l) to remove NaCl from the mercurinucleoside. The white solid was then washed with methanol and diethyl ether as described (94% yield). **Nucleoside 3.** Argon was bubbled through the reaction mixture after treatment with H_2S. The solution was filtered through Celite, concentrated to dryness, purified by flash chromatography (CH_2Cl_2/toluene/EtOH, 8:1:1, v/v/v; TLC: R_f=0.29) and recrystallized as described (68% yield). **Nucleoside 5.** The reaction was carried out at 4°C. Additional portions of 4,4'-dimethoxytrityl (DMT) chloride (0.05 mole equivalents each) were added after 3 hours until TLC indicated complete reaction. The work-up procedure was as described. Azeotropic removal of pyridine was accomplished by the addition of toluene and reconcentration. Flash chromatography (EtOAc/hexane, 2:1, v/v; TLC: R_f=0.29) provided nucleoside **5** (88% yield over 2 steps). **Nucleoside 7.** Nucleoside **6** was prepared as described. The crude product **6** was coupled with 1.5 mole equivalents of EDTA-triethylester-*N*-

hydroxysuccinimide ester. Flash chromatography (eluting first with CH_2Cl_2/EtOH, 19:1, v/v; and then CH_2Cl_2/toluene/EtOH, 80:10:9, v/v/v; TLC: R_f=0.35) provided T* (57% yield).

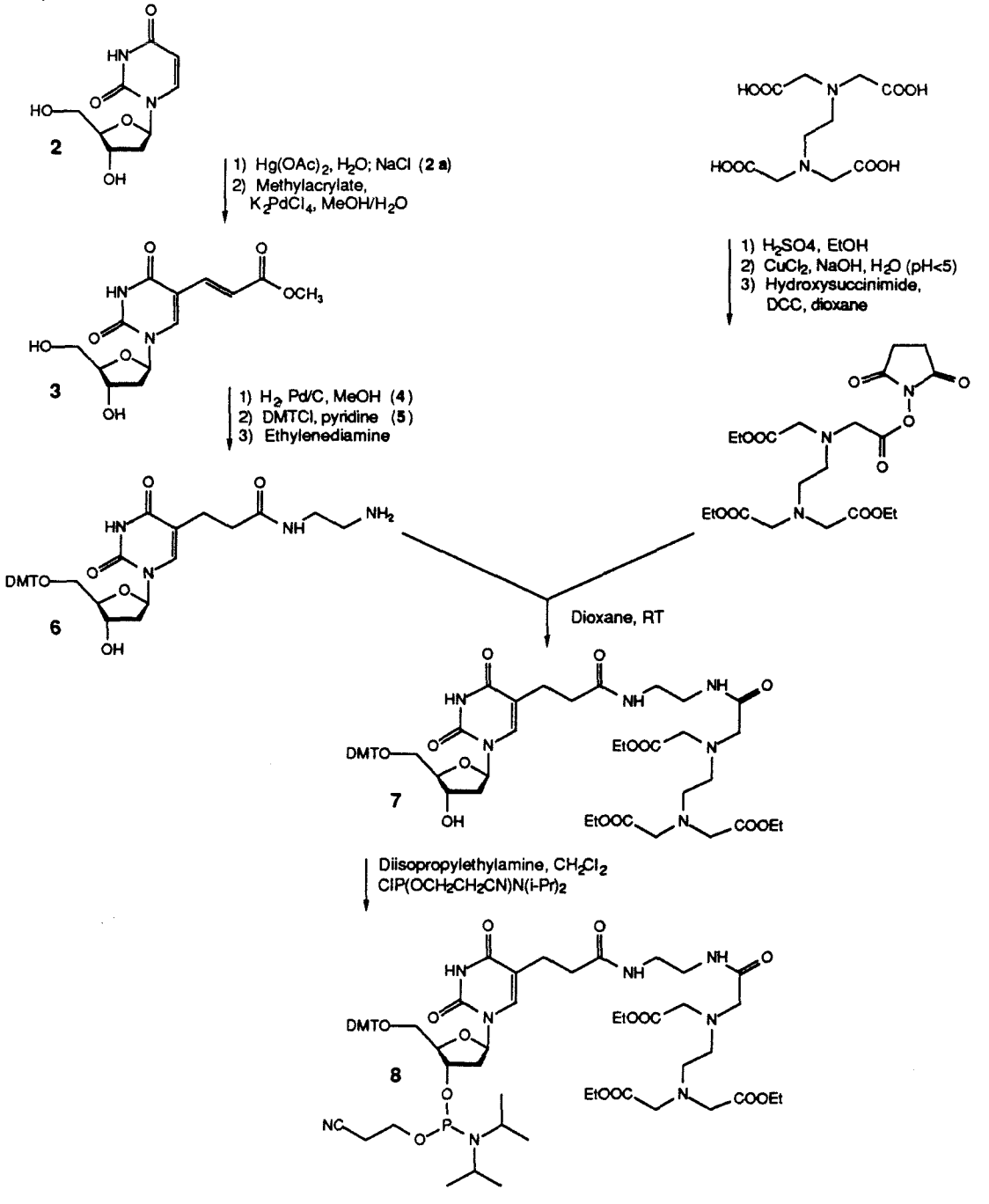


Figure 1. Scheme for the synthesis of DMT-protected T* phosphoramidite as described by Dreyer and Dervan (1985).

Synthesis of T Phosphoramidite* **8**. The phosphoramidite was prepared by the method of McBride & Caruthers (1983). Chloro-*N,N*-diisopropylaminomethoxy phosphine (70 µl, 375 µmol) was added dropwise to a solution of nucleoside **7** (250 mg, 250 µmol; predried for 12 hours at < 0.1 torr) and diisoproplyethylamine (135 µl, 750 µmol) in dry CH_2Cl_2 (1 ml) under Ar. After 1 hour, TLC (CH_2Cl_2/isopropanol/TEA, 97:2:1, v/v/v; R_f=0.32) indicated complete reaction and ethanol (0.5 ml) was added to the solution. The mixture was diluted with ethylacetate (10 ml), then extracted with saturated aqueous $NaHCO_3$ (2 x 10 ml) and saturated aqueous NaCl (2 x 10 ml). The organic layer was dried over anhydrous Na_2SO_4, chromatographed (CH_2Cl_2/isopropanol/TEA, 97:2:1, v/v/v), and concentrated under vacuum. The resulting gum was dissolved in dry CH_2Cl_2 (1 ml) and reconcentrated under vacuum (< 0.1 torr, 3 hours) to yield the T* phosphoramidite (242 mg, 208 µMol; 83% yield) as a white foam. ^1H NMR (acetone-d6) (diastereomers) δ 8.25 (1H, NH), 7.58 (1H, H_6), 7.52 (2H, m, ArH), 7.4-7.3 (7H, m, ArH), 7.25 (1H, t, NH), 7.0-6.9 (5H, m, ArH + NH), 6.32 (1H, m, $H_{1'}$), 4.65 (1H, m, $H_{3'}$), 4.15-4.11 (7H, m, OCH_2*CH_3 + $H_{4'}$), 3.80 (6H, s, Ar-OCH_3), 3.60 (6H, NCH_2CO + PNCH), 3.38-3.23 ((13H, P-OCH_3 2d's) + $H_{5'}$ + $NCH_2CH_2(CH_2)_2$), 2.85 (4H, $CONHCH_2CH_2NCO$), 2.45-2.23 (6H, $H_{2'}$ + CH_2CH_2CO), 1.25-1.08 (21H, CH_3 + OCH_2CH_3*).

Synthesis of **1** *(5'-GCAACGCGAT*TAAGTTGGG-3')*. The synthesis of the fully protected oligomer **1** was accomplished by the phosphoramidite method (Beaucage & Caruthers, 1981; Matteucci & Caruthers, 1981; McBride & Caruthers, 1983; Adams *et al.*, 1984), 1 µM scale, using a Beckman System 1 Plus DNA synthesizer. T* phosphoramidite **8** (200 mg, 0.1 M in CH_3CN) was placed in the synthesizer and added in the usual manner. DMT cation measurements indicated phosphoramidite coupling efficiencies greater than 97% for T*. Treatment of the completed oligomer with *N*-methylpyrrolidone (Phosphate Demethylation Reagent from Beckman, 1.5 ml, 24 hours at room temperature) converted the phosphotriester to the diester. Full deprotection and

removal from the support was accomplished by addition of 0.1 N NaOH (1.5 ml) to the oligomer, which was then heated at 55°C for 24 hours. The supernatent was neutralized with glacial HOAc (6-7 μl) and spun through a G-10-120 Sephadex column chased by 2 x 200 μl H_2O. The crude oligonucleotide-EDTA **1** was lyophilized and purified by electrophoresis (550 V, 20 hours) on a 2 mm-thick 20% polyacrylamide gel (Maxam & Gilbert, 1980). The gel showed a major UV-absorbing band with traces of a ladder of faster running impurities. The major UV-absorbing band was cut out and the DNA isolated according to Maniatis *et al.* (1982) and dialysed. A sample of purified oligodeoxynucleotide-EDTA **1**, 5'-end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP, was homogenous by electrophoresis on a 20% polyacrylamide gel. The sequence was confirmed by Maxam-Gilbert chemical sequencing methods using a 20% polyacrylamide gel.

Preparation of Labeled Single-Stranded M13mp7 DNA. M13mp7 (+) strand phage DNA was digested with BamHI (Been & Champoux, 1983). Treatment with calf alkaline phosphatase, [γ-^{32}P]ATP, and T4 polynucleotide kinase yielded a 5'-end-labeled strand 7214 bases in length, containing the 19-nucleotide complement of the oligonucleotide-EDTA probe 5'-GCAACGCGAT*TAAGTTGGG-3'. Purification of the labeled M13mp7 was accomplished by gel electrophoresis (1.6% low melt agarose, 120 V, 8 hours). The M13mp7 DNA was visualized and isolated according to Maniatis *et al.* (1982) and further purified by dialysis.

Cleavage Conditions and Analysis. Unless otherwise noted, reaction mixtures (20 μl) contained oligonucleotide-EDTA **1** (0.25 μM) premixed (0°C, 10 min) with Fe(II) (20 μM), 5'-^{32}P-end-labeled M13mp7 template (4000 cpm), tris-acetate (25 mM, pH 7.4), NaCl (100 mM), and calf thymus DNA (100 μM bp). Reactions with additional carrier DNA contained unlabeled M13mp7 at a concentration of 8.37 nM in T* oligonucleotide hybridization sites. The reaction mixtures were heated to 65°C and then allowed to cool to room temperature over a period of 1 hour to effect hybridization. The cleavage reactions

were then cooled to 0°C and initiated by addition of an aqueous solution of dithiothreitol (2 μl), such that the final concentration was 3 mM dithiothreitol. The cleavage reactions were allowed to proceed at 0°C for 20 hours and were then terminated by freezing (-78°C) and lyophilization. These samples were suspended in 5 μl of formamide loading buffer, heat denatured, and loaded onto 0.4 mm-thick, 40 cm-long, 8% polyacrylamide (1:20 cross linked)/50% urea high-resolution sequencing gels (Maxam & Gilbert, 1980). Electrophoresis was conducted at 1400 V until xylene cyanol tracking dye had traveled 25 cm. The gels were dried and autoradiographed at room temperature using Kodak SB 5 film. Using the autoradiogram as a guide, each gel lane was cut into three portions. The first portion, 21 bands corresponding to the site covering 10 bases on either side of the T*A base, contained the oligonucleotide-directed cleavage. The second gel portion contained uncleaved DNA. The third portion contained the remainder of the activity, nonspecific background cleavage. The radioactivity of each portion was counted using a Beckman LS 3801 liquid scintillation counter. Cleavage efficiency of DNA by T* oligonucleotide **1** was calculated as the ratio of radioactivity of site-specific cleavage to total radioactivity in each gel lane.

Oligonucleotide Melting Temperature Studies. Melting temperature studies were performed on T* oligonucleotide **1** and its 19-nucleotide complement and the same duplex containing thymidine in place of T*. The complementary 19-mers in 10 mM tris-HCl (pH 7.5), and 50 mM NaCl were mixed in a 1:1 ratio at concentrations ranging from A_{260} = 0.2-2.0 (~1-10 μM in 19-mer). The solutions were heated from 15°C to 75°C and absorbance measured at 5°C intervals on a Beckman model 25 UV-Vis spectrophotometer. The effect of T* on melting temperature was also studied in the presence of 50 mM $MgCl_2$ or 20 μM $NiCl_2$. Absorbance at 260 nm versus temperature was plotted and T_m determined as the temperature at which half the duplex was melted.

RESULTS AND DISCUSSION

Synthesis. (Fig. 1) 5'-DMT-T* triethylester **7** was synthesized as described by Dreyer and Dervan (1985) with a few modifications. The T* phosphoramidite was prepared in a manner similar to that described by McBride & Caruthers (1983). Chloro-*N,N*-diisopropylaminomethoxyphosphine (1.5 mole equivalents) was added to T* (1 mole equivalent) and diisopropylamine (3 mole equivalents) in CH_2Cl_2 to yield the T* phosphoramidite. The T* phosphoramidite has been found to be a suitable monomer for machine oligonucleotide synthesis. It is stable (0.1 M in CH_3CN) in the DNA synthesizer under N_2 for several days and gives excellent coupling efficiencies (>97%). Oligonucleotide-EDTA **1** was machine synthesized and the sequence verified by chemical sequencing methods (Maxam & Gilbert, 1980) (data not shown). Use of the T* phosphoramidite in machine synthesis greatly facilitates T* oligonucleotide synthesis relative to manual techniques (Dreyer & Dervan, 1985).

Cleavage of DNA. (Fig. 2) The sequence specific cleavage of DNA by oligodeoxynucleotide-EDTA•Fe(II) was examined on a 7214 base 5'-^{32}P-end-labeled BamHI single-stranded restriction fragment of M13mp7 phage DNA, containing the 19-base complement to oligonucleotide-EDTA **1**. Single-stranded M13mp7 was chosen as the template for cleavage by T* oligonucleotide to eliminate the extra factor of a competing strand in hybridization when using double stranded DNA (Dreyer & Dervan, 1985) and also to demonstrate sequence-specific cleavage by oligonucleotide-EDTA in a large piece of DNA. The oligonucleotide-EDTA probe **1** (0.25 μM) premixed (0°C, 10 min) with 20 μM Fe(II) was combined with 5'-labeled template (4000 cpm), 25 mM tris-acetate (pH 7.4), 100 mM NaCl, and 100 μM bp calf thymus DNA. The reactions were heated to 75°C and then allowed to cool to 30°C over a period of 1 hour to effect hybridization. The cleavage reactions were cooled to 0°C and initiated by addition of dithiothreitol (3 mM), allowed to proceed at 0°C for 20 hours, and stopped by freezing and lyophilization. The reaction products were analyzed by high resolution gel electrophoresis (Maxam & Gilbert, 1980).

From the autoradiogram we observed cleavage at the site complementary to the oligonucleotide-EDTA probe. No other cleavage sites were observed. These results suggest that oligonucleotide-EDTA•Fe(II) **1** forms a stable duplex with its complementary sequence and in the presence of dioxygen and added reducing agents, such as dithiothreitol, effects localized DNA cleavage. Reactions were performed in the presence and absence of additional unlabeled single-stranded M13mp7 carrier DNA. In reactions with additional unlabeled M13mp7 (8.37 nM in T* probe **1** hybridization sites), significant cleavage was observed with as little as 15-fold excess of oligonucleotide-EDTA **1** to M13mp7 template. The presence of carrier slightly reduced non-specific cleavage of the DNA but did not decrease sequence-specific cleavage relative to reactions with no carrier at the same concentrations of oligonucleotide-EDTA **1** (data not shown).

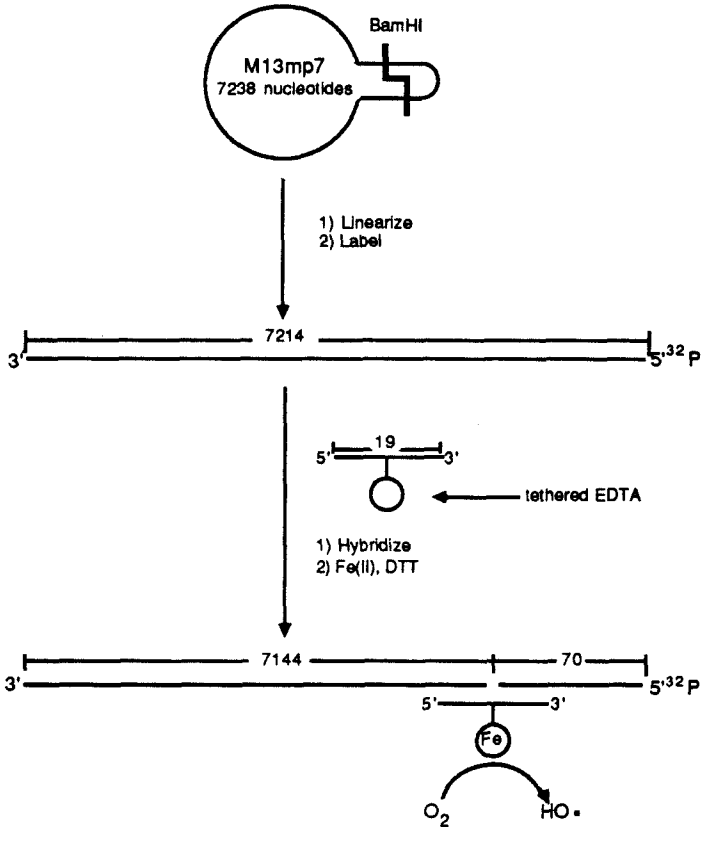
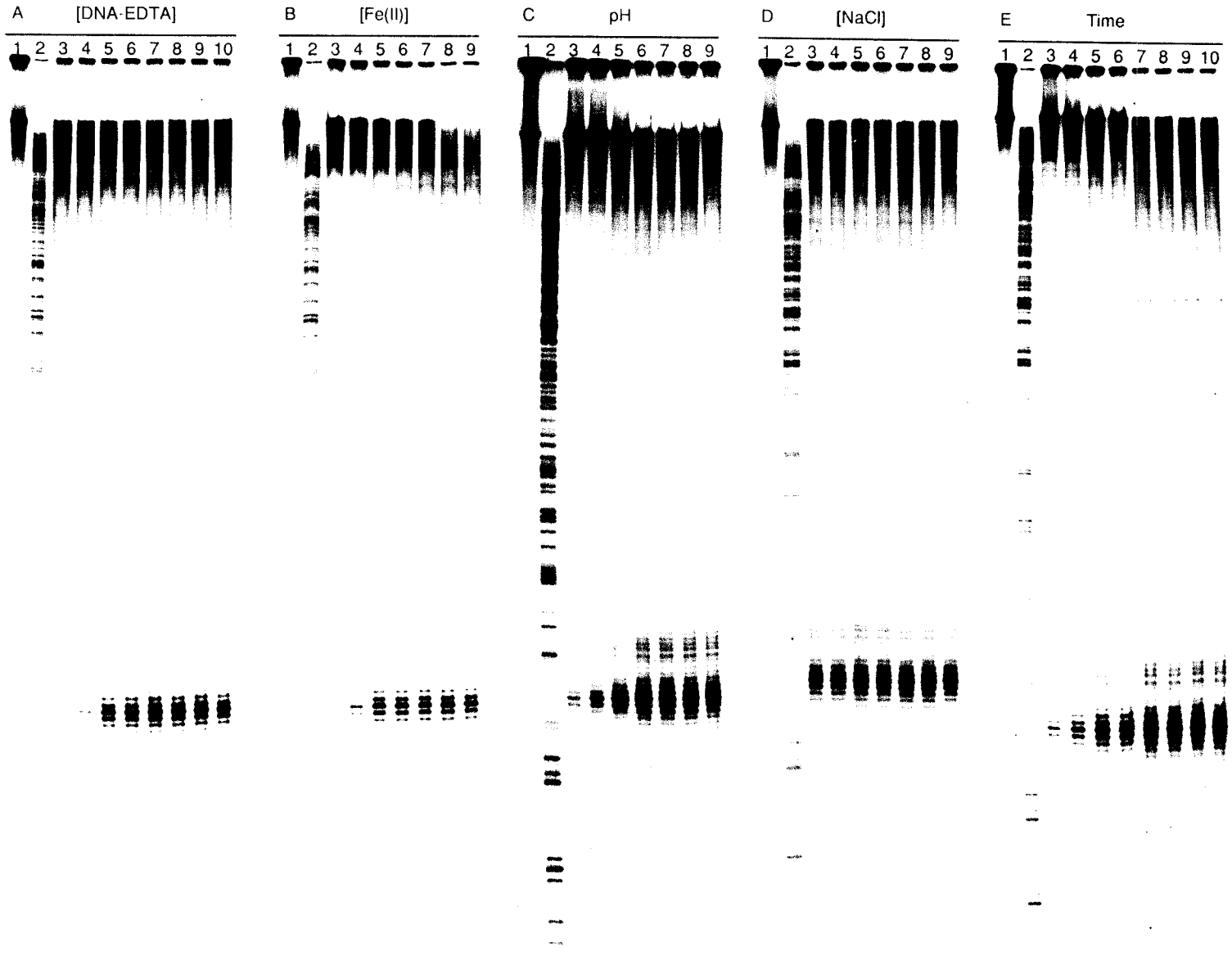


Figure 2. Schematic for oligonucleotide-directed cleavage of single-stranded M13mp7 DNA by a complementary 19-mer in which one thymidine has been replaced by the metal chelating T*.

Cleavage Efficiency of DNA by Oligonucleotide-EDTA **1**. Cleavage efficiency of DNA by oligonucleotide-EDTA **1** was calculated as the ratio of radioactivity of oligonucleotide-directed cleavage to the total radioactivity in each gel lane. Maximum cleavage efficiency was found to be approximately 35%. The maximum cleavage efficiency attainable may be limited by products from autocleavage of the oligonucleotide-EDTA•Fe(II). These products would compete with the intact T* oligonucleotide for hybridization of the M13mp7. Cleavage at the base at which the greatest amount of sequence specific cleavage occurs, four nucleotides to the 5' side of A•T*, is on the average at least 2000 times greater than nonspecific cleavage at any other base. Cleavage efficiency is dependent on several variables:

T Oligonucleotide Concentration.* The effect of T* oligonucleotide concentration was examined in the range of 0-5 μM. The optimum probe concentration was found to be between 0.25-1.00 μM. At concentrations of T* oligonucleotide greater than 2 μM cleavage efficiency gradually dropped off. The cause of the decrease in cleavage efficiency with higher oligonucleotide concentrations may be due to the presence of micromolar amounts of EDTA in the the T* oligonucleotide that would compete with T* for chelation of Fe(II). (Fig. 3A, 4A)

Figure 3 A-E. Autoradiograms of high resolution denaturing gel electrophoresis of 5'-end-labeled 7214 nucleotide M13mp7 (+)strand DNA. Lane 1 [A-E] are intact DNA control lanes after incubation under reaction conditions. Lane 2 [A-E] are chemical sequencing lanes containing the products of Maxam-Gilbert G specific cleavage reactions. The remaining lanes are the products of cleavage reactions (typically 4000 cpm labeled M13mp7 template, 0.25 μM T* oligonucleotide, 20 μM Fe (II), 100 mM NaCl, 100 μM bp calf thymus DNA, 25 mM tris-acetate (pH 7.4), and 3 mM DTT, 16 hr at 0°C, unless otherwise stated). The reactions were stopped ethanol precipitation and the cleavage products were analyzed by gel electrophoresis (1200-2000 v, xylene cyanol tracking dye 26 cm). T* oligonucleotide [A] (lane 3) 0 μM, (lane 4) .0025 μM, (lane 5) .025 μM, (lane 6) 0.25 μM, (lane 7) 0.5 μM, (lane 8) 1.0 μM, (lane 9) 2.0 μM, (lane 10) 5.0 μM; Fe(II) [B] (lane 3) 0 μM, (lane 4) 1 μM, (lane 5) 5 μM, (lane 6) 10 μM, (lane 7) 25 μM, (lane 8) 50 μM, (lane 9) 100 μM; pH [C] (lane 3) pH 5.8, (lane 4) pH 6.2, (lane 5) pH 6.6, (lane 6) pH 7.0, (lane 7) pH 7.4, (lane 8) pH 7.8, (lane 9) pH 8.2; NaCl [D] (lane 3) 0 mM, (lane 4) 20 mM, (lane 5) 50 mM, (lane 6) 100 mM, (lane 7) 200 mM, (lane 8) 500 mM, (lane 9) 1000 mM; Time [E] (lane 3) 1 hr, (lane 4) 2 hr, (lane 5) 4 hr, (lane 6) 8 hr, (lane 7) 16 hr, (lane 8) 24 hr, (lane 9) 32 hr, (lane 10) 42.


A
[DNA-EDTA]
1 2 3 4 5 6 7 8 9 10
B
[Fe(II)]
1 2 3 4 5 6 7 8 9
C
pH
1 2 3 4 5 6 7 8 9
D
[NaCl]
1 2 3 4 5 6 7 8 9
E
Time
1 2 3 4 5 6 7 8 9 10

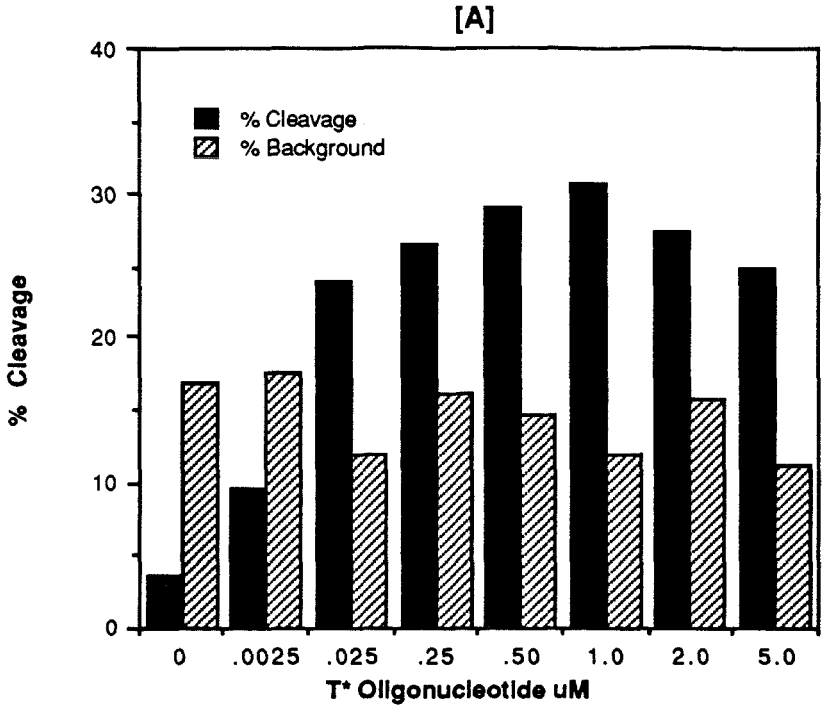
Fe(II) Concentration. The Fe(II) concentration was varied from 0 to 100 mM. A concentration of 20 mM gives optimum specific cleavage with minimum nonspecific cleavage. (Fig. 3B, 4B)

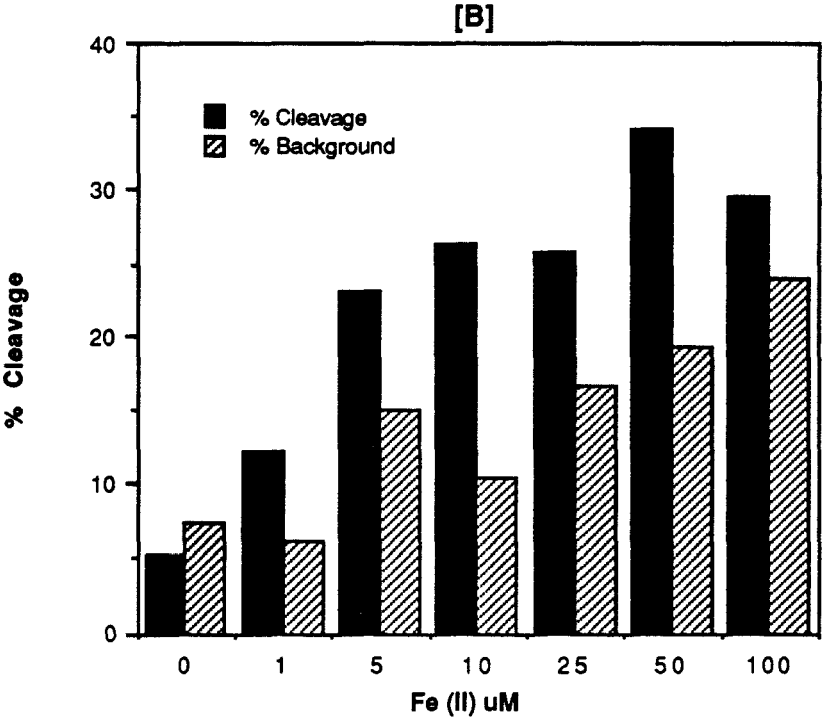
pH. The T* probe was allowed to react with M13mp7 DNA over a pH range of 5.8 to 8.2, using tris-acetate buffer. We find that efficient cleavage occurs between pH 7.0-7.8 with an optimum at 7.4. At pH values below 6 the EDTA becomes significantly protonated and can no longer bind Fe(II) as effectively resulting in little DNA cleavage. (Fig. 3C, 4C)

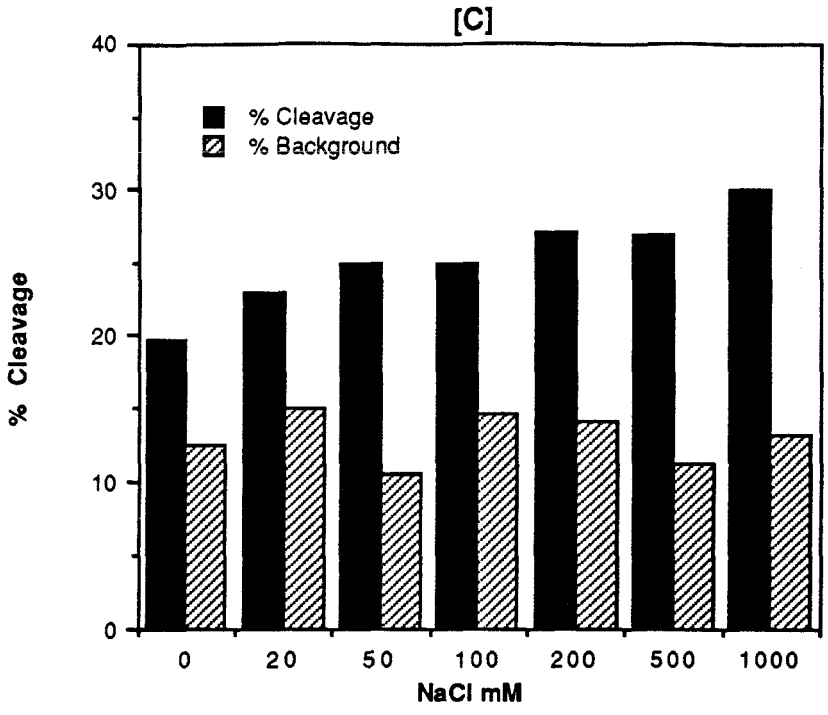
NaCl Concentration. The effect of sodium ion concentration was examined over the range of 0-1 M DNA. DNA cleavage by T* was efficient over the entire range studied. The hybridization of the T* probe to template is fairly insensitive to low NaCl concentration when the buffer concentration is 25 mM as opposed to 10 mM (data not shown). Cleavage efficiency dependence on NaCl concentration differs for a double-stranded system (Dreyer & Dervan, 1985) where the sequence specific cleavage drops off rapidly above 100 mM NaCl. This may be due to the increased ability of a higher NaCl concentration reaction mixture to stabilize the hybridization of the complimentary strand relative to the T* oligonucleotide. (Fig. 3D, 4D)

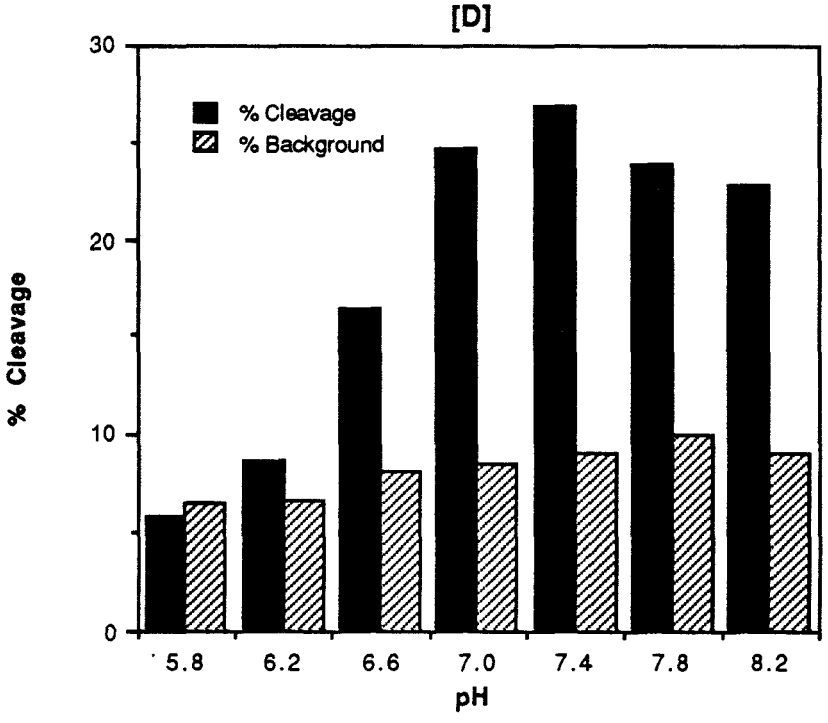
Time. The cleavage reaction of M13mp7 by oligonucleotide-EDTA **1** was stopped at several time intervals in order to follow the rate of cleavage. We find at 0°C the cleavage increases with time up to 16 hours and then levels off. (Fig. 3E, 4E)

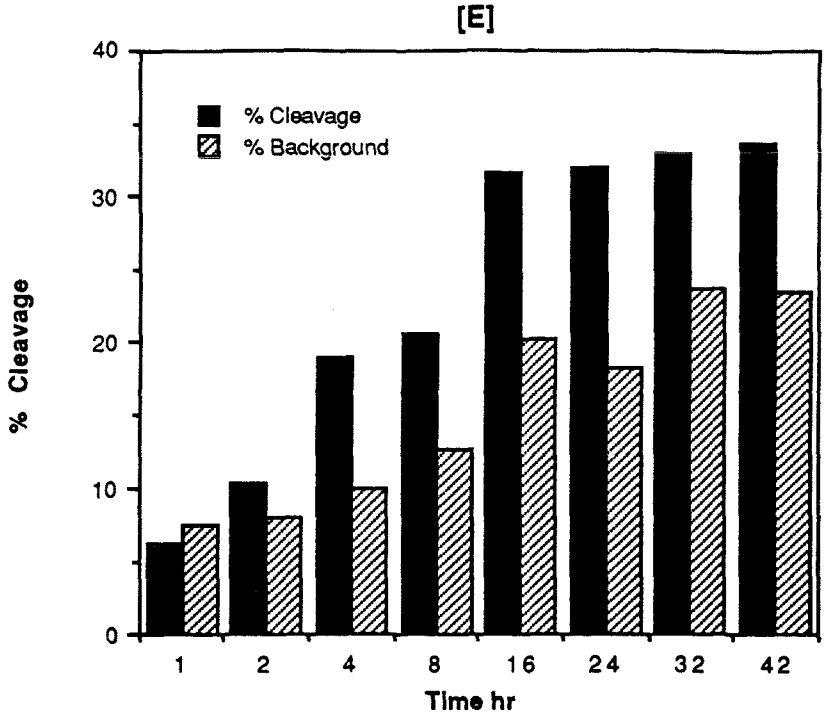
Figure 4 [A-E]. Bar graphs showing the amount of specific and nonspecific (background) cleavage derived from scintillation counting of the gels [A-E] shown in Figure 3. % Cleavage was calculated as the ratio of radioactivity of oligonucleotide directed cleavage to the total radioactivity in a gel lane. Background cleavage was calculated as the total radioactivity in a gel lane minus the radioactivity of the intact DNA and oligonucleotide directed cleavage. The data are reproducible to ± 10% of cleavage.


[A]
% Cleavage
% Background
% Cleavage
40
30
20
10
0
0
.0025
.025
.25
.50
1.0
2.0
5.0
T* Oligonucleotide uM


[B]
% Cleavage
% Background
% Cleavage
40
30
20
10
0
0
1
5
10
25
50
100
Fe (II) uM


[C]
% Cleavage
% Background
40
30
20
10
0
% Cleavage
0
20
50
100
200
500
1000
NaCl mM


[D]
% Cleavage
% Background
30
20
10
0
% Cleavage
5.8
6.2
6.6
7.0
7.4
7.8
8.2
pH



DNA Cleavage Pattern Analysis. The cleavage pattern produced by oligonucleotide-EDTA•Fe(II) **1** extends over 17 contiguous nucleotides of the template strand centered at the position of T*. The cleavage pattern consists of two loci of unequal intensity. The ratio of major and minor site cleavage is approximately 4:1 with maxima occurring four nucleotides to the 5' and 3' side of T*. A histogram of the DNA cleavage pattern obtained from densitometry of the autoradiogram is presented in Figure 5. Since T* is modified at the C-5 position, the tethered EDTA•Fe(II) extends out into the major groove. We believe that the EDTA•Fe(II) cleaves DNA preferentially in the minor groove and therefore the reactive intermediate must diffuse into the adjacent minor grooves to effect cleavage of the DNA (Dervan *et al.*, in preparation). Thus the two cleavage loci arise from diffusion of the reactive species into the minor grooves above and below the EDTA•Fe(II) (Fig. 5). The major site of cleavage is consistent with CPK models showing the base at which maximum cleavage occurs to be closest to the EDTA•Fe(II). The proximity of the

EDTA•Fe relative to the target strand bases is most likely sensitive to structural sequence-composition effects and may explain why different cleavage patterns are observed at other target sequences (Dreyer and Dervan, 1985). The length of the tether and diffusion of the reactive species, believed to be hydroxyl radical, allows the range of cleavage observed.

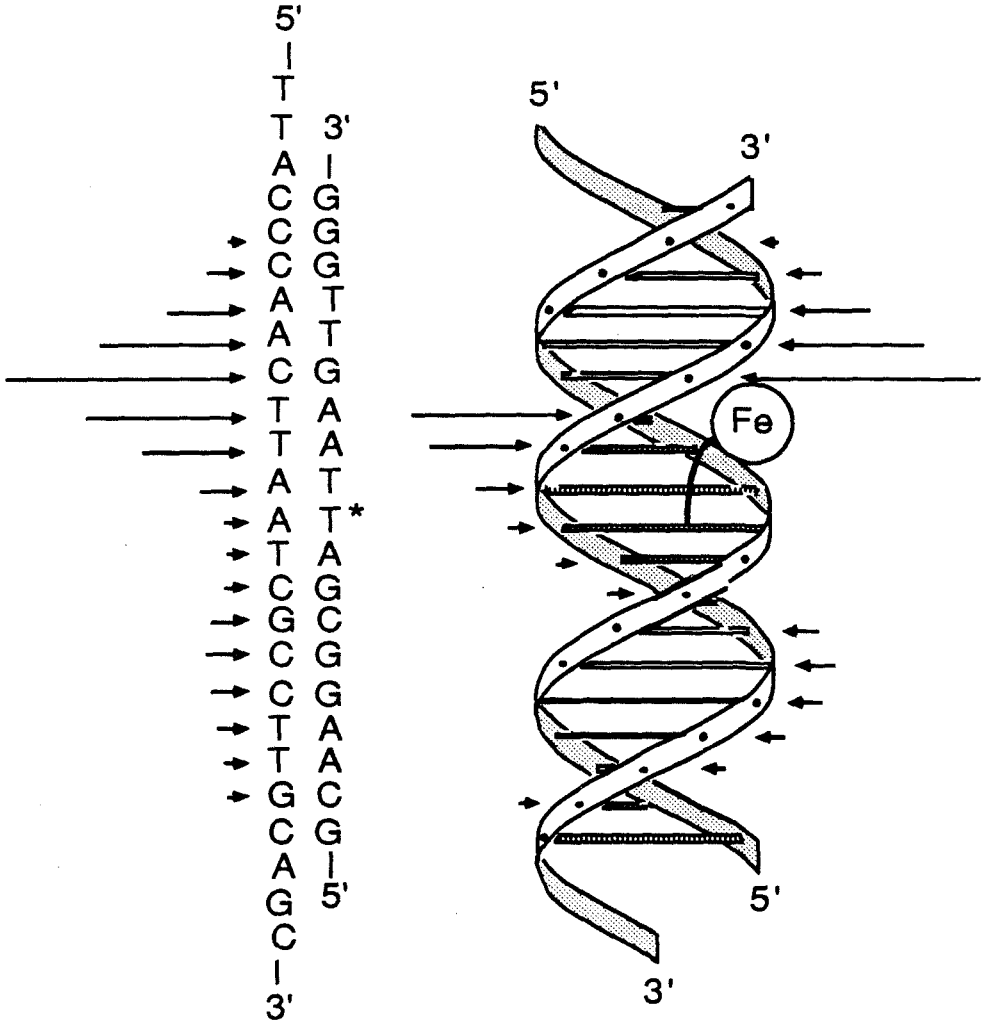


Figure 5. (Left) Histogram of the DNA cleavage pattern derived by densitometry of the autoradiogram shown in Figure 3 [B] lane 6. The heights of the arrows represent the relative cleavage intensities at the indicated bases. **(Right)** Simplified representation of the double-stranded complex between the T* oligonucleotide and the template M13mp7 (+)strand.

Oligonucleotide Melting Temperature Studies. The effect of T* on the melting temperature was studied by optical methods. Melting of doubled-stranded DNA causes an increase in UV absorbance and this effect can be used to determine T_m, the temperature at which half the DNA is single-stranded. Melting temperature studies were performed on T* oligonucleotide **1** and its 19-nucleotide complement and the same duplex containing thymidine in place of T*. The complementary 19-mers were mixed in a 1:1 ratio at concentrations ranging from A_{260}=0.2-2.0 (1-10 μM in 19-mer). The melting temperatures of the 19-mer duplexes in 10 mM tris-HCl (pH 7.5) and 50 mM NaCl alone or with 50 mM $MgCl_2$ or 20 μM $NiCl_2$ are compared in Figure 6. Under reaction concentrations, using Ni(II) or Mg(II) in place of Fe(II), which might give autocleavage of the T* oligonucleotide, the melting temperature of the T* duplex was 54.5°C, 3.5°C lower than that of the duplex without T*. Thus, the tethered EDTA of T* when chelating a divalent metal slightly destabilizes the 19-mer duplex. Chelation of a divalent metal by the EDTA reduces the electrostatic repulsion between the negatively charged EDTA and the DNA. Absence of a divalent metal results in a much lower T_m of 39°C.

SUMMARY

T* phosphoramidite has been shown to be DNA synthesizer compatible by the efficient machine synthesis of oligonucleotide-EDTA **1**. Oligonucleotide-EDTA **1** in the presence of Fe(II), O_2, and DTT cleaves a 7214-nucleotide-long DNA strand at its complementary site on the DNA strand. Optimum conditions for cleavage of M13mp7 (+) strand DNA (4000 cpm) occur at 0.25-1.00 μM T* oligonucleotide, 20 μM Fe(II), 50-1000 mM NaCl, in a 25 mM tris-acetate buffer at pH 7.4. The optimum reaction time is between 16-24 hours at 0°C. The cleavage pattern of two loci appears to arise from cleavage in the minor grooves above and below the major groove located EDTA•Fe. Under reaction conditions the T* 19-mer duplex (T_m=54.5°C) is slightly less stable than the

thymidine 19-mer duplex (T_m=58°C). In the absence of a divalent metal the melting temperature of the T* duplex is only 39°C.

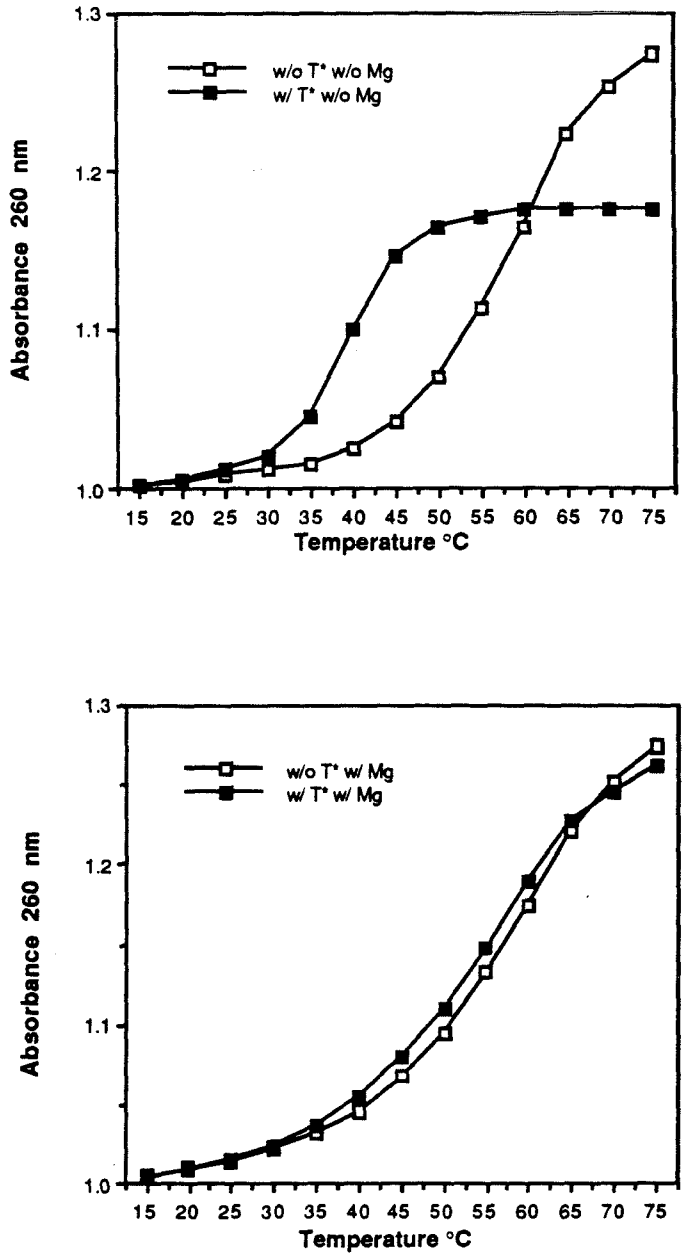


Figure 6. Melting temperature studies of the T* 19-mer **1** and its complement and the same duplex containing thymidine in place of T*. The complementary 19-mers were mixed in a 1:1 ratio at concentrations ranging form A_{260}=0.2-2.0 (1-10 μM in 19-mer) in 10 mM Tris-HCl (pH 7.5) and 50 mM NaCl in the presence or absence of 50 mM $MgCl_2$.

Cleavage of M13mp7 (+) strand DNA, 7214 nucleotides in length, has demonstrated the ability of oligonucleotide-EDTA probes to cleave large pieces of DNA sequence-specifically. This work suggests that these oligonucleotide-EDTA probes might be useful reagents capable of cleaving single-stranded nucleic acid uniquely at any desired site. Potential applications of this class of oligonucleotide-directed DNA cleaving molecule include the mapping of large genomes, site directed mutagenesis, sequence specific cleavage of RNA, diagnostic probes, and novel chemotherapeutic agents.

ACKNOWLEDGMENTS

We are grateful to Dr. Heinz Moser for suggestions for improvements in the synthesis of T* and to Dr. Geoff Dreyer for instruction in phosphoramidite oligonucleotide methods.

REFERENCES

Adams, S.P., Kavka, K.S., Wykes, E.J., Holder, S.B. & Galluppi, G.R. (1983) *J. Am. Chem. Soc. 105, 661-663.*

Beaucage, L.S. & Caruthers, M.H. (1981) *Tetrahedron Lett. 22*, 1859-1862.

Been, M.D. & Champoux, J.J. (1983) *Methods Enzymol. 101*, 90-98.

Boidot-Forget, M., Thuong, N.G., Chassignol, M. & Helene, C. (1986) *C.R. Acad. Sc., Ser. 2, 302 (2)*, 75-80.

Boutorin, A.S., Vlassov, V.V., Kasakov, S.A., Kutiavin, I.V. & Podyminogin, M.A. (1984) *FEBS Lett. 172*, 43-46.

Chu, B.C.F. & Orgel, L.E. (1985) *Proc. Natl. Acad. Sci. USA 82*, 963-967.

Dreyer, G.B. & Dervan, P.B. (1985) *Proc. Natl. Acad. Sci. USA 82*, 968-972.

Grineva, N.I. & Karpova G.G. (1973) *FEBS Lett. 32*, 351-355.

Grineva, N.I, Karpova, G.G., Kuznetsova, L.M., Venkstern, T.V. & Bayev, A.A. (1977) *Nucleic Acids Res. 4*, 1609-1631.

Hertzberg, R.P. & Dervan P.B. (1982) *J. Am. Chem. Soc. 104*, 6861-6863.

Hertzberg, R.P. & Dervan P.B. (1984) *Biochemistry 23*, 3934-3945.

Kim, S.C., Podhajska, A. & Szybalski, W. **Science 240**, 504 (1988).

Knorre, D.G. & Vlassov, V.V. (1985) *Progress in Nucleic Acid Research and Molecular Biology 32*, 291-321.

Maniatis, T., Fritsch, E.F. & Sambrook J. (1982) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Lavoratory, Cold Spring Harbor, NY.

Martin, R.F. & Haseltine, W.A. (1981) *Science 213*, 896-898.

Matteucci, M.D. & Caruthers, M.H. (1981) *Tetrahedron Lett. 22*, 1859-1862.

Maxam, A.M. & Gilbert, W. (1980) *Methods Enzymol. 65*, 599.

Maxam, A.M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA 74*, 560-564.

McBride L.J. & Caruthers, M.H. (1983) *Tetrahedron Lett. 24*, 245.

Modrich, Pl (1982) *Crit. Rev. Biochem. 13*, 288-323.

Podhajska, A.J. & Szybalski, W. *Gene 40*, 175 (1985).

Roberts, R.J. (1983) *Nucleic Acids Res. 11*, r135-167.

Schultz, P.G. & Dervan, P.B. (1984) *J. Biomol. Struct. Dyn. 1*, 1133-1147.

Schultz, P.G. & Dervan, P.B. (1983) *Proc. Natl. Acad. Sci. USA 80*, 6834-6837.

Schultz, P.G., Taylor, J.S. & Dervan, P.B. *J. Am. Chem. Soc. 104*, 6834-6837 (1982).

Smith, H.O. (1979) *Science 205*, 455-462.

Szybalski, W. *Gene 40*, 169 (1985).

Taylor, J.S., Schultz, P.G. & Dervan, P.B. (1984) *Tetrahedron 40*, 457-465.

Vlassov, V.V., Zarytova, V.F., Kutiavin, I.V., Mamaev, S.V. & Podyminogin, M.A. (1986) *Nucleic Acids Res.* 10, 4065-4076.

Youngquist, R.S. & Dervan, P.B. (1984) *Proc. Natl. Acad. Sci. USA 82*, 2565-2569.

Part II

Oligonucleotide-Directed Cleavage of Double-Stranded DNA by Triple Helix Formation

Overview

"The sequence-specific cleavage of double helical DNA by restriction endonucleases is essential for many techniques in molecular biology, including gene isolation, DNA sequence determination, and recombinant DNA manipulations (2,3). With the advent of pulsed-field gel electrophoresis, the separation of large pieces of DNA is now possible (4,5). However, the binding site sizes of naturally occurring restriction enzymes are in the range of 4 to 8 base pairs, and hence their sequence specificities may be inadequate for mapping genomes over very large distances. The design of sequence-specific DNA cleaving molecules that go beyond the specificities of the natural enzymes depends on a detailed understanding of the chemical principles underlying two functions: recognition and cleavage of DNA (6). Synthetic sequence-specific binding moieties for double helical DNA that have been studied are coupled analogs of natural products (5), transition metal complexes (7), and peptide fragments derived from DNA binding proteins (8,9) (1)."

"The DNA cleaving function used in our laboratories is EDTA•Fe(II), which cleaves the DNA backbone by oxidation of the deoxyribose with a short-lived diffusible hydroxyl radical (6,10). The fact that hydroxyl radical is a relatively nonspecific cleaving species is useful when studying recognition because the cleavage specificity is due to the binding moiety alone, not some combination of cleavage specificity superimposed on binding specificity. The most sequence-specific molecules characterized so far, with regard to the natural product analog approach, is bis (EDTA-distamycin) fumaramide, which binds in the minor groove and cleaves at sites containing 9 bp of contiguous A,T DNA (11). A synthetic peptide containing 52 residues from the DNA binding domain of Hin protein with EDTA at the amino terminus binds and cleaves at the 13-bp Hin site (9). Despite this progress, our understanding of molecular recognition of DNA is still sufficiently primitive that the elucidation of the chemical principles for creating specificity at the ≥ 15-bp level

may be slow when compared to the time scale for and interest in mapping large genome (1)."

Triple helix formation. One way to approach sequence-specific cleavage of large double helical DNA is with modified oligonucleotides that bind in the major groove, forming a triple helix structure (Fig. 1). "The first triplex of nucleic acids was reported three decades ago (12). Poly(U) and poly(A) were found to form a stable 2:1 complex in the presence of $MgCl_2$. After this, several triple-stranded structures were discovered (13,14). Poly(C) forms a triple-stranded complex at pH 6.2 with guanine oligonribo-

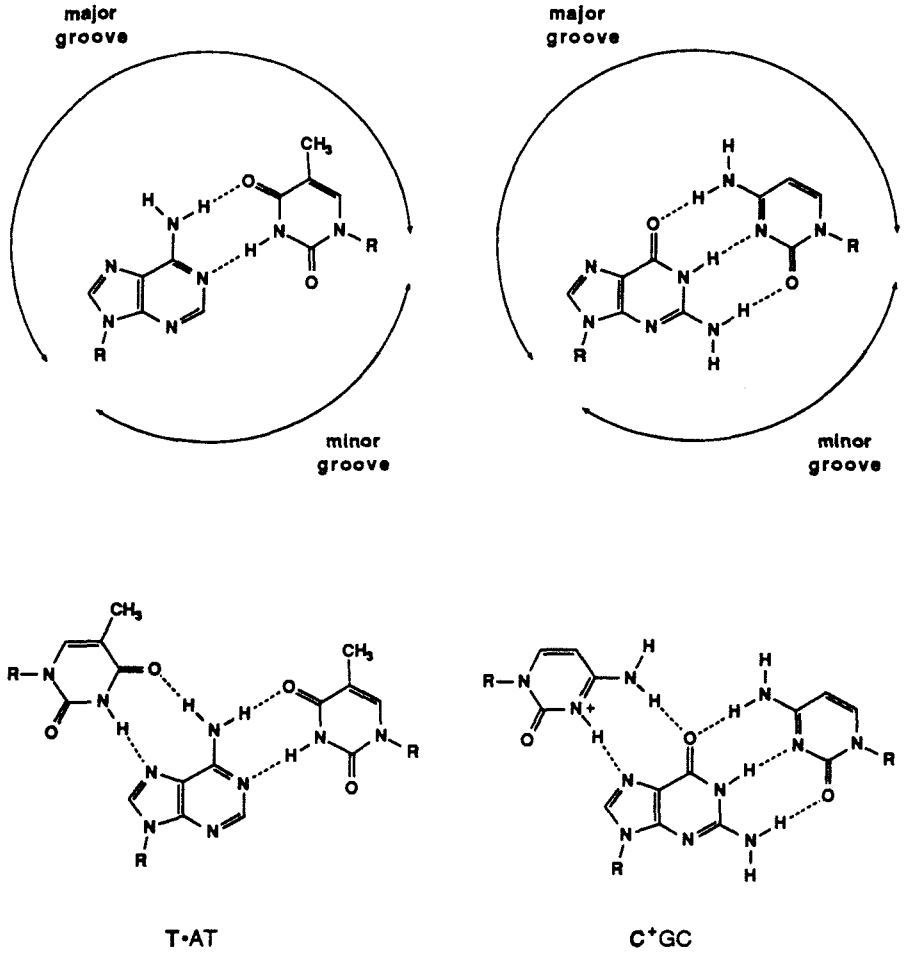


Figure 1. (Top) Watson-Crick base pairs. **(Bottom)** Isomorphous base triplets T•AT and C+GC. The additional pyrimidine base is bound in the major groove by Hoogsteen hydrogen bonds to the purine base in the Watson-Crick duplex. Protonation of the Hoogsteen bound cytosine is required.

nucleotides. One of the pyrimidine strands is in the protonated form (15-18). In principle, isomorphous base triplets (T•AT and C+GC) can be formed between any homopurine•homopyrimidine duplex and a corresponding homopyrimidine strand (19-21) (Fig. 1). The DNA duplex poly (dT-dC)•poly(dG-dA) associates with poly(U-C) or poly (dT-dC) below pH 6 in the presence of $MgCl_2$ to afford a triple-stranded complex (20, 21). Several investigators proposed an antiparallel orientation for the two polypyrimidine strands on the basis of an *anti* conformation of the bases (19-21). The x-ray diffraction patterns of triple-stranded fibers (poly(A)•2poly(U) and poly(dA)•2poly(dT)) supported this hypothesis (22-24) and suggested an A' RNA-like conformation of the two Watson-Crick base paired strands with the third strand in the same conformation, bound parallel to the homopurine strand of the duplex by Hoogsteen hydrogen bonds (25). The 12-fold helix with dislocation of the axis by almost 3Å, C3'-endo sugar puckering, and small base-tilts result in a large and deep major groove that is capable of accommodating the third strand (26). A high-resolution x-ray structure of a triple helical DNA or RNA has not yet been reported (1)."

Oligonucleotide-EDTA probes. Oligonucleotides equipped with a DNA cleaving moiety have been described, which produce sequence-specific cleavage of single-stranded DNA (27-30). An example of this is oligonucleotide-EDTA•Fe hybridization probes, which cleave the complementary single strand sequence (28,29). Moser and Dervan have reported that homopyrimidine oligodeoxyribonucleotides with EDTA•Fe attached at a single position bind the corresponding homopyrimidine•homopurine tracts within large double-stranded DNA by triple helix formation and cleave at that site (Fig. 2) (1). Oligonucleotides with EDTA•Fe at the 5' end cause a sequence-specific double strand break. The location and asymmetry of the cleavage pattern reveal that the homopyrimidine-EDTA probes bind in the major groove parallel to the homopurine strand of Watson-Crick double helical DNA. Cleavage by triple strand formation is not observed in the absence of cations such as spermine or $Co(NH_3)_6^{3+}$, which are required to overcome the repulsion

between two anionic chains of Watson-Crick duplex and a third negatively charged phosphodiester backbone. The efficiency of oligonucleotide duplex cleavage is increased by a factor of 10 on addition of ethylene glycol (40% by volume) or other organic solvents that lower the relative humidity. Dehydration of the DNA favors a B to A transition and can allow the formation of a triple helix (A' RNA-like) to occur more readily (26). Oligonucleotide-EDTA•Fe(II) probes do not bind-cleave Watson-Crick DNA in slightly basic solutions (pH ≥ 8). An important factor is the required protonation of cytosines at N-3 in the third strand to enable the formation of two Hoogsteen hydrogen bonds (24). Cleavage efficiency of oligonucleotide-EDTA•Fe(II) decreases sharply below pH 6 (10), presumably as a result of partial protonation of the EDTA and the resulting loss of Fe(II) or some pH dependence of the cleavage reaction. Footprinting experiments confirm that the triple helix is forming at acidic pH values (31). The sequence-specific recognition of double helical DNA by homopyrimidine probes is sensitive to single base mismatches.

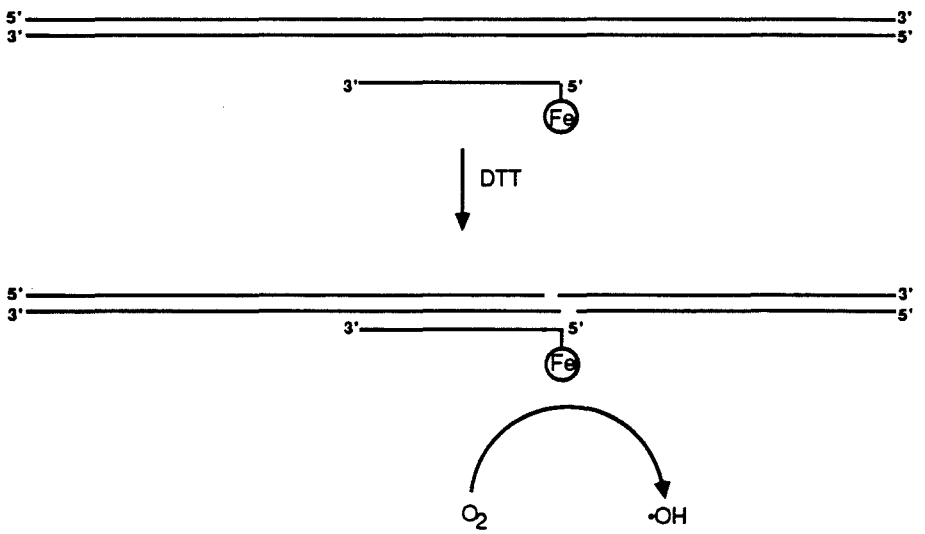


Figure 2. Oligonucleotide-directed cleavage of double helical DNA by a triple helix-forming oligonucleotide-EDTA•Fe probe. One thymidine has been replaced by thymidine with the iron chelator EDTA covalently attached at C-5. Reduction of dioxygen generateslocalized hydroxyl radical at this position (28).

Strobel, Moser, and Dervan have now reported the double strand cleavage of genomic DNA at a single site by triple helix formation (32). An oligonucleotide-EDTA•Fe probe (0.8 μM) equipped with thymidine-EDTA (T*) at the 5' end, 5'-T*$T_3CT_6CT_4CT$-3', causes double strand cleavage at a single homopurine site 18 base pairs in size (5'-$A_4GA_6GA_4GA$-3') within 48,502 base pairs of bacteriophage λ DNA (1 μM in base pairs). The double strand cleavage efficiency is 25% (100 mM NaCl, 25 mM tris-acetate, pH 7.0, 1 mM spermine, 24°C). No secondary cleavage sites (at partially homologous sequences) were detected under these reaction conditions. The oligonucleotide-EDTA•Fe mediated site-specific double strand cleavage of DNA can also be carried out in a low melting point agarose matrix.

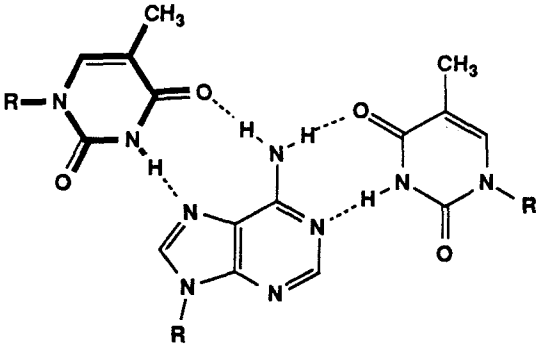
Povsic and Dervan recently reported triple helix formation by oligonucleotides on DNA extended to the physiological pH range (33). The incremental effect on the affinity and pH dependence of triple helix formation by substitution at the 5-position of pyrimidines in the Hoogsteen strand was studied. Oligonucleotide-EDTA•Fe probes containing 5-bromouracil and 5-methylcytosine bind and cleave polypurine sequences in duplex DNA with specificities comparable to those of their thymine-cytosine analogues, but with greater affinities and over an extended pH range. Oligonucleotides containing uracil bind with lower affinity. Because oligonucleotide specificity could provide a basis for the design of artificial repressors of gene expression and viral diseases, this work has important implications for controlling triple helix formation *in vivo* where the intracellular pH range is strictly regulated and compartmentalized.

Maher, Wold, and Dervan have demonstrated the inhibition of DNA binding proteins by oligonucleotide-directed triple helix formation (34). Oligonucleotides that bind to duplex DNA in a sequence specific manner by triple helix formation offer an approach to the experimental manipulation of sequence specific protein binding. Micromolar concentrations of pyrimidine oligodeoxyribonucleotides are shown to block recognition of double helical DNA by prokaryotic modifying enzymes and a transcription factor at a

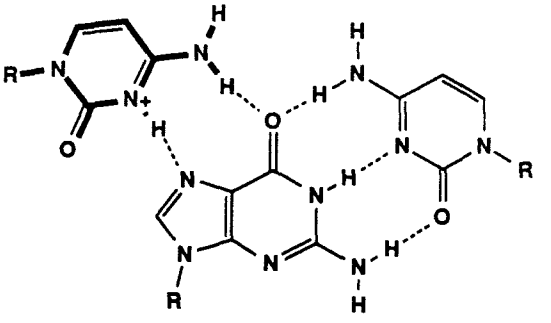
homopurine target site. Inhibition is sequence specific. Oligonucleotides containing 5-methylcytosine provide substantially more efficient inhibition than oligonucleotides containing cytosine. These results have implications for gene-specific repression by oligonucleotides or their analogs.

Less well understood is purine oligonucleotide recognition of double helical DNA (purine•purine•pyrimidine triplets) (35-38). Recently, purine oligonucleotides have been postulated to bind parallel to purines in duplex DNA by triple helix formation (A•AT and G•GC base triplets) (38). Cooney *et al.* have demonstrated *in vitro* gene repression using a G-rich oligonucleotide to bind a G-rich target site (38). α-oligonucleotides have also been observed to form triple-stranded structures (39,40).

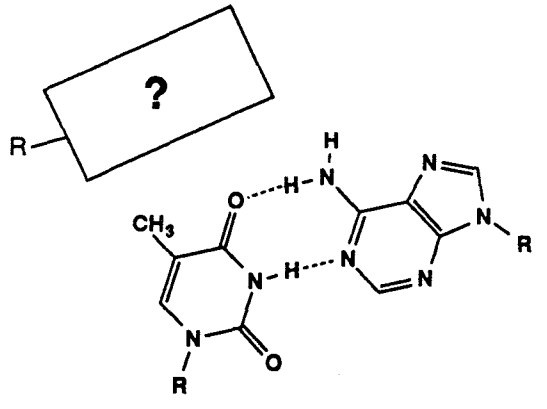
Studies in the area of oligonucleotide-directed recognition and cleavage of duplex DNA by triple strand formation clearly demonstrate the power and potential of this form of duplex DNA recognition. However, because triple strand formation is presently limited to homopurine•homopyrimidine tracts of DNA (AT and GC base pairs), a general solution to duplex DNA recognition is still lacking. The extension of recognition by triple helix formation to any sequence (AT, GC, TA and CG base pairs, Fig. 3) is an interesting and challenging problem. The quest for a general solution will test and refine our understanding of molecular recognition. We have chosen to approach the problem in two ways. The first approach is to test readily available bases for TA and/or CG recognition (Chapter 2) and the second is a rational design synthetic approach (Chapter 3).



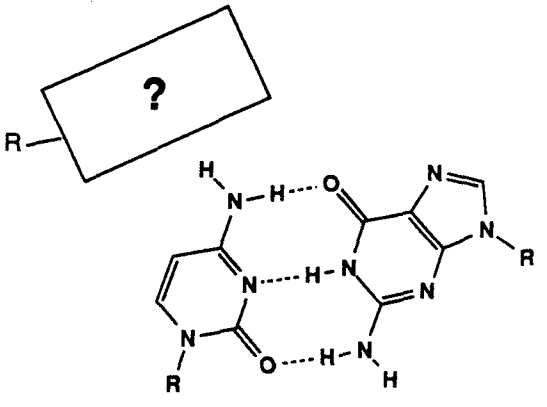
T•AT



C⁺GC



?•TA



?•CG

Figure 3. (Top) Isomorphous base triplets **T•AT** and **C⁺GC**. **(Bottom)** Unknown new base triplets **?•TA** and **?•CG**.

References and Notes

1. H.E. Moser and P.B. Dervan, *Science* **238**, 645 (1987).
2. H.O. Smith, *Science* **205**, 455 (1979).
3. P. Modrich, *Crit. Rev. Biochem.* **13**, 287 (1982).
4. D. Schwartz and C.R. Cantor, *Cell* **37**, 67 (1984).
5. G. F. Carle, M. Frank, M.V.C. Olsen, *Science* **232**, 65 (1986).
6. P. B. Dervan, *ibid.*, p. 464.
7. J. K. Barton, *ibid.* **233**, 727 (1986).
8. M. Bruist, S.J. Horvath, L.E. Hood, T.A. Steitz, M.I. Simon, *ibid.* **235**, 777 (1987).
9. J. Sluka, M. Bruist, S.J. Horvath, M.I. Simon, P.B. Dervan, *ibid.* **238**, 1129 (1987).
10. R.P. Hertzberg and P.B. Dervan, *Biochemistry* **23**, 3934 (1984).
11. R.S. Youngquist and P.B. Dervan, *J. Am. Chem. Soc.* **107**, 5528 (1985).
12. G. Felsenfeld, D.R. Davies, A. Rich, *ibid.* **79**, 2023 (1957).
13. A.M. Michaelson, J. Massoulie, W. Guschlbauer, *Prog. Nucleic Acids Res. Mol. Biol.* **6**, 83, (1967).
14. G. Felsenfeld and H.T. Miles, *Annu. Rev. Biochem.* **36**, 407 (1967).
15. M.N. Lipsett, *Biochem. Biophys. Res. Commun.* **11**, 224 (1963).
16. M.N. Lipsett, *J. Biol. Chem.* **239**, 1256 (1964).
17. F.B. Howard, J. Frazier, M.N. Lipsett, H.T. Miles, *Biochem. Biophys. Res. Commun.* **17**, 93 (1964).
18. P. Rajagopal and J. Feigon, *Nature* **239**, 637 (1989).
19. J.H. Miller and J.M. Sobell, *Proc. Natl. Acad. Sci. USA* **55**, 1201 (1966).
20. A.R. Morgan and R.D. Wells, *J. Mol. Biol.* **37**, 63 (1968).
21. J.S. Lee, D.A. Johnson, A.R. Morgan, *Nucleic Acids Res.* **6**, 3073 (1979).

22. S. Arnott and P.J. Bond, *Nature (London) New Biol.* **244**, 99 (1973).

23. S. Arnott and E. Selsing, *J. Mol. Biol.* **88**, 509 (1974).

24. S. Arnott, P.J. Bond, E. Selsing, P.J.C. Smith, *Nucleic Acids Res.* **3**, 2459 (1976).

25. K. Hoogsteen, *Acta Cryst.* **12**, 822 (1959).

26. W. Saenger, Principles of Nucleic Acid Structure, C.R. Cantor, Ed. (Springer-Verlag, New York, 1984).

27. G.D. Knorre and V.V. Vlassov, *Prog. Nucleic Acid Res. Mol. Biol.* **32**, 291 (1985), and references therein.

28. G.B. Dreyer and P.B. Dervan, *Proc. Natl. Acad. Sci. USA* **82**, 968 (1985).

29. C.F. Chu and L.E. Orgel, *ibid.* **82**, 963 (1985).

30. B.L. Iverson and P.B. Dervan, *J. Am. Chem. Soc.* **109**, 1241 (1987).

31. T. Povsic and P.B. Dervan, unpublished observations.

32. S.A. Strobel, H.E. Moser, P.B. Dervan, *J. Am. Chem. Soc.* **110**, 7927 (1988).

33. T.J. Povsic and P.B. Dervan, *ibid.* **111**, 3059 (1989).

34. L.J. Maher, B.J. Wold and P.B. Dervan, Science in press (1989).

35. C. Marck and D. Thiele, *Nucleic Acids Res.* **5**, 1017 (1978).

36. S.L. Broitman, D.D. Im, and J.R. Fresco, *Proc. Natl. Acad. Sci. USA* **84**, 5120 (1987).

37. A.J. Letai, *et al. Biochemistry* **27**, 9108 (1988).

38. M. Cooney, *et al. Science* **241**, 456 (1988).

39. T.L. Doan, *et al. Nucleic Acids Res.* **15**, 7749 (1987).

40. C. Praseuth, *et al. Proc. Natl. Acad. Sci. USA*, **85**, 1349 (1988).

Chapter 2

Recognition of Thymine•Adenine Base Pairs by Guanine in a Pyrimidine Triple Helix Motif

The sequence-specific recognition of double helical DNA is essential for the regulation of cellular functions including transcription, replication, and cell division. The ability to design synthetic molecules that bind sequence specifically to unique sites on human DNA could have major implications for the treatment of genetic, neoplastic, and viral diseases (1-3). Pyrimidine oligodeoxyribonucleotides (15-18 nucleotide (nt) oligomers) bind homopurine sites within large double-strand DNA by triple helix formation (4-9). Pyrimidine oligonucleotides bind in the major groove, parallel to the purine strand of Watson-Crick double helical DNA (4). Specificity is due to Hoogsteen hydrogen bonding, wherein thymine (T) recognizes adenine•thymine (AT) base pairs (T•AT triplet) and protonated cytosine (C) recognizes guanine•cytosine (GC) base pairs (C+GC triplet (10-23) (Overview Fig. 1). In addition to length and sequence composition, the binding affinity and specificity of the pyrimidine oligonucleotide for duplex DNA is sensitive to pH, organic cosolvent, added cations, and temperature (4-8). Less well understood is purine oligonucleotide recognition of double helical DNA (pur•pur•pyr triplets) (21,24-26). Recently, purine oligonucleotides have been postulated to bind parallel to purines in duplex DNA by triple helix formation (A•AT and G•GC base triplets) (26).

We examined the relative affinities of common bases for all four base pairs within a pyrimidine triple helix motif. We report that G in a pyrimidine oligonucleotide specifically recognizes TA base pairs within mixed purine/pyrimidine sites. We believe that this G•TA triplet represents a new specific interaction stabilizing triplex formation in mixed purine/pyrimidine sequences. Although there will undoubtedly be sequence composition effects, this finding extends specific recognition within the pyrimidine triple helix motif to three of the four possible base pairs in double helical DNA.

Recognition of TA by G was revealed by the study of the effects on triple strand formation of 20 possible base triplets at a single common position (Table I). The use of oligonucleotides equipped with the DNA cleaving moiety, thymidine-EDTA•Fe(II) (T*) (27,28), allowed the relative stabilities of triple helix formation between 30 base pair (bp) DNA duplexes containing the site $d(A_7XA_7)$•$d(T_7YT_7)$ (XY=AT, GC, CG, or TA) and a series of 15 nt oligomers differing at one base position $d(T_7ZT_7)$ [Z=T,C, A, G, or I (I=inosine] to be determined by the affinity cleaving method (12) (Fig. 1A). The 30-bp duplexes were labeled with ^{32}P at the 5' end of the target site strand $d(T_7YT_7)$. The DNA binding-cleaving reactions were performed under conditions that were sensitive to the stability of the variable base triplet in the middle of a thymine 15 nt fragment upon triple helix formation (pH 7.0, 23°C, 40% ethanol). The most intense cleavage patterns were observed for the combinations Z=T, XY=AT; Z=C, XY=GC; and Z=G, XY=TA (Fig. 1B and Table I). The cleavage observed for two of these combinations (lanes 3 and 8)

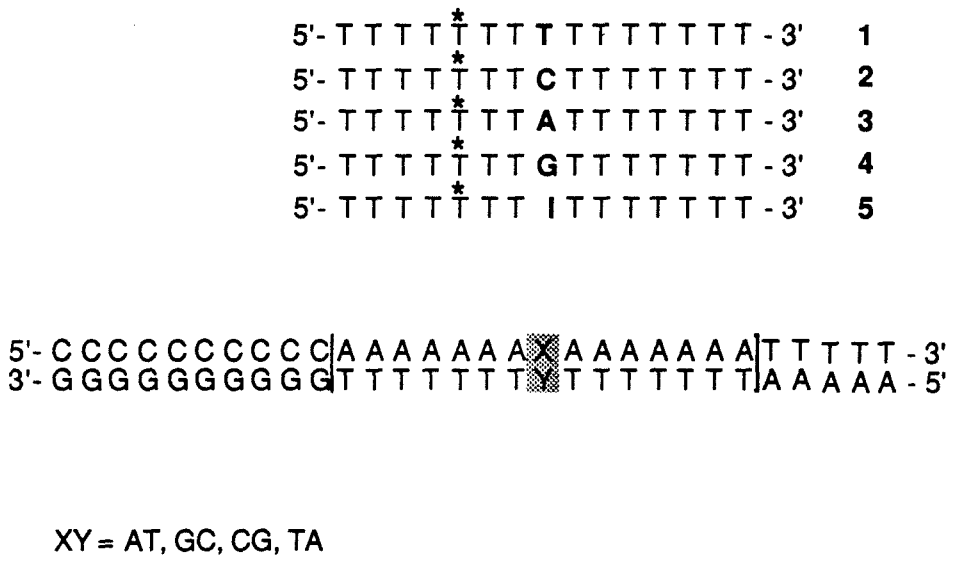


Figure 1A. **(Above)** Sequences of oligonucleotide-EDTA **1-5** where T* is the position of the thymidine-EDTA. The oligonucleotides differ at one base position indicated in bold type. **(Below)** The box indicates the double stranded sequence bound by oligonucleotide-EDTA•Fe(II) **1-5**. The Watson-Crick base pair (AT, GC, CG, or TA) opposite the variant base in the oligonucleotide is shaded. (See Table 1.)

represents the known ability of T and C to form T•AT and C+GC base triplets, respectively. Remarkably, intense cleavage is also observed for a G•TA base triplet (lane 18) and implies that G may form specific hydrogen bonds to TA. The lack of binding-cleavage observed for base triplets G•AT, G•GC, and G•CG (lanes 15,16, and 17, Fig. 2A, 2C) demonstrates that G is specific for a TA base pair. The lack of binding-cleavage observed for base triplets T•TA, C•TA, A•TA, and I•TA show that the base pair TA is also specific for G in the third strand (lanes 6,10, 14, and 22, Fig. 2A, 2C) (Table I). In order to test for the importance of the amino group of G in the G•TA base triplet, G was replaced with inosine (I), a purine base that lacks NH2 at position 2. The lack of cleavage observed with oligonucleotide 5 [Z=I] indicates that the amino group of guanine is important for recognition of TA base pairs. No binding is observed for Z=8-bromoguanosine, XY=TA (see appendix). Because 8-bromoguanosine prefers the *syn* conformation (29), the simplest model is a G•TA base triple structure where G is in the *anti* conformation with respect to the sugar ring and forms at least one hydrogen bond between G and T (G-N2 to T-O4) (Fig. 2). Finally, minor cleavage was observed for triplets T•GC, T•CG, C•CG, and G•CG (lanes 4, 5, 9, and 17). These bases may hydrogen bond to the Watson-Crick base pair to a lesser extent or may be sterically compatible with triple strand formation.

Figure 1B. Autoradiogram of the 20 percent denaturing polyacrylamide gel. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (1 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled 30-mer duplex [~0.5 μM (bp) (~12,500±600 cpm)] in a solution of tris-acetate, pH 7.0 (25 mM), NaCl (100 mM), calf thymus DNA [100 μM (bp)], and 40% ethanol and incubated at 0°C for 30 minutes and at 23°C for 30 minutes. Cleavage reactions were initiated by addition of DTT (3mM) and allowed to proceed for 6 hours at 23°C. The reactions were stopped by freezing and lyophilization and the cleavage products were analyzed by gel electrophoresis (1200-2000 V, BPB 23 cm). (Lanes 1-22) Duplexes containing 5' end-labeled $d(A_5T_7YT_7G_{10})$. (Lane 1) Control showing intact 5' labeled 30 bp DNA standard obtained after treatment according to the cleavage reactions in the absence of oligonucleotide-EDTA. (Lane 2) Products of Maxam-Gilbert G+A sequencing reaction. (Lanes 3-22) DNA cleavage products produced by oligonucleotide-EDTA•Fe(II) (**1-5**); **1** (Lanes 3-6); **2** (lanes 7-10); **3** (lanes 11-14); **4** (lanes 15-18); **5** (lanes 19-22). **XY=AT** (Lanes 3,7,11,15, and 19); **XY=GC** (lanes 4,8,12,16, and 20); **XY=CG** (lanes 5,9,13,17, and 21); **XY=TA** (lanes 6,10,14,18, and 22).

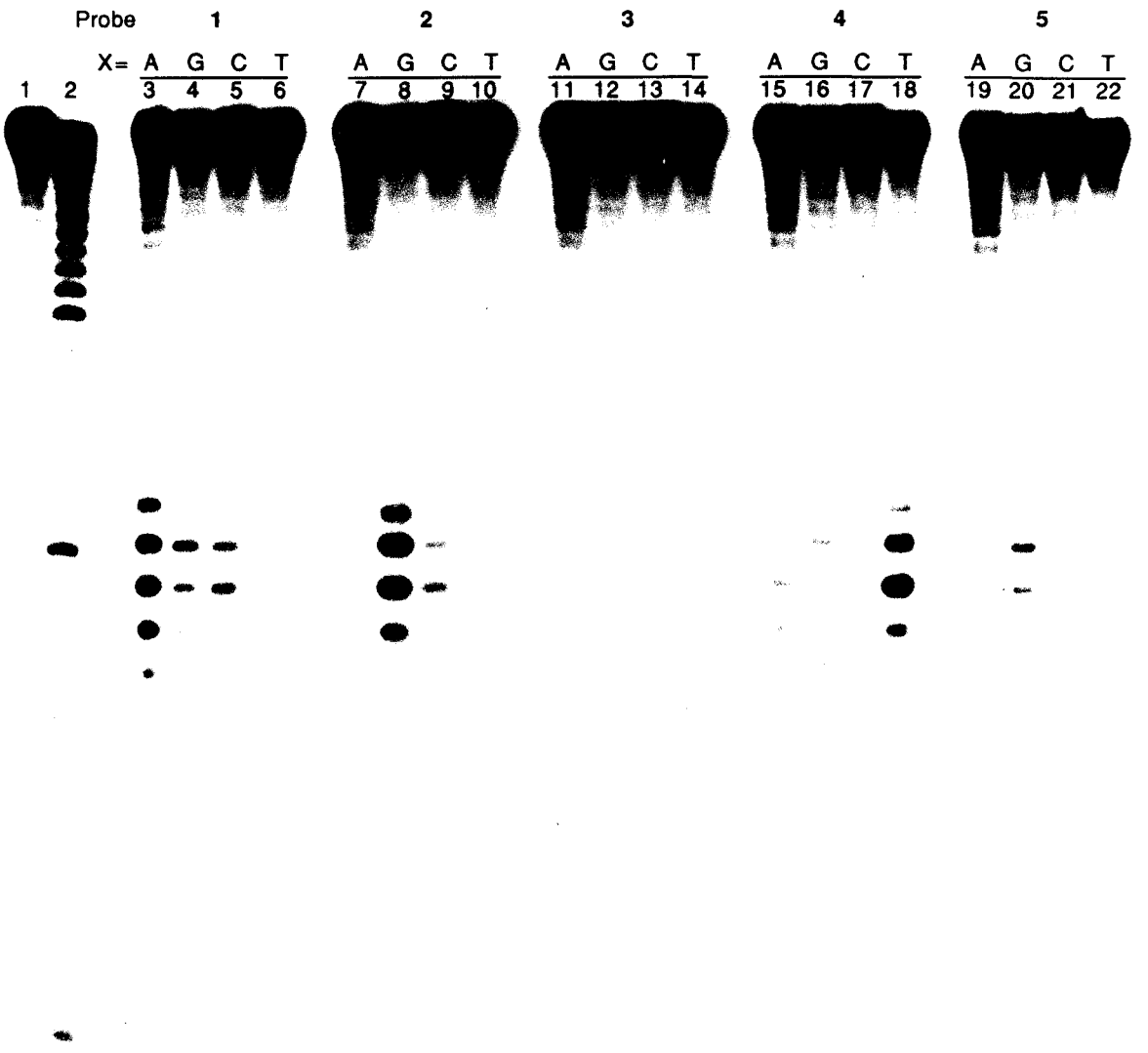

Probe
1
2
3
4
5
X= A G C T
1 2
3 4 5 6
7 8 9 10
11 12 13 14
15 16 17 18
19 20 21 22

Table 1

WC Duplex (XY)	Pyrimidine Third Strand (Z)				
	T	C	A	G	I
AT	+++	-	-	-	-
GC	+	+++	-	-	-
CG	+	+	-	+	-
TA	-	-	-	+++	-

Table 1. Twenty base triplets were examined for binding specificity compatible with pyrimidine-Hoogsteen triple helix motif by the experiment described in Figure 1A and 1B. (+++) ≥ 25-30 (±3)% cleavage, (+) > 10-15 (±1.5)% cleavage, (-) < 5 (±0.5)% cleavage. The data are from scintillation counting and densitometric analysis of the autoradiogram shown in Fig. 1B.

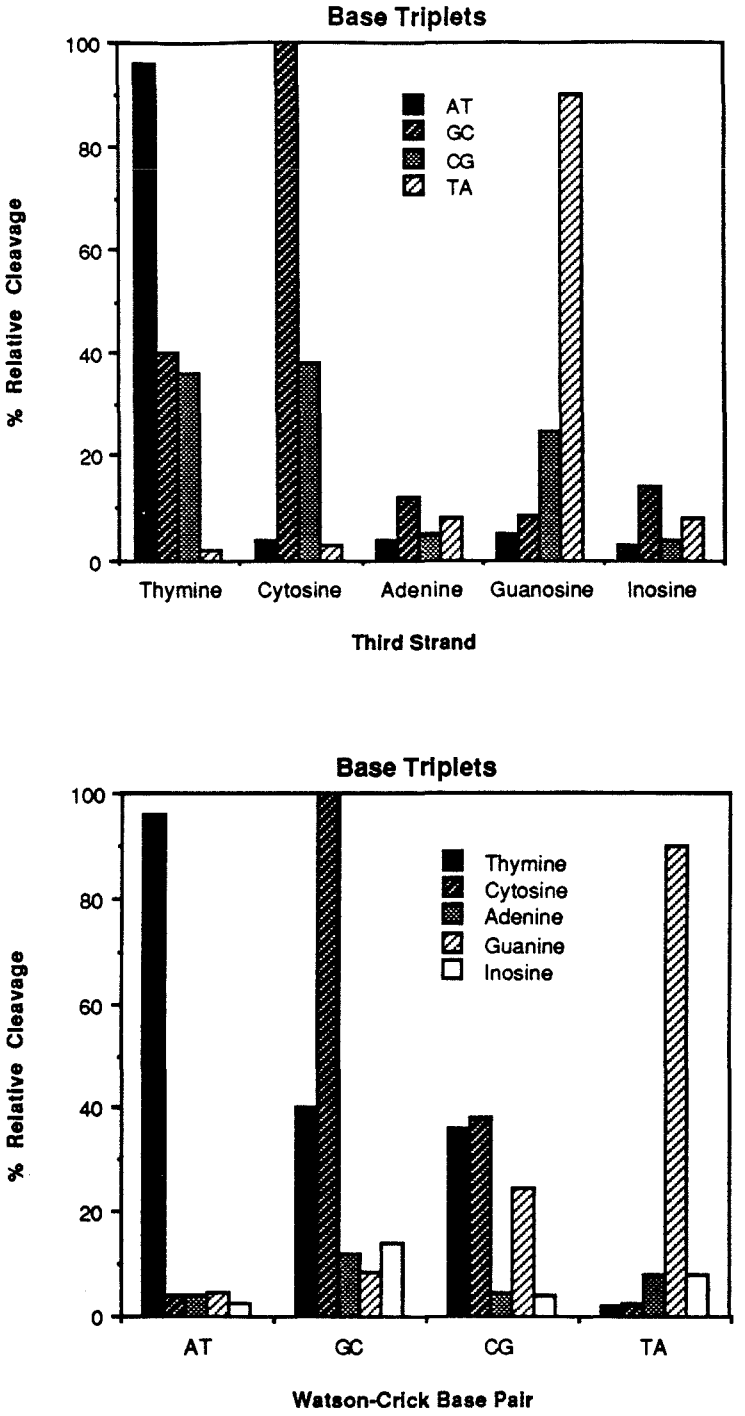
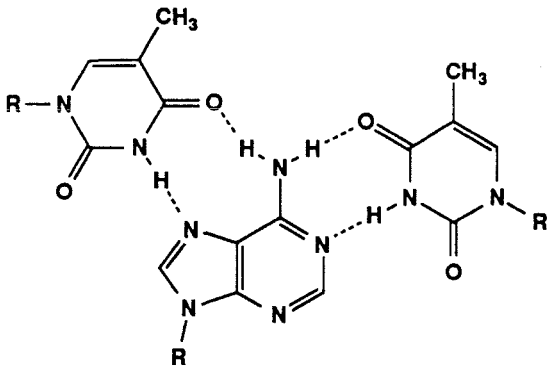
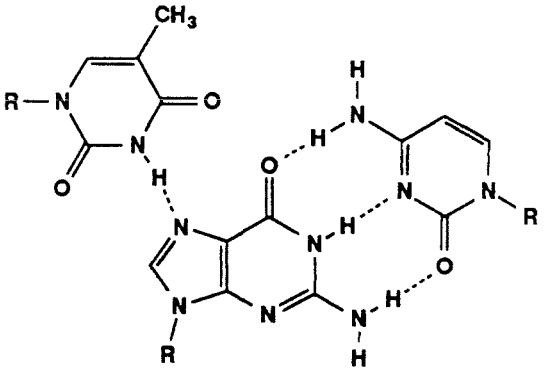


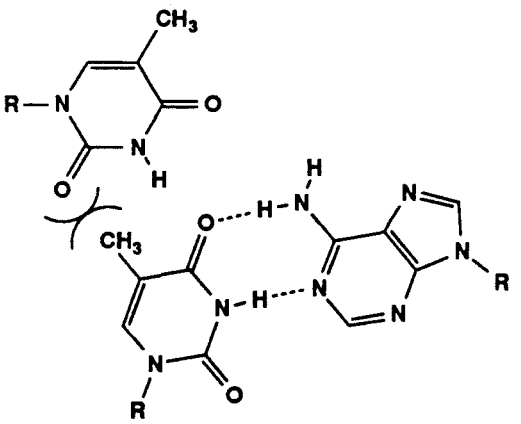
Figure 1C. Bar graphs presenting the average relative cleavage efficiencies obtained with oligonucleotide-EDTA•Fe(II) **1-5** with the four possible Watson-Crick base pairs contained in 30-mer duplexes (Fig. 1A, 1B). The values were determined by measuring the radioactivity at each cleavage site by scintillation counting and also densitometric analysis of the autoradiograms. The data are corrected for the background that resulted from the untreated 30-mer duplexes and are reproducible within ± 10% of reported values.



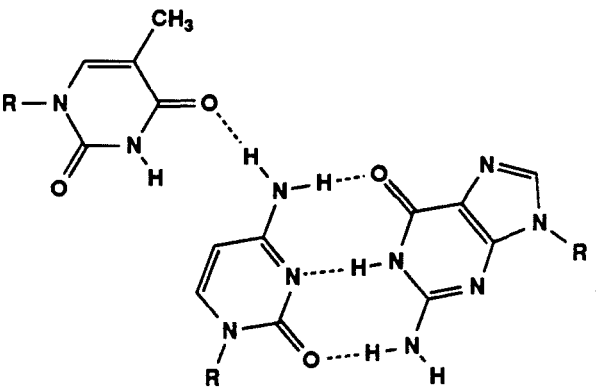
T•AT (96)



T•GC (40)

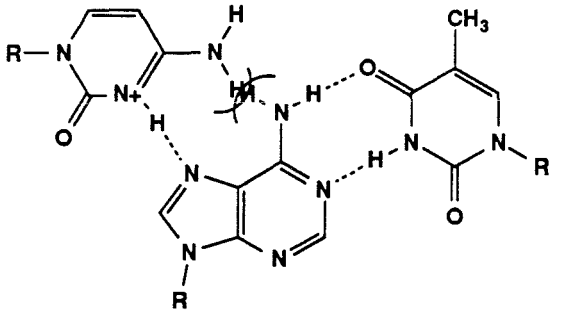


T•TA (36)

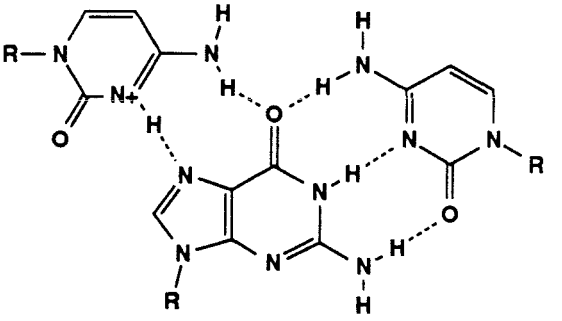


T•CG (2)

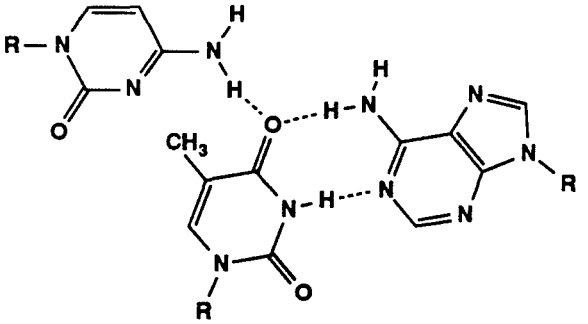
Figure 1D. Possible base triplets T•AT, T•GC, T•CT, T•TA. For each base triplet the positioning of the third base with respect to the Watson-Crick base pair is based upon forming possible hydrogen bonds, an *anti* conformation at the glycosidic link, and a backbone alignment close to the T•AT and C+GC base triplets.



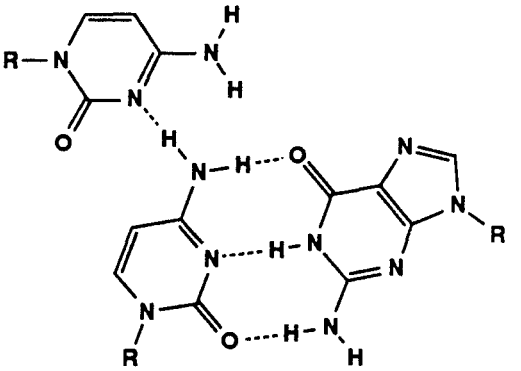
C+AT (4)



C+GC (100)

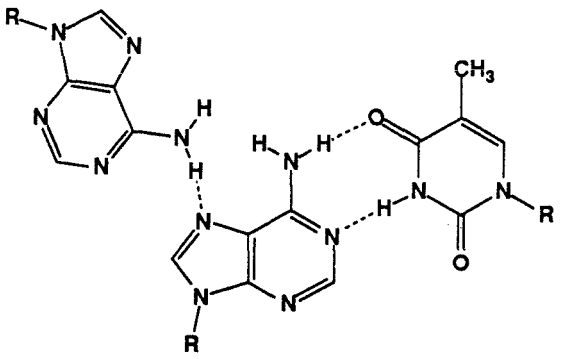


C•TA (3)

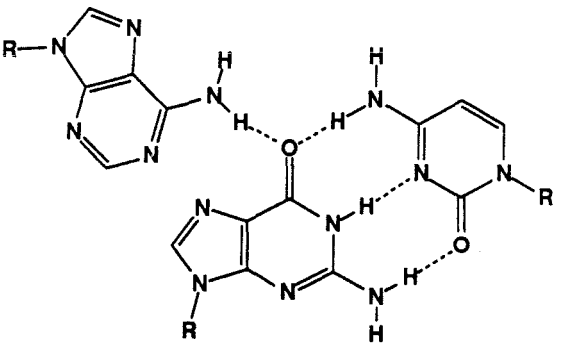


C•CG (38)

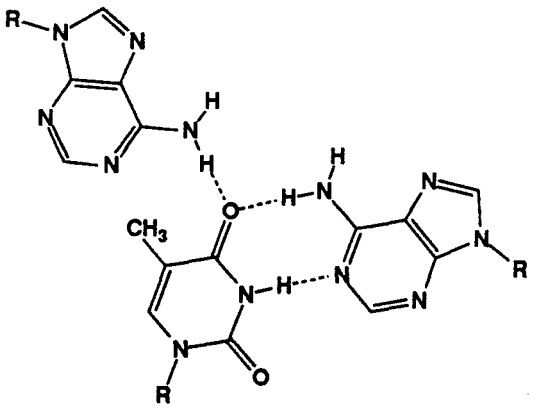
Figure 1E. Possible base triplets C^+AT, C^+GC, C^+CT, C^+TA. For each base triplet the positioning of the third base with respect to the Watson-Crick base pair is based upon forming possible hydrogen bonds, an *anti* conformation at the glycosidic link, and a backbone alignment close to the T•AT and C+GC base triplets.



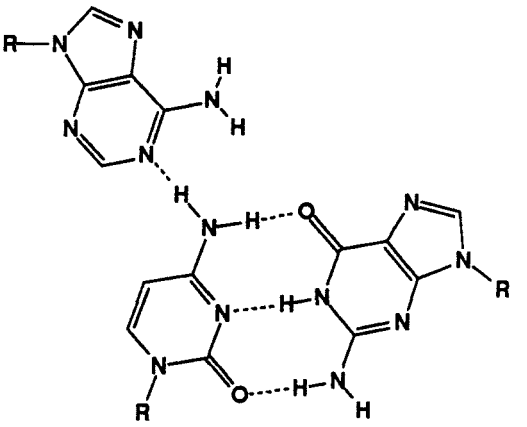
A•AT (4)



A•GC (12)

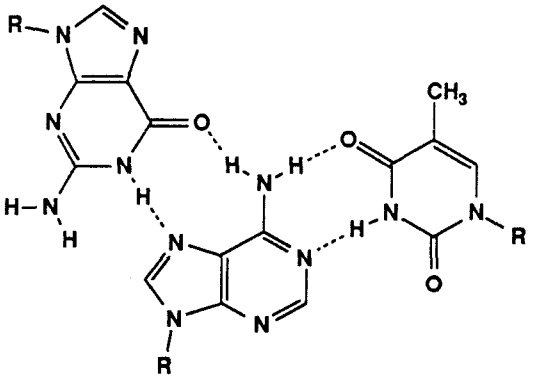


A•TA (5)

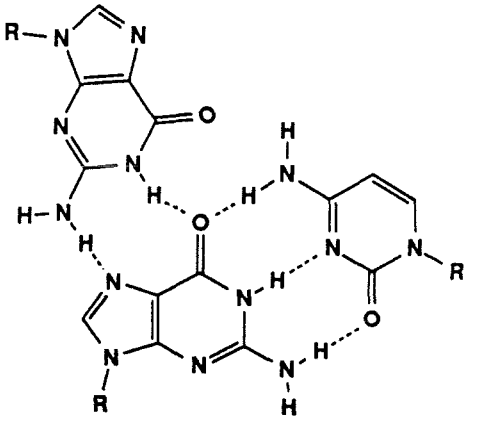


A•CG (8)

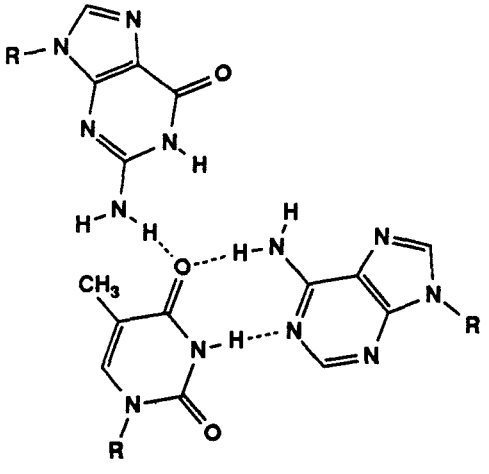
Figure 1F. Possible base triplets A•AT, A•GC, A•CT, A•TA. For each base triplet the positioning of the third base with respect to the Watson-Crick base pair is based upon forming possible hydrogen bonds, an *anti* conformation at the glycosidic link, and a backbone alignment close to the T•AT and C+GC base triplets.



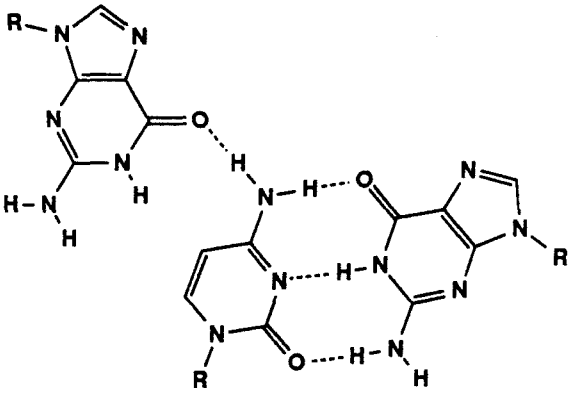
G•AT (5)



G•GC (9)

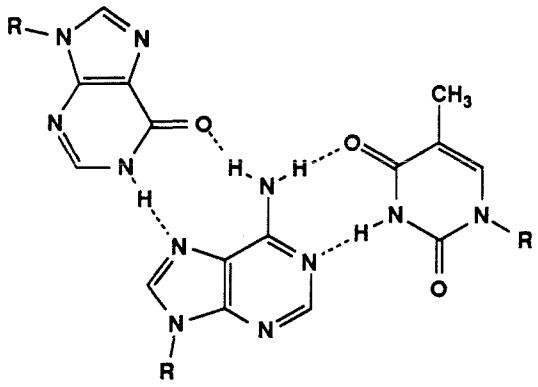


G•TA (90)

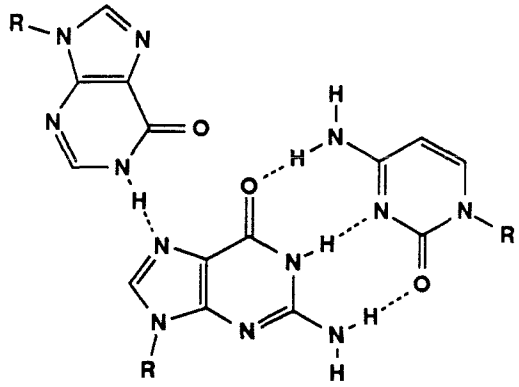


G•CG (25)

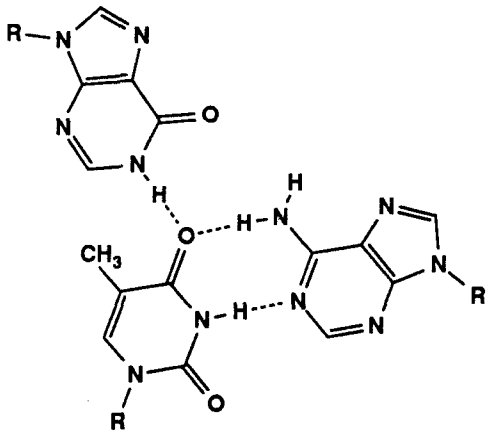
Figure 1G. Possible base triplets G•AT, G•GC, G•CT, G•TA. For each base triplet the positioning of the third base with respect to the Watson-Crick base pair is based upon forming possible hydrogen bonds, an *anti* conformation at the glycosidic link, and a backbone alignment close to the T•AT and C+GC base triplets.



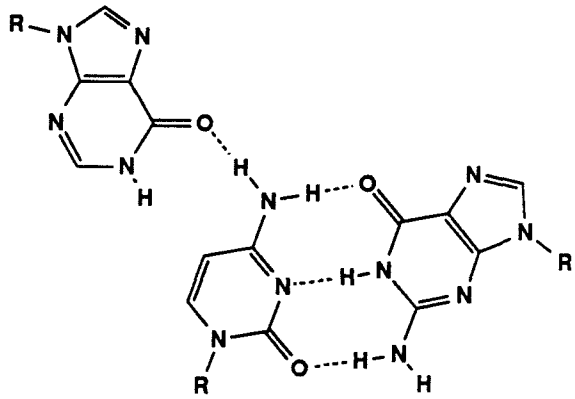
I•AT (3)



I•GC (14)

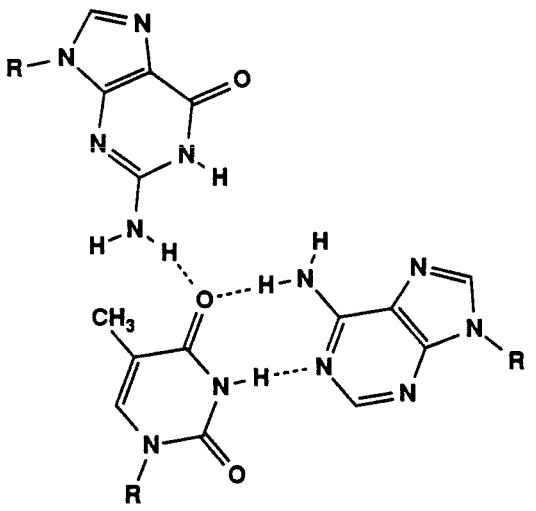


I•TA (8)

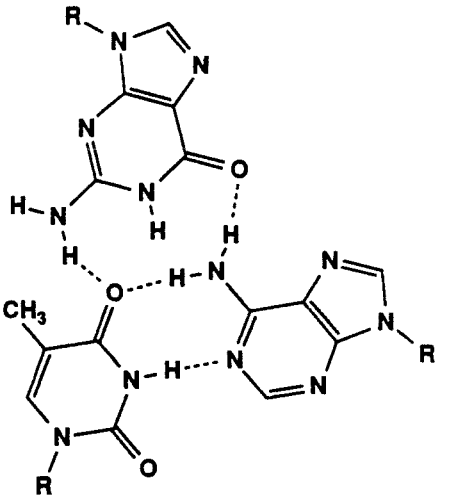


I•CG (4)

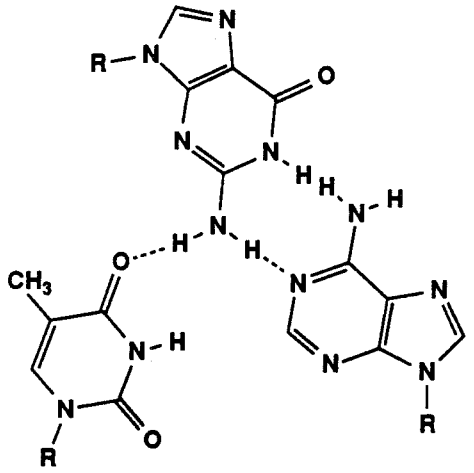
Figure 1H. Possible base triplets I•AT, I•GC, I•CT, I•TA. For each base triplet the positioning of the third base with respect to the Watson-Crick base pair is based upon forming possible hydrogen bonds, an *anti* conformation at the glycosidic link, and a backbone alignment close to the T•AT and C+GC base triplets.



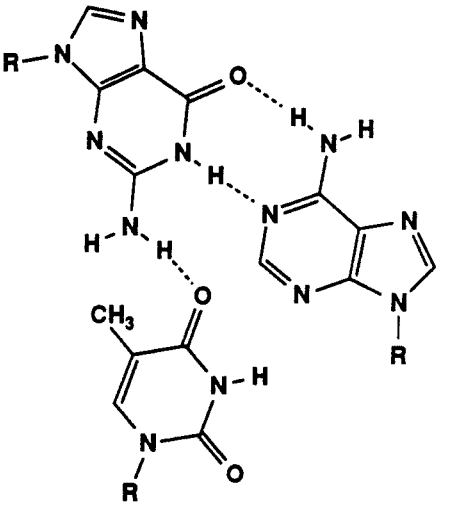
G•TA



G•TA



G•TA



G•TA

Figure 2. Simple hypothetical model for G•TA base triplet within a pyrimidine triple helix motif where G (N-2) is bound in the major groove by a hydrogen bond to T (O-4) of the Watson-Crick TA base pair. With more severe perturbation of the third strand DNA backbone or disruption of the Watson-Crick hydrogen bonds between T and A, additional hydrogen bonds can conceivably be formed between G and A.

To determine whether this specificity can be observed at mixed sequences within larger DNA fragments, cleavage of a 799 bp SV40 restriction fragment by the oligonucleotide-EDTA•Fe(II) series (**6-10**) (5'-T*T3C2T2T2ZTCT2C-3') was examined. This restriction fragment contains the 17-bp sequence d(AAAAGGAGGAATAGAAG), which represents a purine-rich site containing one pyrimidine (T). The cleavage efficiencies of oligonucleotides **6-10**, differing at one base position opposite the TA Watson-Crick base pair, were examined under conditions sensitive to the stability of the base triplet at the TA site (pH 6.6, 37°C). Oligonucleotide-EDTA **9**, but not **6**, **7**, **8**, or **10**, produced significant site specific cleavage on the 799-bp fragment (Fig. 3B). Oligonucleotide **9** contains G opposite the TA base pair. Site specific cleavage by oligonucleotides **6**, **7**, **8**, and **10** can be increased under less stringent binding conditions for triple helix fomation such as lower pH or temperature or added ethanol (4).

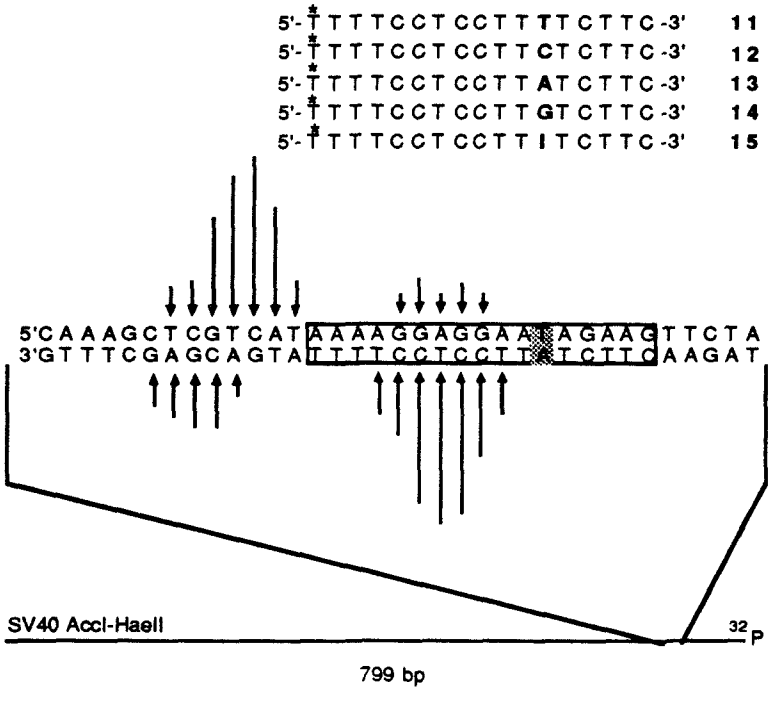


Figure 3A. (Above) Sequence of oligonucleotide-EDTA **11-15** where T* is the position of thymidine-EDTA. The oligonucleotides differ at one base position indicated by bold type. **(Below)** Histogram of the DNA cleavage patterns derived by densitometry of the autoradiogram in Figure 3B (lanes 6 and 13) from the cleavage of the SV40 AccI-HaeII restriction fragment with oligonucleotide-EDTA **9**.

Figure 3B. Autoradiogram of an 8 percent denaturing polyacrylamide gel with cleavage reactions on an SV40 restriction fragment. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (1 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled restriction fragment [~100 nM (bp) (~12,500±600 cpm)] in a solution of tris-acetate, pH 6.6 (50 mM), NaCl (100 mM), and calf thymus DNA [100 μM (bp)]. This was incubated at 0°C for 30 minutes and at 23°C for 30 minutes. Cleavage reactions were initiated by addition of DTT (3 mM) and allowed to proceed for 1.5 hours at 37°C. The reactions were stopped by precipitation with ethanol and the cleavage products were analyzed by gel electrophoresis (1200-2000 V, BPB 30 cm). (Lanes 1-7) 3' end-labeled AccI-HaeII restriction fragment of SV40; (lanes 8-14) 5' end-labeled restriction fragment. (Lanes 1 and 8) Maxam-Gilbert G sequencing reactions. (Lanes 2 and 9) Controls containing no oligonucleotide-EDTA•Fe(II). (Lanes 3-7 and 10-14) DNA cleavage products by oligonucleotide-EDTA•Fe(II) (**11-15**): **11** (Lanes 3 and 10); **12** (lanes 4 and 11); **13** (lanes 5 and 12); **14** (lanes 6 and 13); **15** (lanes 7 and 14). (See next page.)

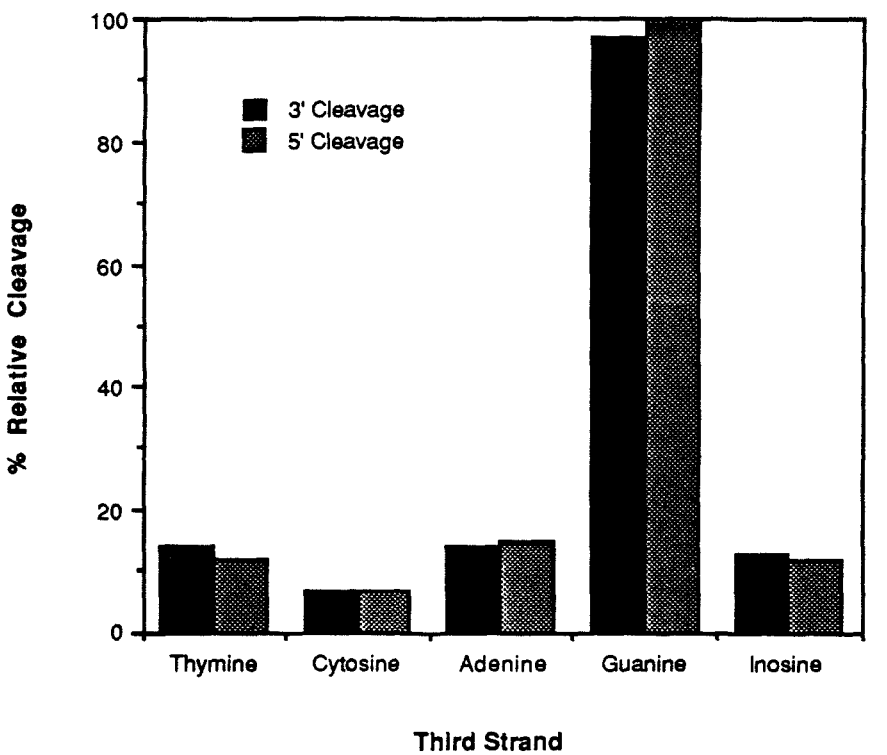
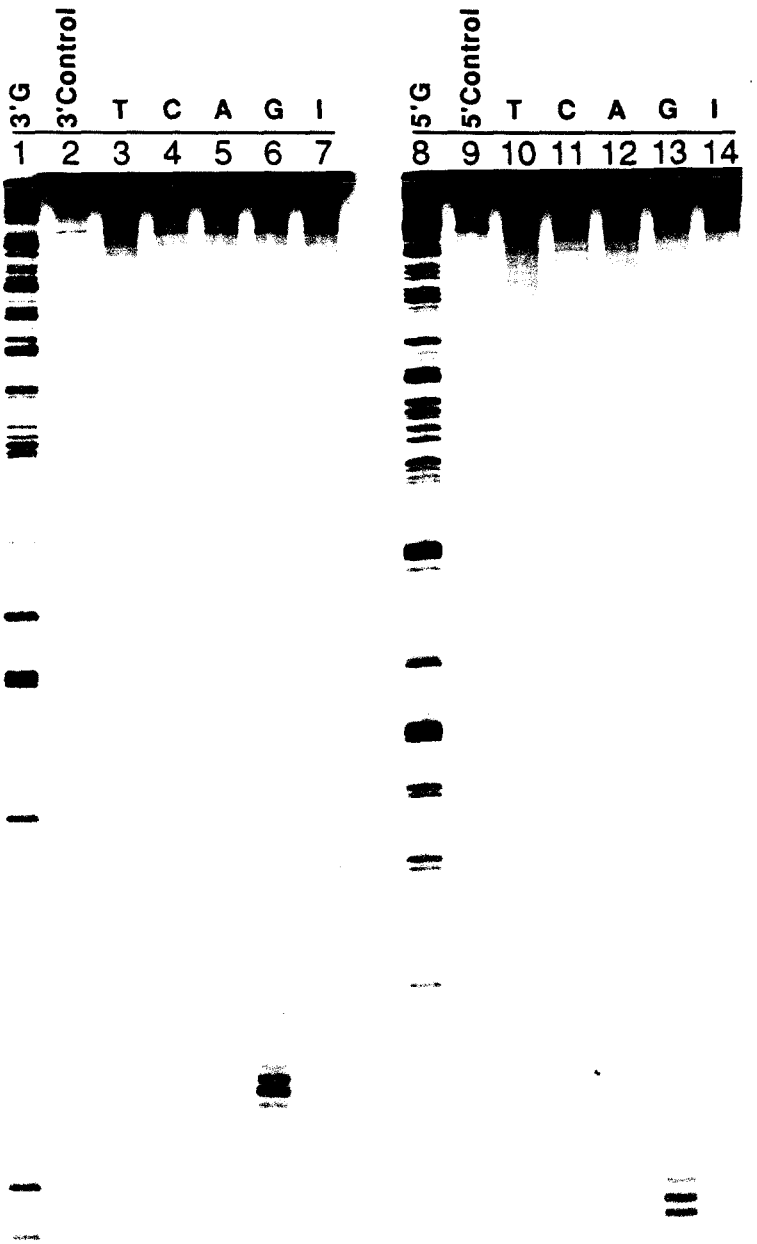


Figure 3C. Bar graph presenting the relative cleavage efficiencies obtained with oligonucleotide-EDTA•Fe(II) **11-15** with SV40 DNA (Fig. 3A, 3B). The values were determined by densitometric analysis of the autoradiogram in Fig. 3B. The data are corrected for the background that resulted from the untreated DNA and are reproducible within ± 10% of reported values.


3'G
3'Control
T C A G I
1 2 3 4 5 6 7
5'G
5'Control
T C A G I
8 9 10 11 12 13 14

The formation of G•TA triplets at DNA sites containing more than one T within a local purine tract, was tested by examining a site in the 3' long terminal repeat (LTR) of HIV DNA (Fig. 4A). The 17-bp sequence d(AGATAAGATAGAAGAGG) in a pHIV-CAT HindIII-XhoI restriction fragment is a mixed purine duplex target containing two thymines (29). DNA cleavage was observed by oligonucleotide-EDTA **11** (5'-T*CTGT$_2$CTGTCT$_2$CTC$_2$-3') at 0°C and 23°C (pH range 6.2 to 7.0) in the presence of 10% ethanol. Strand scission was observed only at the target site, with maximal cleavage occurring at pH 6.2, 0°C. For comparison, comparable cleavage was observed at pH 6.6, 37°C for oligonucleotide **9**, which recognizes a purine site of similar size with one pyrimidine (Fig. 2B).

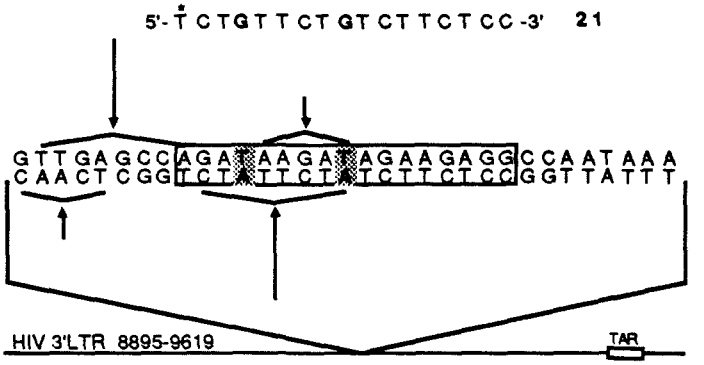
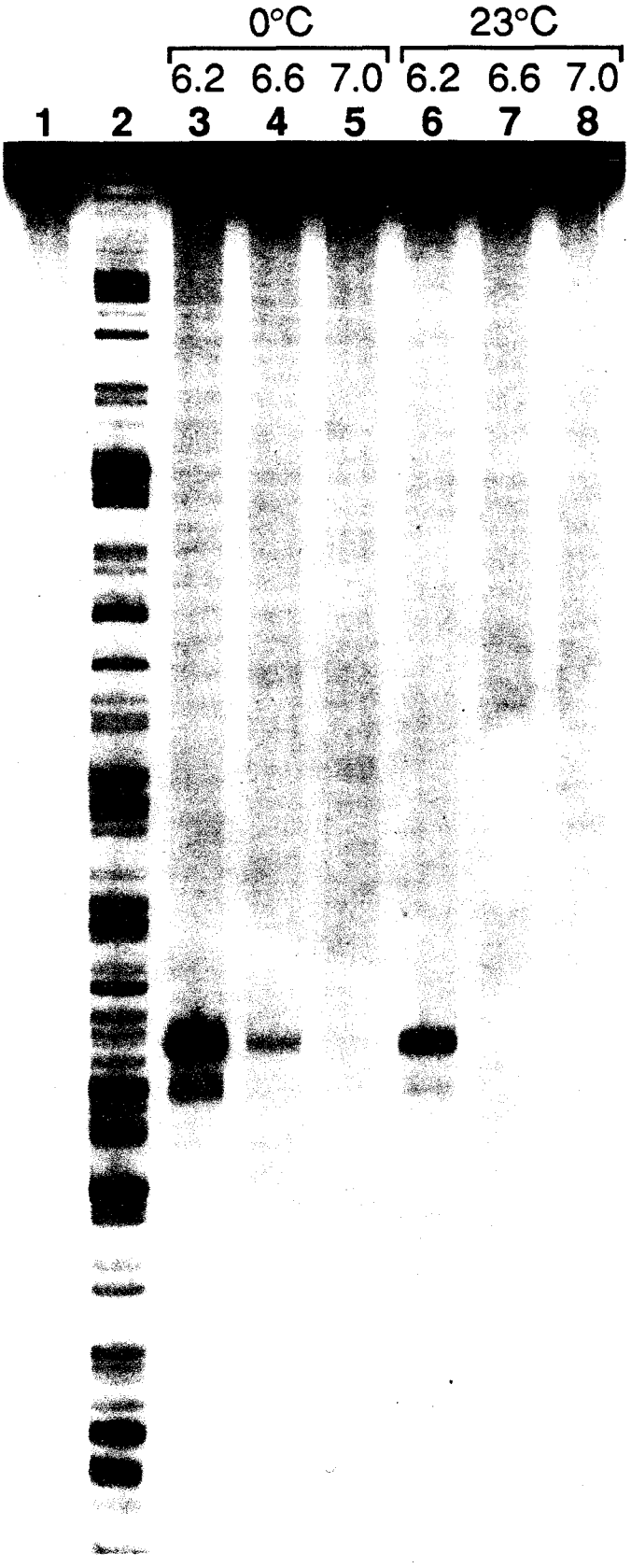


Figure 4A. (Above) Schematic of plasmid HIV-CAT, which contains the entire HIV promoter region on a XhoI-HindIII fragment derived from the 3' Long Terminal Repeat (LTR) of HIV DNA . Within the LTR is the regulatory element known as the TAR site. This element is necessary for transactivation by the trans-activator (tat) proteins that increase expression of viral mRNAs. **(Below)** Oligonucleotide-EDTA **16-20** synthesized to bind at or near the TAR site. Very weak cleavage was observed for oligonucleotide **17** (pH 6.2, 0°C, 40% EtOH).

Figure 4B. Autoradiogram of a 4 percent denaturing polyacrylamide gel with cleavage reactions on an HIV restriction fragment. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA **21** (1.67 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled restriction fragment [~100 nM (bp) (~8000±300 cpm)] in a solution of tris-acetate, pH 6.6 (50 mM), NaCl (100 mM), calf thymus DNA [100 μM (bp)], 10% ethanol, and incubated 1 hour at the reaction temperature. Cleavage reactions were initiated by addition of DTT (3mM) and allowed to proceed for 16.5 hours at 0°C or 23°C. The reactions were stopped by precipitation with ethanol and the resulting cleavage products were analyzed by gel electrophoresis (1600 V, 5 hr). (Lanes 1-8) 3' end-labeled HindIII-XhoI restriction fragment of pHIV-CAT; (Lanes 1) Control containing no oligonucleotide-EDTA•Fe(II). (Lane 2) Maxam-Gilbert G sequencing reaction. (Lanes 3-5) at 0°C, (Lanes 6-8) at 23°C, (Lanes 3 and 6) at pH 6.2, (Lanes 4 and 7) at pH 6.6, (Lanes 5 and 8) at pH 7.0.


0°C
23°C
6.2
6.6
7.0
6.2
6.6
7.0
1
2
3
4
5
6
7
8

The ability of oligonucleotide-EDTA•Fe(II) **11** to cause site-specific double strand breaks in HIV-CAT plasmid DNA is documented in Figure 5. pHIV-CAT was digested with BamHI to produce a 4.95-kbp fragment, which contained the 3' LTR of HIV with the site d(AGATAAGATAGAAAGAGG) located 1.54 kbp and 3.41 kbp from the ends.

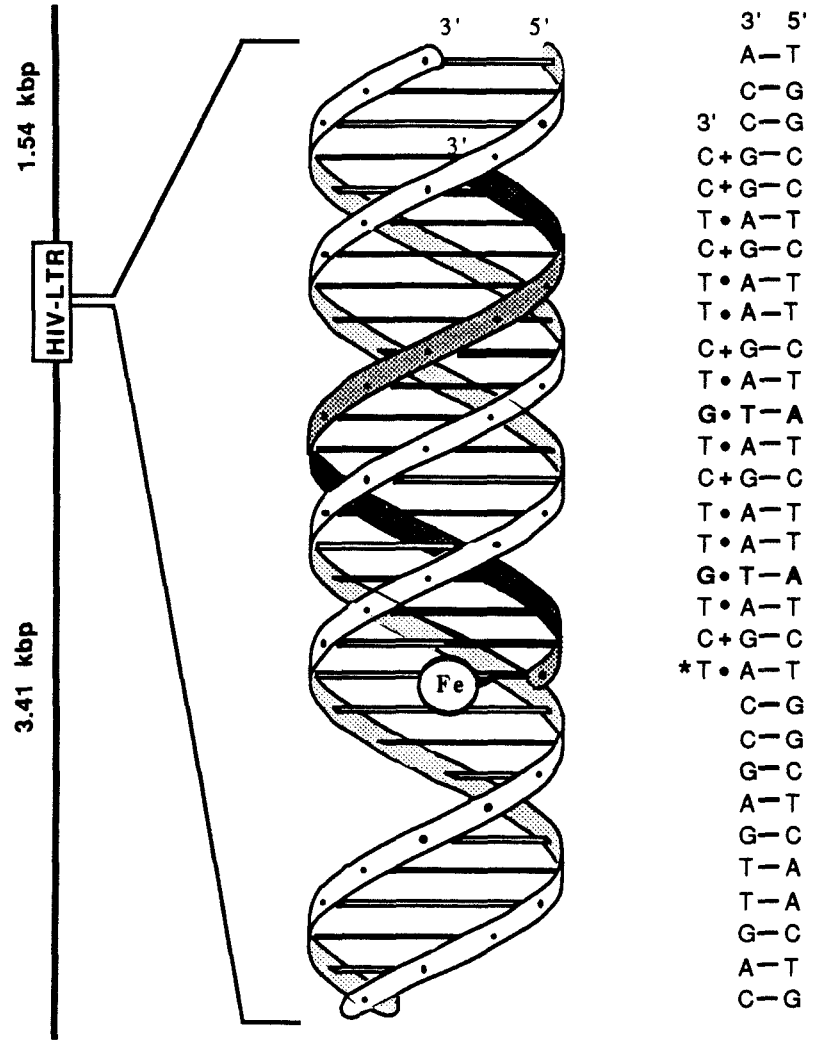


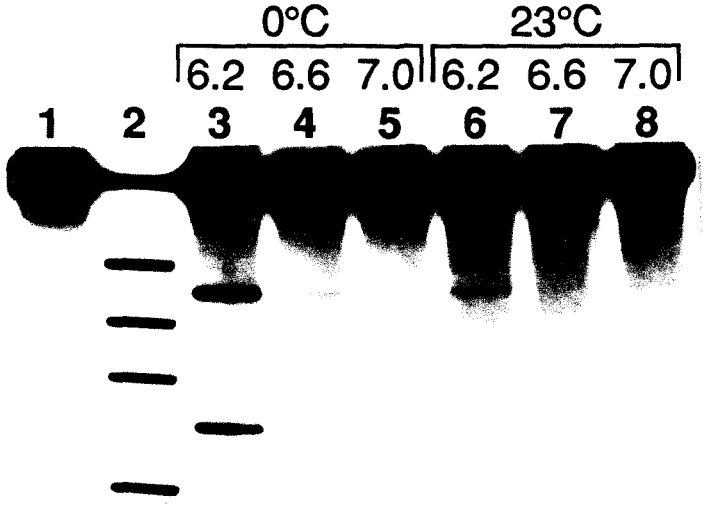
Figure 5A. (Left) The coarse resolution cleavage pattern from Figure 5B. **(Right)** Simplified model of the triple helix complex between the bound oligonucleotide-EDTA•Fe(II) **21** and a single site within the 4.95 kb plasmid DNA.

The ^{32}P end-labeled DNA was allowed to react with oligonucleotide-EDTA•Fe(II) in the presence of ascorbate at 0°C or 23°C (pH 6.2 to 7.0). Separation of the cleavage products by agarose gel electophoresis revealed one major cleavage site producing two DNA fragments, 1.54 and 3.41 kb in size (Fig.5). The HIV-DNA data indicate that an oligonucleotide binding a purine site with two thymines requires lower temperature and pH and the presence of ethanol. Nevertheless, under these less stringent binding conditions, additional cleavage sites do not appear. This result suggests a sequence composition limitation on the number of G residues within a pyrimidine oligonucletide that is tolerated for TA recognition (such as T>C>>G). Assuming an A'-like conformation (30), model-building studies suggest that a G•TA triplet within a pyrimidine triple helix motif displaces the position of the deoxyribosyl group and hence the phosphodiester backbone in the third strand,which is likely energetically unfavorable.

Although the G•TA triplet within a pyrimidine oligonucleotide extends, in a formal sense, triple helix specificity to three of the four possible base pairs of double helical DNA, some limitations on sequence composition are likely. It is not clear how far one can deviate from hompurine•homopyrimidine target sequences and still obtain triple helix formation. Rather, this result may provide structural leads for the design of deoxyribonucleosides (or their analogs) with nononatural heterocycles directed toward a more general solution. With

---

Figure 5B. Autoradiogram of double strand cleavage of pHIV-CAT DNA (4.95 kbp) analyzed on a 1 percent agarose gel. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA **21** (1 μM), spermine (1 mM), and Fe(II) (1.25 μM) with the ^{32}P labeled linearized plasmid [~100 nM (bp) (~27,250±250 cpm)] in a solution of tris-acetate (50 mM), NaCl (100 mM), calf thymus DNA [100 μM (bp)], 10% ethanol, and incubated 1 hour at the reaction temperature. Cleavage reactions were initiated by addition of ascorbate (1 mM) and allowed to proceed for 18 hours at 0°C or 23°C. The reactions were stopped by precipitation with ethanol and the cleavage products were analyzed by gel electrophoresis (155 V, BPB run near the bottom of the gel). (Lanes 1-8) pHIV-CAT linearized with BamHI and 3' end-labeled at both ends. (Lanes 1) Control containing no oligonucleotide-EDTA•Fe(II). (Lane 2) DNA size markers obtained by digestion of BamHI linearized pHIV-CAT with HindIII and XhoI: 4950 (undigested DNA), 3725, 3003, 1947, 1225. Lanes 3-8) DNA cleavage products produced by oligonucleotide-EDTA•Fe(II) **21**. (Lanes 3-5) at 0°C, (Lanes 6-8) at 23°C, (Lanes 3 and 6) at pH 6.2, (Lanes 4 and 7) at pH 6.6, (Lanes 5 and 8) at pH 7.0.


0°C
23°C
6.2 6.6 7.0 6.2 6.6 7.0
1 2 3 4 5 6 7 8

regard to the putative purine•purine•pyrimidine motif (A•AT and G•GC triplets) (21, 25, 26), this data suggests that there may be distinct conformational families of intermolecular triple helices (such as hompyrimidine vs homopurine donor third strand), which are not mutually compatible. The specificity of base triplets may differ for each structural motif due to different alignment of the deoxyribosyl-phosphodiester backbone and hence the heterocyclic bases along the major groove, which would create different optimal hydrogen bonding patterns. For example, from the data presented here, within a primidine rich oligonucleotide, G prefers to bind TA base pairs, whereas within a purine rich oligonucleotide, G apparently binds GC base pairs.

References and Notes

1. P.B. Dervan, *Science* **232**, 464 (1986).
2. P.B. Dervan in *Nucleic Acids and Molecular Biology*, F. Eckstein and D.M.J. Lilley, Ed. (Springer-Verlag, Heidelberg, 1988) vol. 2, pp. 49-64.
3. J.K. Barton, *Science* **233**, 727 (1986).
4. H.E. Moser and P.B. Dervan, *ibid.* **238**, 645 (1987).
5. V.I. Lyamichev, *et al. Nucleic Acids Res.* **16**, 2165 (1988).
6. S.A. Strobel, H.E. Moser, P.B. Dervan, *J. Am. Chem. Soc.* **110**, 7927 (1988).
7. T.J. Povsic and P.B. Dervan, *ibid.* **111**, 3059 (1989).
8. L.J. Maher, B.J. Wold, P.B. Dervan, *Science* in press (1989).
9. For α-oligodeoxyribonucleotides see: a) T.L. Doan, *et al. Nucleic Acids Res.* **15**, 7749 (1987). b) D. Praseuth, *et al. Proc. Natl. Acad. Sci. USA* **85**, 1349 (1988).
10. G. Felsenfeld, D.R. Davies, A. Rich, *J. Am. Chem. Soc.* **79**, 2023 (1957).
11. K. Hoogsteen, *Acta Cryst.* **12**, 822 (1959).
12. A.M. Michelson, J. Massouilie, W. Guschlbauer, *Prog. Nucl. Acid Res. Mol. Biol.* **6**, 83, (1967).

13. G. Felsenfeld and H.T. Miles, *Annu. Rev. Biochem.* **36**, 407 (1967).

14. M.N. Lipsett, *Biochem. Biophys. Res. Commun.* **11**, 224 (1963). M.N. Lipsett, *J. Biol. Chem.* **239**, 1256 (1964).

15. F.B. Howard, J. Frazier, M.N. Lipsett, H.T. Miles, *Biochem. Biophys. Res. Commun.* **17**, 93 (1964).

16. J.H. Miller and H.M. Sobell, *Proc. Natl Acad. Sci. USA* **55**, 1201 (1966).

17. A.R. Morgan and R.D. Wells, *J. Mol. Biol.* **37**, 63 (1968).

18. J.S. Lee, D.A. Johnson, A.R. Morgan, *Nucleic Acid Res.* **6**, 3073 (1979).

19. S.M. Mirkin, *et al. Nature* **330**, 495 (1987).

20. H. Htun and J.E. Dahlberg, *Science* **241**, 1791 (1988).

21. Y. Kohwi and T. Kohwi-Shigematsu, *Proc. Nat. Acad. Sci. USA* **85**, 3781 (1988).

22. J.C. Hanvey, *et al.*, *ibid* **85**, 6292 (1988). J.C. Hanvey, *et al. J. Biol. Chem.* **263**, 7386 (1988).

23. P. Rajagopal and J. Feigon, *Nature* **239**, 637 (1989).

24. C. Marck and D. Thiele, *Nucleic Acids Res.* **5**, 1017 (1978).

25. S.L. Broitman, D.D. Im, J.R. Fresco, *Proc. Natl. Acad. Sci. USA* **84**, 5120 (1987). A.G. Letai, *et al.*, *Biochemistry* **27**, 9108 (1988).

26. M. Cooney, *et al.*, *Science* **241**, 456 (1988).

27. R.P. Hertzberg and P.B. Dervan, *Biochemistry* **23**, 3934 (1984).

28. G.B. Dreyer and P.B. Dervan, *Proc. Natl. Acad. Sci. USA* **82**, 963 (1985).

29. S.S. Tavale and H.M. Sobell, *J. Mol. Biol.* **48**, 109 (1970). F. Jordan and H. Niv, *Biochim. Biophys. Acta* **476**, 265 (1977).

30. S. Arnott, P. J. Bond, E. Selsing., P.J.C. Smith, *Nucleic Acids Res.* **3**, 2459 (1976).

31. The pH values are not corrected for temperature or presence of ethanol and are given for the tenfold concentrated buffer solutions at 25°C.

32. T. Maniatis, E.F. Fritsch, J. Sambrook, *Molecular Cloning, a laboratory manual*, Cold Spring Harbor Laboratory (1982).

33. A.M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).

34. We are grateful to the National Institutes of Health, the Office of Naval Research and the Parsons Foundation for generous support. We thank Scott Singleton for valuable assistance in model building, Dr. Alan Frankel for the gift of pHIV-CAT, and Tom Povsic for large scale preparation.

Appendix
Chapter 2

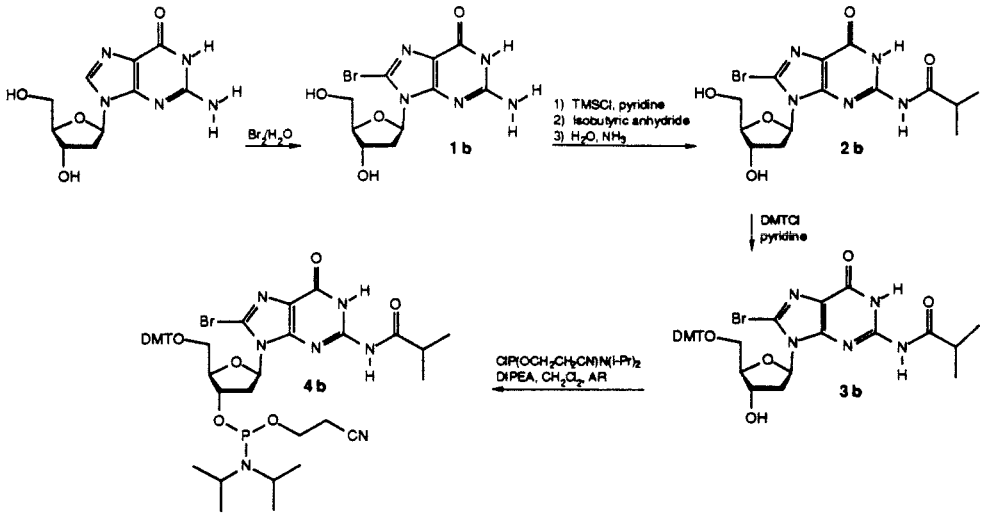


Figure 1A. Scheme for the synthesis of DMT protected 8-Bromo-2-deoxyguanosine phosphoramidite (8-Bromo-2-deoxyguanosine, T. Lin, J. Cheng, K. Ishiguro, and A.C. Sartorelli, *J. Med. Chem.* **1985**, *28*, 1194-1198.).

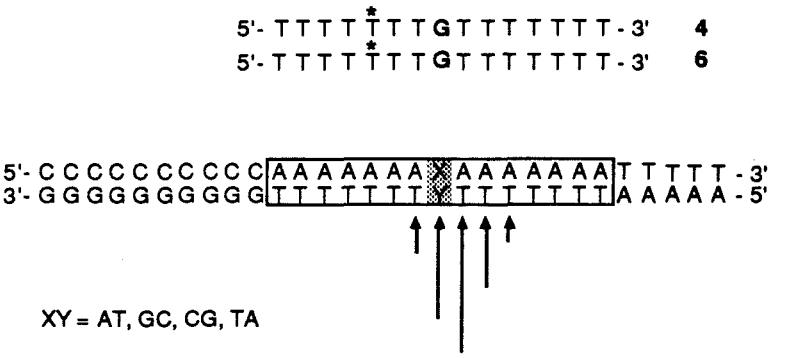
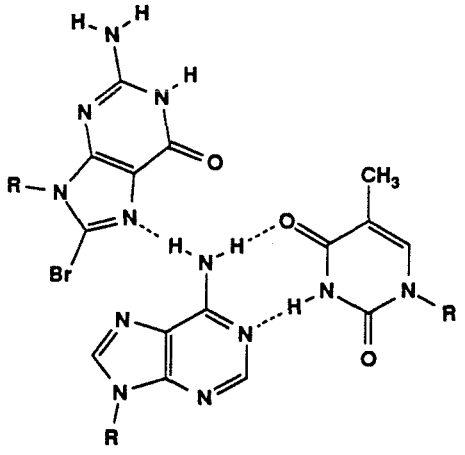


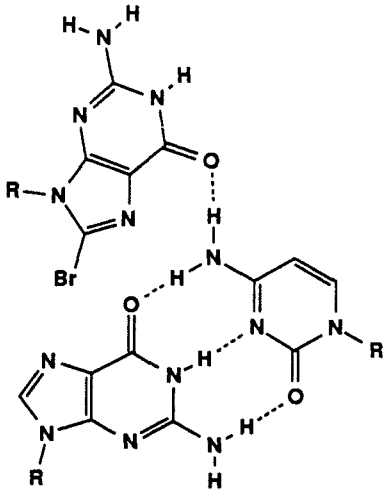
Figure 1B. (Above) Sequences of oligonucleotide-EDTA **4** and **6** where T* is the position of the thymidine-EDTA. The oligonucleotides differ at the central base position where **G** is guanine and **𝔾** is bromoguanine. **(Below)** Histogram of the DNA cleavage patterns derived by densitometry of the autoradiogram shown in Figure 1C (lane 18, oligonucleotide **4**). The heights of the arrows represent the relative cleavage intensity at the indicated bases. The box indicates the double stranded sequence bound by oligonucleotide-EDTA•Fe(II) **4**. The Watson-Crick base pair (AT, GC, CG, or TA) opposite the variant base in the oligonucleotide is shaded.

Figure 1C. Autoradiogram of the 20 percent denaturing polyacrylamide gel. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled 30-mer duplex [~0.5 μM (bp) (~3,400±300 cpm)] in a solution of tris-acetate, pH 6.6 (25 mM), NaCl (100 mM), calf thymus DNA [100 μM (bp)], and 35% ethanol and incubated at 10°C for 1 hr. Cleavage reactions were initiated by addition of DTT (3mM) and allowed to proceed for 4.5 hours at 23°C. The reactions were stopped by freezing and lyophilization and the cleavage products were analyzed by gel electrophoresis (1200-2000 V, BPB 23 cm). (Lanes 1-8) 5' end-labeled d($A_5T_7YT_7G_{10}$) DNA cleavage products produced by oligonucleotide-EDTA•Fe(II) **4** and **6**; **4** (Lanes 1-4); **6** (lanes 5-8). **XY=AT** (Lanes 1 and 5); **XY=GC** (lanes 2 and 6); **XY=CG** (lanes 3 and 7); **XY=TA** (lanes 4 and 8).

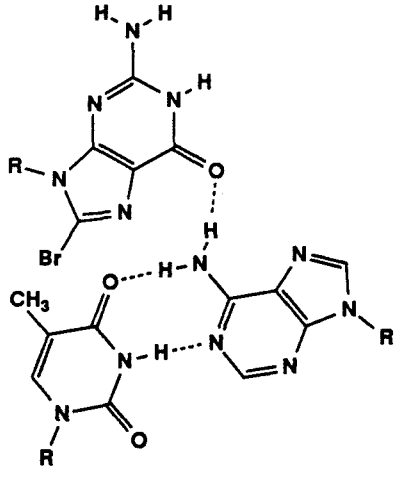

G
BrG
1 2 3 4 5 6 7 8



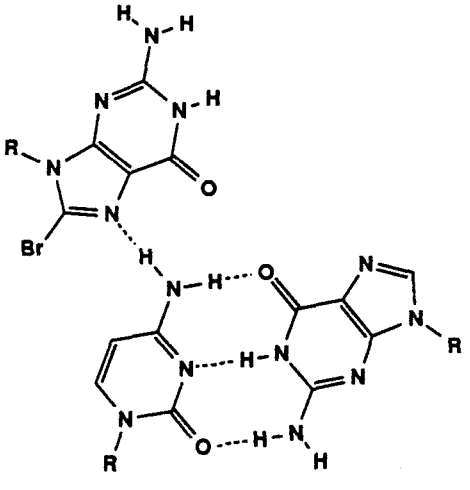
BrG•AT



BrG•GC



BrG•TA



BrG•CG

Figure 1D. Possible base triplets **BrG•AT**, **BrG•GC**, **BrG•CT**, **BrG•TA**. For each base triplet the positioning of the third base with respect to the Watson-Crick base pair is based on forming possible hydrogen bonds, a *syn* conformation at the glycosidic link, and a backbone alignment close to the T•AT and C+GC base triplets.

Figure 2A. (Above) Sequences of oligonucleotide-EDTA **7-10** where T* is the position of the thymidine-EDTA. **(Below)** Histogram of the DNA cleavage patterns derived by densitometry of the autoradiogram shown in Figure 2B (lane 1, oligonucleotide **7**). The heights of the arrows represent the relative cleavage intensity at the indicated bases. The box indicates the double stranded sequence bound by oligonucleotide-EDTA•Fe(II) **7** and **8**.

Figure 2B. Autoradiogram of the 20 percent denaturing polyacrylamide gel. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled 30-mer duplex [~0.5 μM (bp) (~4400±150 cpm)] in a solution of tris-acetate, pH 7.0 (25 mM), NaCl (100 mM), calf thymus DNA [100 μM (bp)], and 40% ethanol and incubated at 21°C for 1 hr. Cleavage reactions were initiated by addition of DTT (3mM) and allowed to proceed for 12 hours at 21°C. The reactions were stopped by freezing and lyophilization and the cleavage products were analyzed by gel electrophoresis (1200-2000 V, BPB 23 cm). (Lanes 1-4) 5' end-labeled d($A_5T_7YT_7G_{10}$) DNA cleavage products produced by oligonucleotide-EDTA•Fe(II) **7-10**; **7** (Lanes 1); **8** (lanes 2); **9** (Lanes 3); **10** (lanes 4).

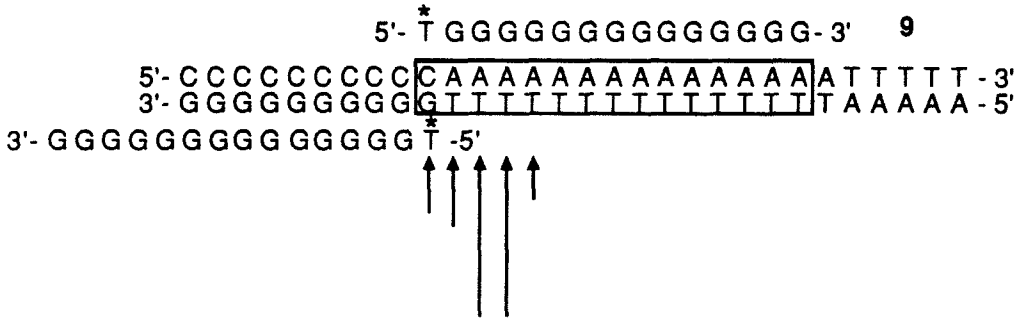
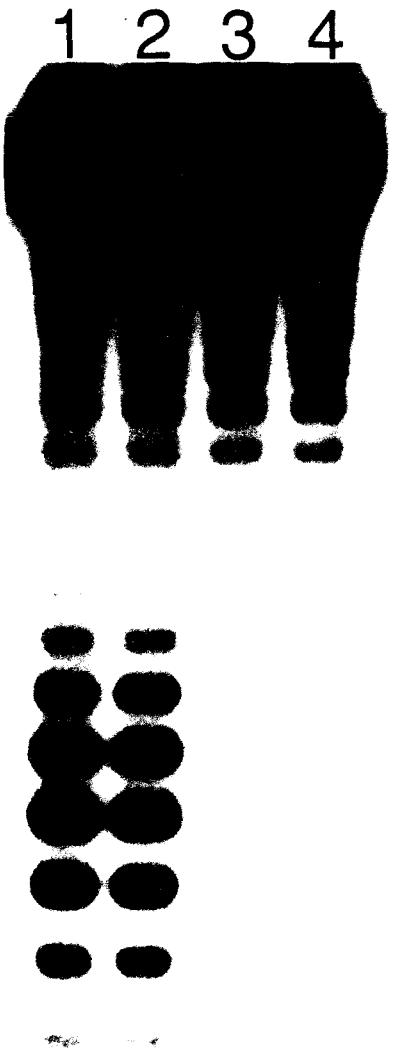


Figure 2C. (Above) Sequence of oligonucleotide-EDTA **9** where T* is the position of the thymidine-EDTA. **(Below)** Histogram of the DNA cleavage patterns derived by densitometry of the autoradiogram shown in Figure 2.19 (lane 3, oligonucleotide **9**). The heights of the arrows represent the relative cleavage intensity at the indicated bases. The box indicates the double stranded sequence bound by oligonucleotide-EDTA•Fe(II) **9** (note that it is shifted one base pair from the sequence bound by oligonucleotide-EDTA•Fe(II) **7** and **8**). It is possible that the cleavage observed is the result of G recgntion of G (G•GC base triplets).


1 2 3 4

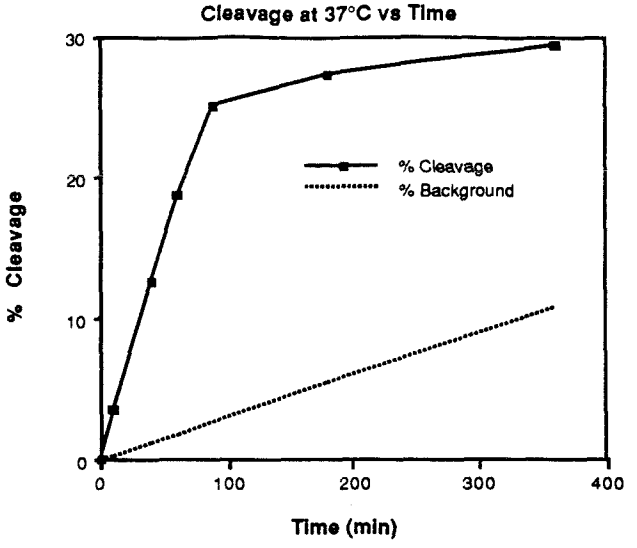


Figure 3. Graphical representation of percent cleavage at 37°C vs time. Data derived by densitometry and scintillation counting. % Cleavage was calculated as the ratio of radioactivity of oligonucleotide directed cleavage to the total radioactivity in a gel lane. Background cleavage increases linearly with time.

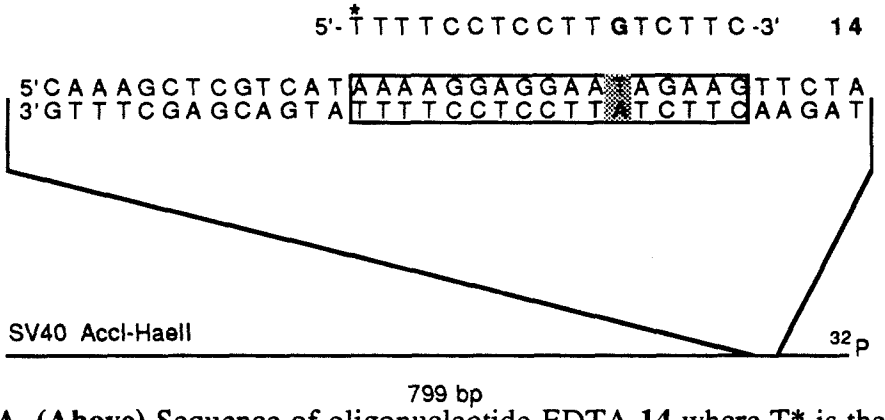
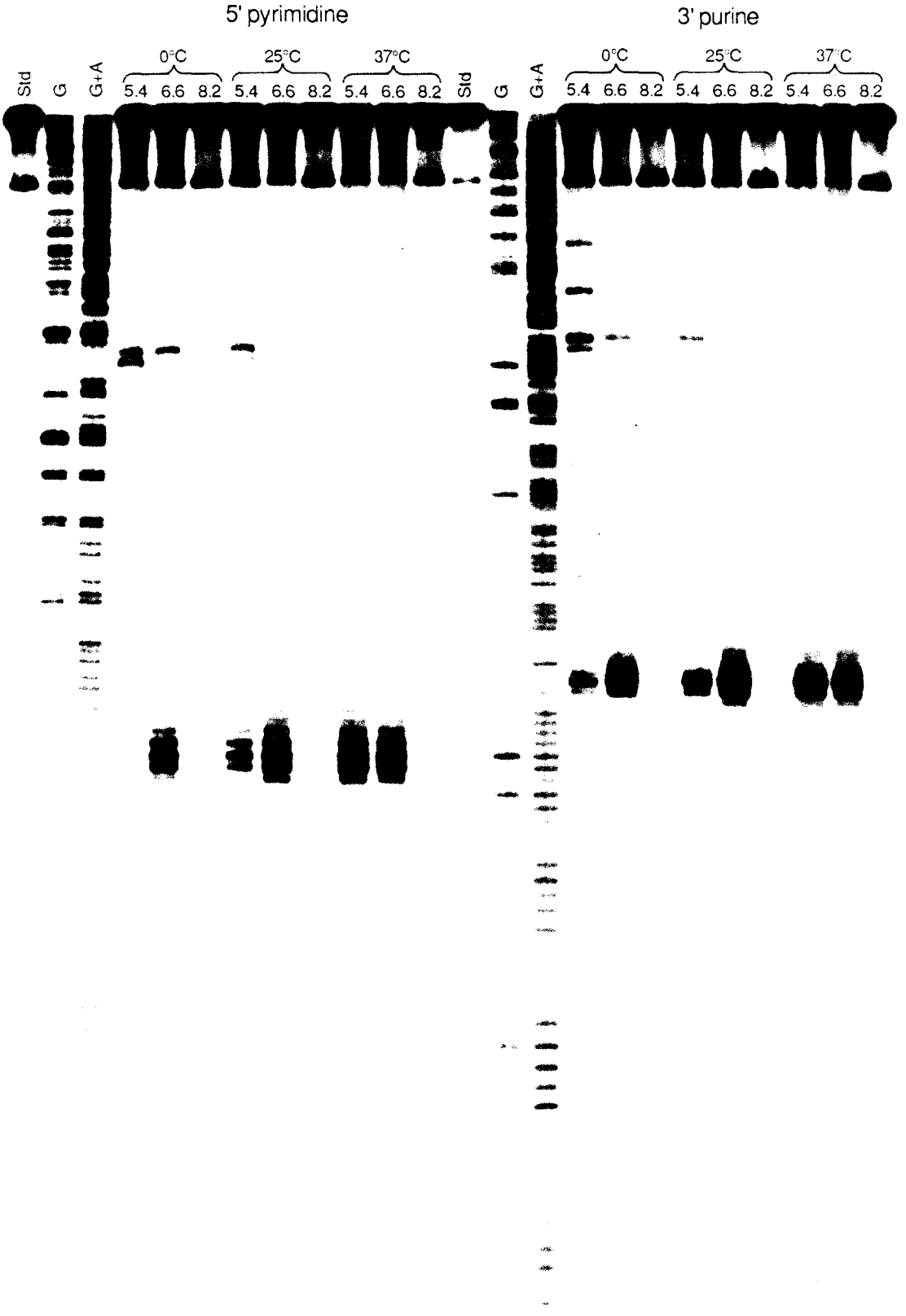


Figure 4A. (Above) Sequence of oligonucleotide-EDTA **14** where T* is the position of thymidine-EDTA . (Below) The target site in the 799 bp SV40 AccI-HaeII restriction fragment. The box indicates the double stranded sequence bound by oligonucleotide-EDTA•Fe(II) **14**.

Figure 4B. Autoradiogram of the 10 percent denaturing polyacrylamide gel. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled restriction fragment [~100 nM (bp) (~3,500±300 cpm)] in a solution of tris-acetate, pH 5.4, 6.6, or 8.2 (50 mM), NaCl (100 mM), calf thymus DNA [100 μM (bp)], and 30% ethanol. This was incubated at 0°C for 45 minutes. Cleavage reactions were initiated by addition of DTT (2 mM) and allowed to proceed for 16 hours at 0°C, 25°C, or 37°C. The reactions were stopped by precipitation with ethanol and the cleavage products were analyzed by gel electrophoresis (1200-2000 V, BPB 27 cm). **(Left)** 5' end-labeled AccI-HaeII restriction fragment of SV40; **(Right)** 3' end-labeled restriction fragment.

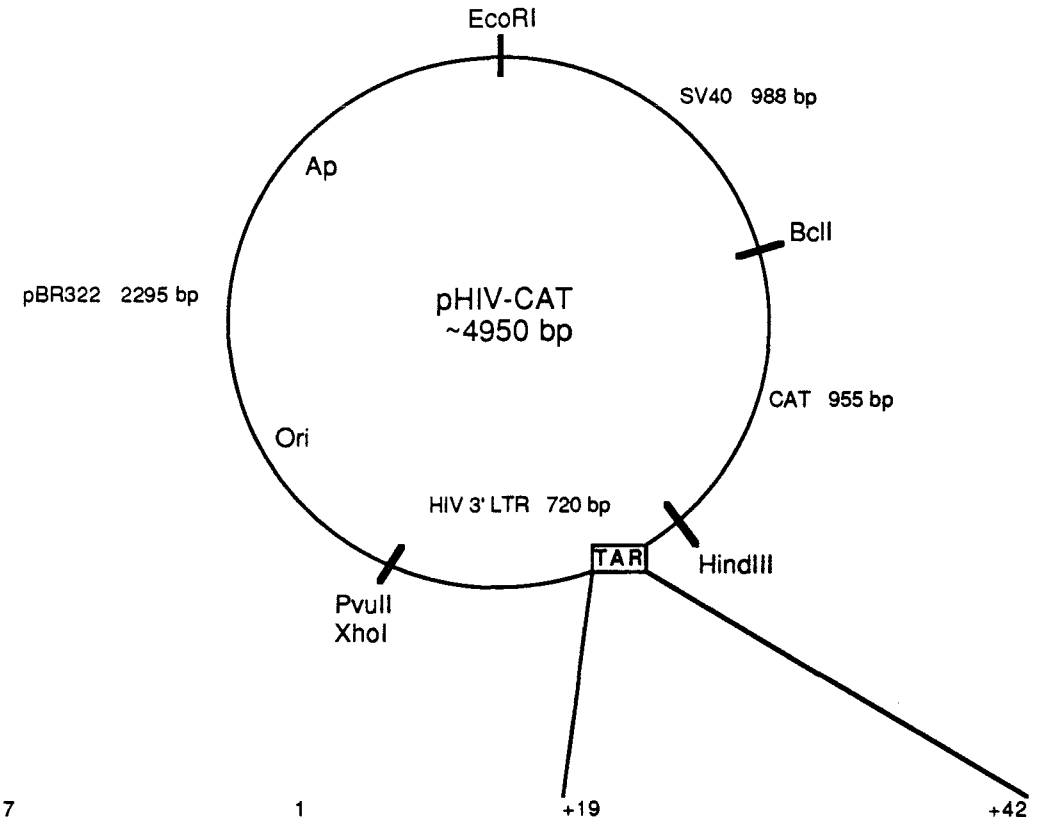
SV40 Acc I - Hae II fragment



SV40

1430	1440	1450	1460
ATCAAAGAAC	TGCTCCTCAG	TGGATGTTGC	CTTTACTTCT
1470	1480	1490	1500
AGGCCTCTAC	GGAAGTGTTA	CTTCTGCTCT	AAAAGCTTAT
1510	1520	1530	1540
GAAGATGGCC	CCAACAAAAA	GAAAAGGAAG	TTGTCCAGGG
1550	1560	1570	1580
GCAGCTCCCA	AAAAACCAAA	GGAACCAGTG	CAAGTGCCAA
1590	1600	1610	1620
AGCTCGTCAT	**AAAAGGAGGA**	**ATAGAAG**TTC	TAGGAGTTAA
1630			
AACTGGAGTA			

Figure 4C. SV40 DNA 1421-1630. Under less stringent conditions (ethanol, lower pH, lower temperature) partially homologous sites (in outline) are cleaved by oligonucleotide-EDTA **14** whose target site is indicated in bold.



-17 1 +19 +42

5'-GCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGC-3'

TAR

3'-CTTTTT GCCT GTG CT * -5' **16**

3'-CTTTTT GCCT GT GCT GGG TCTCTCT * -5' **17**

3'-TCTCTCT GGTT G GGCC GGG TCT * -5' **18**

19 3'-CC GGG TCT GGG CCT GGGGG CTCTCT * -5'

20 5'- T *CT GGG CTC GGG CCCTC GGGGGG CC-3'

Figure 5. (Above) Schematic of plasmid HIV-CAT, which contains the entire HIV promoter region on a XhoI-HindIII fragment derived from the 3' Long Terminal Repeat (LTR) of HIV DNA . Within the LTR is the regulatory element known as the TAR site. This element is necessary for transactivation by the trans-activator (tat) proteins that increase expression of viral mRNAs. **(Below)** Oligonucleotide-EDTA **16-20** synthesized to bind at or near the TAR site. Very weak cleavage was observed for oligonucleotide **17** (pH 6.2, 0°C, 40% EtOH).

Chapter 3

Recognition of all Four Base Pairs of Duplex DNA by Triple Helix Formation • Design of Pyrimidine Specific Bases

The sequence-specific recognition of double helical DNA is essential for the regulation of cellular functions including transcription, replication, and cell division. The ability to design synthetic molecules that bind specifically to unique sequences on human DNA could have major implications for the treatment of genetic, neoplastic, and viral diseases (1-3).

Triple Helix Formation. Oligonucleotide recognition by triple helix formation offers a powerful chemical approach for the sequence-specific binding of double helical DNA. In the pyrimidine-Hoogsteen triple strand model, a binding site size of >15 homopurine base pairs affords >30 discrete sequence-specific hydrogen bonds to duplex DNA. An oligonucleotide 15 nucleotides (nt) in length could conceivably bind uniquely to one site on a human chromosome. Binding specificity is due to Hoogsteen hydrogen bonding, wherein thymine (T) recognizes adenine•thymine (AT) base pairs (T•AT triplet) and protonated cytosine (C) recognizes guanine•cytosine (GC) base pairs (C+GC triplet) (4-19) (Fig. 1). Moser and Dervan have reported that homopyrimidine oligodeoxyribonucleotides with EDTA•Fe attached at a single position bind the corresponding homopyrimidine•homopurine tracts within large double-stranded DNA by triple helix formation and cleave at that site (20, 21). The sequence-specific recognition of double helical DNA by homopyrimidine probes is sensitive to single base mismatches (20). Oligonucleotides with EDTA•Fe at the 5' end cause a sequence-specific double strand break. The location and asymmetry of the cleavage patterns reveal that the homopyrimidine-EDTA probes bind in the major groove parallel to the homopurine strand of Watson-Crick double helical DNA (20). Recently Maher, *et al.*, have shown the inhibition of DNA binding proteins by oligonucleotide-directed triple helix formation (29).

Cleavage by triple strand formation is not observed in the absence of cations such as spermine or $Co(NH_3)_6^{3+}$ that are necessary to overcome the repulsion between the two anionic chains of Watson-Crick duplex and a third negatively charged phosphodiester backbone. The efficiency of oligonucleotide duplex cleavage is increased by a factor of 10 on addition of ethylene glycol (40% by volume) or other organic solvents that lower the relative humidity. Dehydration of the DNA favors a B to A transition and might allow the formation of the triple helix structure (A' RNA-like, 22-24) to occur more readily (25). Oligonucleotide-EDTA•Fe(II) probes do not bind-cleave Watson-Crick DNA in slightly basic solutions (pH ≥ 8). An important factor is the required protonation of the cytosine N-3 in the third strand to enable the formation of two Hoogsteen hydrogen bonds (17-19).

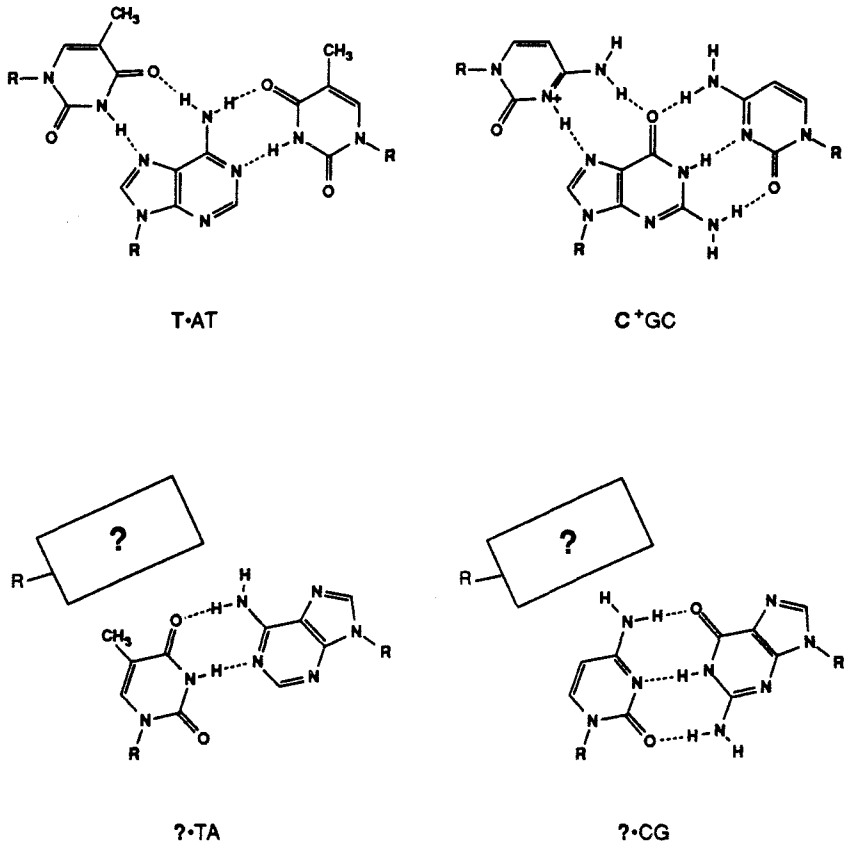


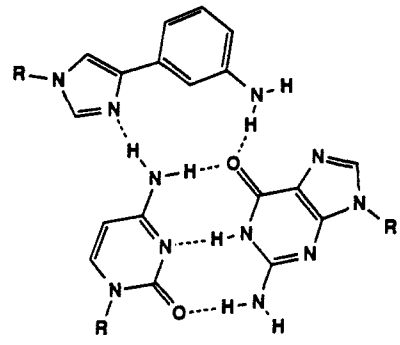
Figure 1. (Top) Isomorphous base triplets **T•AT** and **C+GC**. The additional pyrimidine is bound in the major groove by Hoogsteen hydrogen bonds to the purine base in the Watson-Crick duplex. **(Bottom)** Unknown new base triplets **?•TA** and **?•CG**.

Povsic and Dervan have shown oligonucleotide-EDTA•Fe probes containing 5-bromouracil and 5-methylcytosine bind and cleave polypurine sequence in duplex DNA with specificities comparable to those of their thymine-cytosine analogs, but with greater affinities and over an extended pH range (26). Cleavage efficiency of oligonucleotide-EDTA•Fe(II) decreases sharply below pH 6 (27), presumably as a result of partial protonation of the EDTA and the resulting loss of Fe(II) or some pH dependence of the cleavage reaction. Footprinting experiments confirm that the triple helix forms at acidic pH values (28).

Less well-understood is purine oligonucleotide recognition of double helical DNA (purine•purine•pyrimidine triplets) (16, 30-33). Recently, purine oligonucleotides have been postulated to bind parallel to purines in duplex DNA by triple helix formation (A•AT and G•GC base triplets) (33). Cooney, *et al.* have demonstrated *in vitro* gene repression using a G-rich oligonucleotide to bind a G-rich target site (33). α-Oligonucleotides have also been observed to form triple-stranded structures (34, 35).

Extension of Triple Strand Formation. Studies in the area of pyrimidine or purine oligonucleotide-directed recognition and cleavage of duplex DNA by triple strand formation clearly demonstrate the power and potential of this form of duplex DNA recognition. However, because homopyrimidine/homopurine oligonucleotides limit triple-helix formation to homopurine tracts AT and GC base pairs but not TA and CG base pairs) a general solution is lacking. It is desirable to determine whether oligonucleotides can be used to bind all four base pairs of DNA. A general solution would allow targeting of oligonucleotides (or their analogs) to any given sequence. The examination of readily available bases revealed that the triple-helix can be extended from homopurine to mixed target sequences (36). Guanine contained within a pyrimidine oligonucleotide was found to specifically recognize thymine•adenine base pairs in duplex DNA. Although there are sequence-composition effects, this finding extended specific recognition within the pyrimidine triple-helix motif to three of the four possible base pairs in double helical DNA.

The fourth base pair CG is recognized weakly in a pyrimidine-Hoogsteen motif by bases T and C (36). We wanted to design and synthesize a novel base that could bind CG base pairs strongly and selectively. Our design rationale was to use a molecular frame that would position hydrogen bond donors and acceptors at appropriate distances and angles to form hydrogen bonds to CG base pairs while maintaining a geometry compatible with the pyrimidine triple helix motif. From model building we chose as initial targets 4-phenylimidazole (Z) (1 H-bond) and 4-(3-amino)phenylimidazole (Z') (2 H-bonds) (Fig. 2). The Z' molecule was designed to form hydrogen bonds to both strands of the Watson-Crick duplex. This differs from the pyrimidine-Hoogsteen motif where the third strand forms hydrogen bonds only to the homopurine strand. Both bases Z and Z' were found to recognize CG base pairs selectively but weakly relative to T•AT and C•GC base triplets. We also examined 4-(3-benzamido)phenylimidazole (Z"). It was expected that this molecule would provide the free amino base Z' after deprotection. However, the benzamido group of Z" unlike benzamido protected bases A and C is very stable to the deprotection conditions (0.1N NaOH, 55°C, 24 hours). Interestingly, this novel base strongly recognizes pyrimidine•purine base pairs over purine•pyrimidine base pairs within the pyrimidine triple helix motif (TA>CG>AT>>GC). The utility of 4-(3-benzamido)phenylimidazole (Z") is demonstrated in the site specific binding of an 18 base pair sequence in SV40 DNA (pH 7.4, 40 °C) containing all four base pairs.



Z'•CG

Figure 2. Designed base **Z'** for CG recognition.

Results and Discussion

Synthesis of 1-β-(4-phenylimidazole)-5-DMT-2-deoxyribose-*N,N*-diisopropyl-β-cyano-phosphoramidite. The synthetic scheme for nucleoside-phosphoramidite **1** (Z) is shown in Figure 3. The sodium salt of 4-phenylimidazole was condensed with 1-chloro-2-deoxy-3,5-di-O-p-toluoyl-α-D-erythropentofuranose (37) to give the imidazole N-1 β anomer **3** in 53% yield (38). The sugar toluoyl protecting groups were removed in 1% NaOH/MeOH (39) and the 5' hydroxyl group was protected with dimethoxytrityl chloride (DMTCl) to give **4**. Reaction of **4** with 2-cyanoethyl-*N,N*-diisopropyl-chlorophosphoramidite and diisopropylethylamine (DIPEA) afforded the 5'-DMT-protected Z phosphoramidite **1** (43).

1-β-(4-(3-trifluoroacetamido or 3-benzamido)phenylimidazole)-5-DMT-2-deoxyribose-*N,N*-diisopropyl-β-cyano-phosphoramidite. The synthetic scheme for nucleoside-phosphoramidites **2a** (Z') and **2b** (Z") is shown in Figure 4. 4-(3-Nitro)phenylimidazole **5** was prepared according to the method of Bredereck *et al.* (40). The sodium salt of 4-(3-nitro)phenylimidazole was condensed with 1-chloro-3,5-di-toluoyl-sugar (37) to give the imidazole N-1 β anomer **6** in 53% yield (38). Crystal structure data of a 2-methyl-4-phenylimidazole analog confirmed the imidazole N-1, β–glycosidic linkage (41). The sugar toluoyl protecting groups were removed in 1% NaOH/MeOH (38). Hydrogenation (50 psi, H_2, 10% Pd/C) reduced the nitro group to an amine. The amine was selectively acylated by transiently protecting the sugar hydroxyls with trimethylchlorosilane (TMSCl) followed by treatment with either trifluoroacetic anhydride or benzoyl chloride (42). The 5' hydroxyl group was protected with DMTCl to give **7a** (Z') (trifluoroacetamido protected) or **7b** (Z") (benzamido protected). Reaction of **7a** or **7b** with 2-cyanoethyl-*N,N*-diisopropyl-chlorophosphoramidite and DIPEA afforded the 5'-DMT-protected Z' phosphoramidite **2a** or 5'-DMT-protected Z" phosphoramidite **2b**.

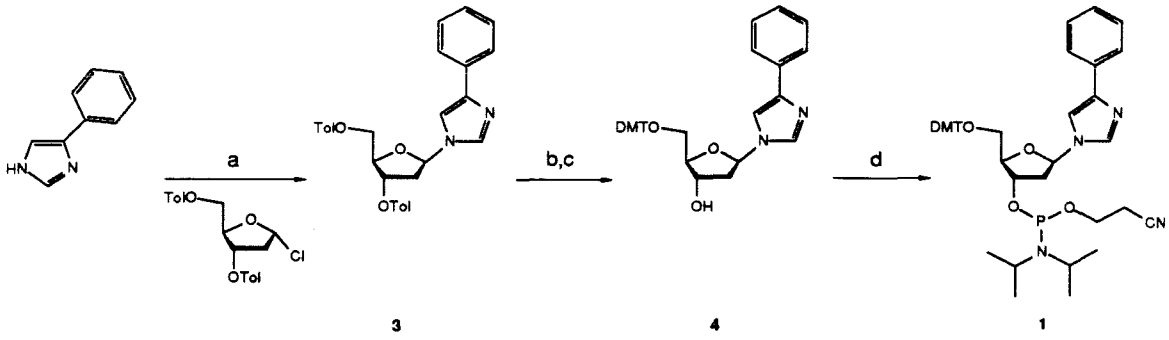


Figure 3. Scheme for the synthesis of 1-β-(4-phenylimidazole)-5'-dimethoxytrityl-2'-deoxyribose-*N,N*-diisopropyl-β-cyano-phosphoramidite (**Z**). **a)** NaH, CH_3CN, Ar; 1-chloro-2-deoxy-3,5-di-O-p-toluoyl-α-D-erythropentofuranose; **b)** 1% NaOH/MeOH **c)** DMTCl, pyridine **d)** $Cl(OCH_2CH_2CN)N(i\text{-}Pr)_2$, DIPEA, CH_2Cl_2, Ar.

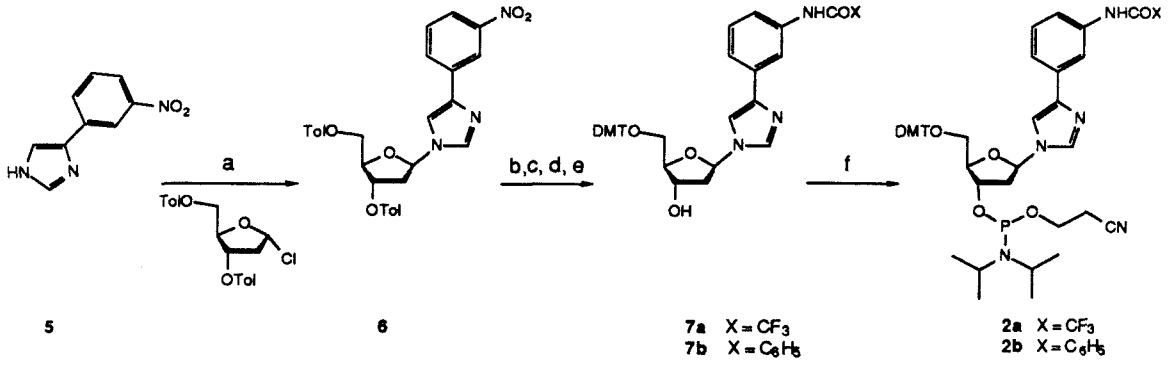


Figure 4. Scheme for the synthesis of 1-β-4-[3-trifluroacetamidophenylimidazole]-5'-dimethoxytrityl-2'-deoxyribose-*N,N*-diisopropyl-β-cyano-phosphoramidite (**Z'**) and 1-β-4-[3-benzamidophenylimidazole]-5'-dimethoxytrityl-2'-deoxyribose-*N,N*-diisopropyl-β-cyano-phosphoramidite (**Z"**). **a)** NaH, CH_3CN, Ar; 1-chloro-2-deoxy-3,5-di-O-p-toluoyl-α-D-erythropentofuranose **b)** 1% NaOH/MeOH **c)** H_2, Pd/C, MeOH, 50 psi **d)** TMSCl, pyridine; $(CF_3CO)_2O$ or C_6H_5COCl; H_2O or H_2O, NH_3 **e)** DMTCl, pyridine **f)** $Cl(OCH_2CH_2CN)N(i\text{-}Pr)_2$, DIPEA, CH_2Cl_2, Ar.

Synthesis of oligonucleotides 1-34 using novel base phosphoramidites. Oligonucleotides **1-34** were machine synthesized using β-cyano-phosphoramidite chemistry (43). The novel base phosphoramidites Z, Z', Z" typically coupled with an efficiency equal to A, G, C, and T phosphoramidites (≥97%). Because oligonucleotides **1-34** contained T* (thymidine-EDTA), 0.1N NaOH was substituted in the deprotection step for NH_4OH to circumvent possible ammoniolysis of the T* ethyl esters (44). The 4-mers T-T-Z-T, T-T-Z'-T, and T-T-Z"-T were synthesized on a 10 μmol scale for NMR studies to examine the stability of the novel bases through several machine synthesis coupling cycles and base deprotection. The glycosidic bonds of Z, Z', and Z" were found to be stable to these conditions. The trifluoroacetamido group of Z' was almost completely removed (>95%) to give the free amine, while the benzamido group of Z" was stable to the deprotection conditions (.1N NaOH, 55°C, 24 hours) (see appendix).

Binding-cleavage of oligonucleotide 30-mer duplexes. The relative affinities of the novel bases Z, Z', and Z" for all four base pairs within a pyrimidine triple-helix motif were examined. The use of oligonucleotides equipped with the DNA cleaving moiety, thymidine-EDTA•Fe(II) (T*) (27,44), allowed the relative stabilities of triple helix formation between 30 base pair (bp) DNA duplexes containing the site $d(A_7XA_7)•d(T_7YT_7)$ (XY=AT, GC, CG, or TA) and a series of 15 nt oligomers differing at one base position $d(T_7NT_7)$ [N=T, C, Z, Z', Z"] to be determined by the affinity cleaving method (1) (Fig. 5A). The 30-bp duplexes were labeled with ^{32}P at the 5' end of the target site strand $d(T_7YT_7)$. The DNA binding-cleaving reactions were performed under conditions that were sensitive to the stability of the variable base triplet in the middle of a thymine 15 nt fragment upon triple helix formation (pH 7.4, 35 °C, 40% ethanol). The most intense cleavage patterns were observed for the combinations N=T, XY=AT; N=C, XY=GC; N=Z", XY=CG ; and N=Z", XY=TA (Fig. 5B, 5C). The cleavage observed for two of these combinations (Fig. 5B, lanes 3 and 8) represents the known ability of T and C to form T•AT and C+GC base triplets, respectively. T and C also show moderate

binding-cleavage in T•GC, T•CG, and C•CG base triplets (Fig. 5B, lanes 2, 3, and 7). Strong cleavage patterns are observed for the base Z" (4-(3-benzamido)phenylimidazole) in Z"•TA and Z"•CG triplets (Fig. 5B, lanes 19 and 20) and a moderated cleavage pattern is also observed in the triplet Z"•AT (Fig. 5B, lane 17). This novel base shows significant binding-cleavage with a preference for pyrimidine•purine base pairs over purine•pyrimidine base pairs (Z"•TA > Z"•CG > Z"•AT >> Z"•GC). The 4-phenylimidazole base Z and 4-(3-amino)phenylimdazole base Z' show selective but weak binding-cleavage (Fig. 5B, lanes 13 and 17) of the base pair CG. Base Z and Z' were designed to form one and two hydrogen bonds respectively to the CG base pair. The similar weak intensities observed in lanes 13 and 17 (Fig. 5B, 5C) indicate that more than one hydrogen bond is not formed by Z'. Because the phenyl ring can freely rotate with respect to the imidazole ring, the amino group may be unavailable to hydrogen bond with the guanosine in the CG base pair

5'- T T T T T* T T T T T T T T T T - 3' 1
5'- T T T T T* T T C T T T T T T T - 3' 2
5'- T T T T T* T T Z T T T T T T T - 3' 3
5'- T T T T T* T T Z' T T T T T T T - 3' 4
5'- T T T T T* T T Z" T T T T T T T - 3' 5

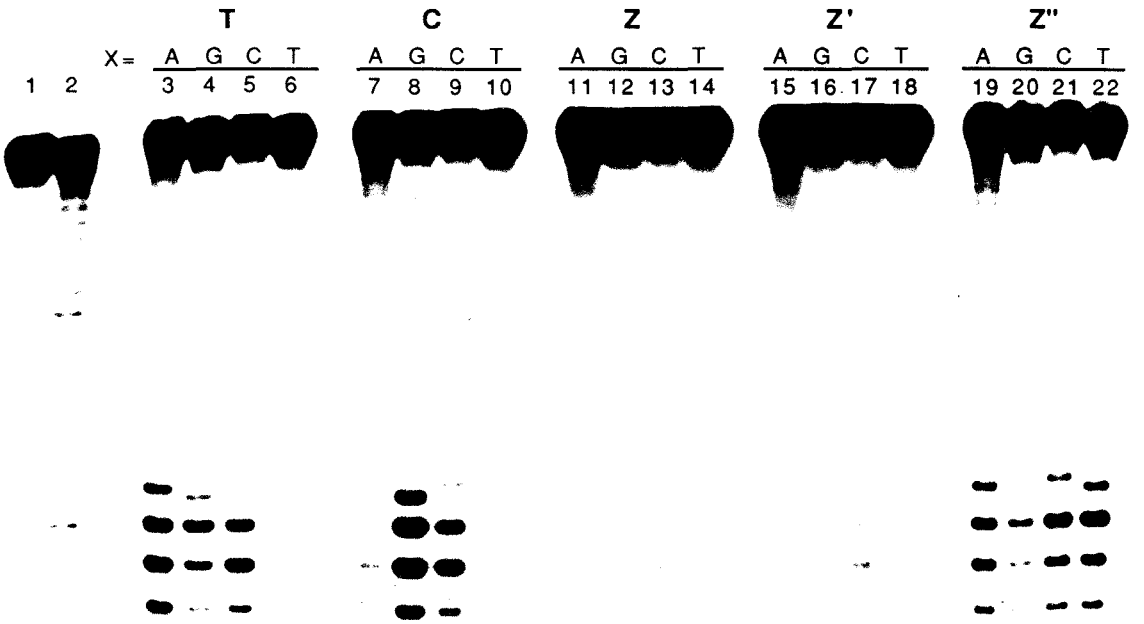
5'- C C C C C C C C C C C A A A A A A A X A A A A A A A T T T T T - 3'
3'- G G G G G G G G G G G T T T T T T T Y T T T T T T T A A A A A - 5'

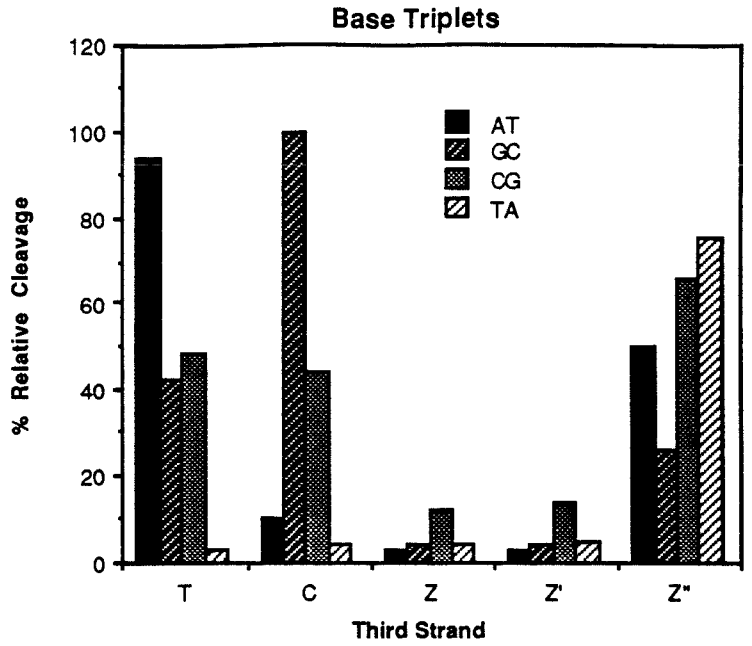
XY= AT, GC, CG, TA

Figure 5A. (Above) Sequences of oligonucleotide-EDTA **1-5** where T* is the position of the thymidine-EDTA. The oligonucleotides differ at one base position indicated in bold type. **(Below)** The box indicates the double stranded sequence bound by oligonucleotide-EDTA•Fe(II) **1-5**. The Watson-Crick base pair (AT, GC, CG, or TA) opposite the variant base in the oligonucleotide is also in bold type.

(Fig.2). A bridged-planar phenylimidazole structure may be a better target for multi-hydrogen bond formation to both bases of a Watson-Crick base pair. Perhaps no hydrogen bonds are formed in the case of Z and Z'. For steric reasons these bases might simply prefer CG base pairs. The base Z", 4-(3-benzamido)phenylimidazole, is less selective but binds-cleaves more efficiently than Z and Z'. This result suggests that the benzamido protecting group is largely responsible for a specific interaction stabilizing triple helix formation. Modeling studies suggest Z" lies out of the plane of the opposite Watson-Crick base pair to avoid steric clashes. The lowest energy structures generated by computer modeling of triple-helices containing Z" opposite all four Watson-Crick base pairs place the benzamido group lying along the major groove of the triple helix (Fig. 6) (45). Space filling models suggest good van der Waals interactions between the walls of the major groove and the benzamido group (45).

Figure 5B. Autoradiogram of the 20 percent denaturing polyacrylamide gel. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled 30-mer duplex [~0.5 μM (bp) (~13,400±700 cpm)] in a solution of tris-acetate, pH 7.4 (25 mM), NaCl (50 mM), calf thymus DNA [100 μM (bp)], and 40% ethanol and incubated at 35 °C for 1 hr (51). Cleavage reactions were initiated by addition of DTT (3 mM) and allowed to proceed for 6 hours at 35 °C. The reactions were stopped by freezing and lyophilization and the cleavage products were analyzed by gel electrophoresis (1200-2000 V, BPB 23 cm). (Lanes 1-22) Duplexes containing 5' end-labeled d($A_5T_7YT_7G_{10}$). (Lane 1) Control showing intact 5' labeled 30 bp DNA standard obtained after treatment according to the cleavage reactions in the absence of oligonucleotide-EDTA. (Lane 2) Products of Maxam-Gilbert G+A sequencing reaction. (Lanes 3-22) DNA cleavage products produced by oligonucleotide-EDTA•Fe(II) (**1-5**); **1** (Lanes 3-6); **2** (lanes 7-10); **3** (lanes 11-14); **4** (lanes 15-18); **5** (lanes 19-22). **XY=AT** (Lanes 3,7,11,15, and 19); **XY=GC** (lanes 4,8,12,16, and 20); **XY=CG** (lanes 5,9,13,17, and 21); **XY=TA** (lanes 6,10,14,18, and 22).


T
C
Z
Z'
Z"
X= A G C T
1 2
3 4 5 6
7 8 9 10
11 12 13 14
15 16 17 18
19 20 21 22



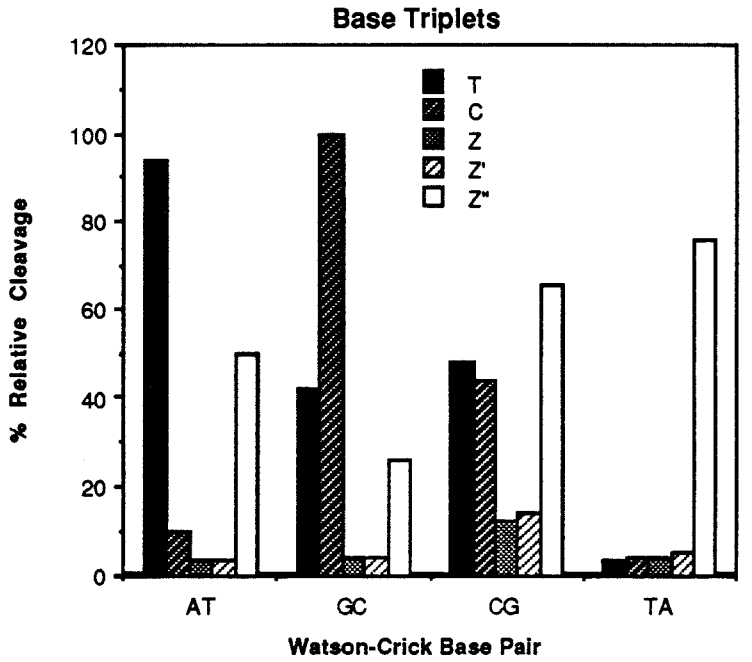


Figure 5C. Bar graph presenting the relative cleavage data from densitometric analysis of Fig. 5B. Twenty base triplets were examined for binding specificity compatible with pyrimidine-Hoogsteen triple helix motif by the experiment described in Fig. 5A, B. The data are reproducible within ± 10% of reported values.

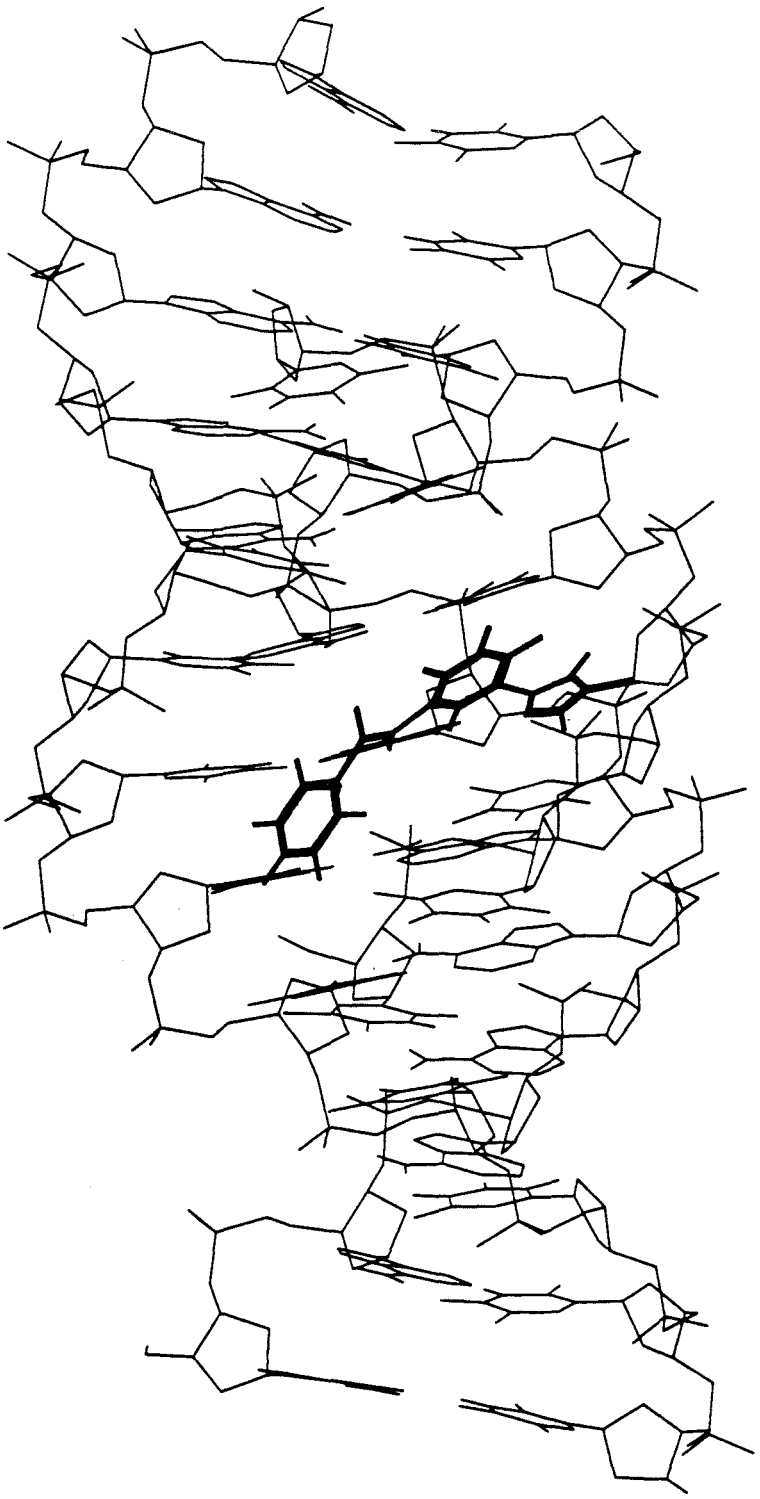


Figure 6. Model of nucleoside **Z"** (outlined in bold) in a triple helix.

Binding-cleavage of SV40 AccI-HaeII restriction fragment at a site containing a TA base pair. The general utility of Z" as a base to use opposite TA and CG base pairs at mixed sequences within larger DNA fragments was examined at three different sites within SV40 DNA. The cleavage of a 799 bp SV40 restriction fragment by the oligonucleotide-EDTA•Fe(II) series (**6-12**) (5'-T*TTTCCTCCTTNTCTTC-3') was first examined (Fig. 7A). This restriction fragment contains the 17-bp sequence d(AAAAGGAGGAATAGAAG), which represents a purine rich site containing one pyrimidine (T). The cleavage efficiencies of oligonucleotides **6-12**, differing at one base position opposite the TA Watson-Crick base pair, were examined under conditions sensitive to the stability of the base triplet at the TA site (pH 6.6, 37 °C). Oligonucleotide-EDTA **9** and **12**, but not **6**, **7**, **8**, **10** or **11** produced significant site specific cleavage on

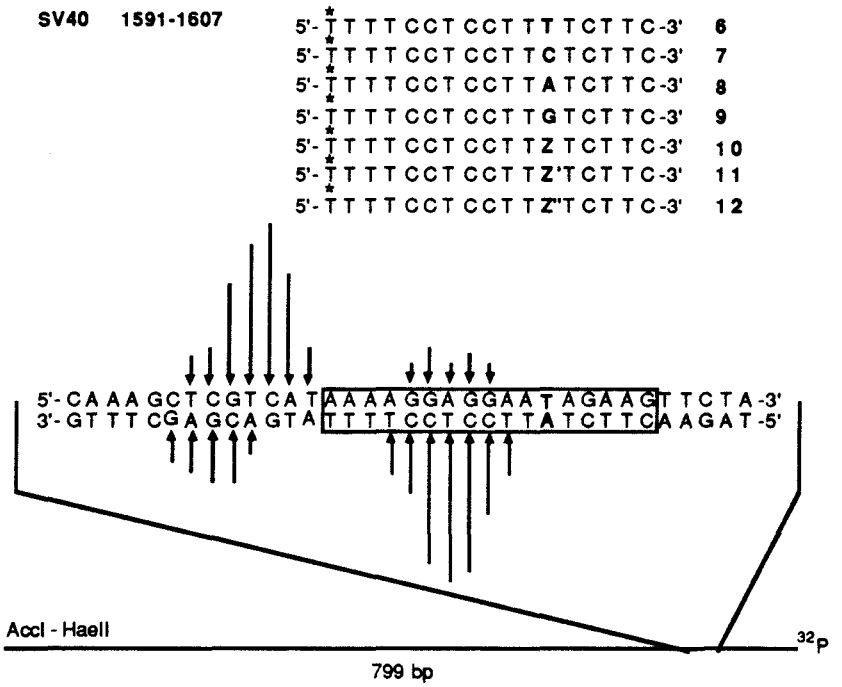


Figure 7A. (Above) Sequence of oligonucleotide-EDTA **6-12** where T* is the position of thymidine-EDTA. The oligonucleotides differ at one base position indicated by bold type. **(Below)** Histogram of the DNA cleavage patterns derived by densitometry of the autoradiogram in Fig. 7B (lane 9) and the 5' labeled strand (not shown) from the cleavage of the 799 bp SV40 restriction fragment with oligonucleotide-EDTA **12**.

the 799-bp fragment (Fig. 7B). Oligonucleotide **9** (Fig. 7B, lane 6) contains G opposite the TA base pair. We have recently reported TA recognition by G (36). Oligonucleotide **11** (Fig. 7B, lane 9) contains Z" opposite the TA. The Z"•TA triplet appears to be slightly more stable than the G•TA triplet. The relative cleavage data is presented in a bar graph in Figure 7C. Oligonucleotides **10** and **11**, which contain Z and Z' respectively, both bind the target site weakly.

Figure 7B. Autoradiogram of the 8 percent denaturing polyacrylamide gel of cleavage reactions on an AccI-HaeII restriction fragment of SV40 DNA. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled restriction fragment [~100 nM (bp) (~8,700±600 cpm)] in a solution of tris-acetate, pH 6.6 (50 mM), NaCl (50 mM), and calf thymus DNA [100 μM (bp)] (51). This was incubated at 37 °C for 1 hr. Cleavage reactions were initiated by addition of DTT (3 mM) and allowed to proceed for 3 hours at 37 °C. The reactions were stopped by precipitation with ethanol and the cleavage products were analyzed by gel electrophoresis (1200-2000 V, BPB 30 cm). (Lanes 1-9) 3' end-labeled AccI-HaeII restriction fragment of SV40. (Lane 1) Control containing no oligonucleotide-EDTA•Fe(II). (Lane 2) Maxam-Gilbert G sequencing reactions. (Lanes 3-9) DNA cleavage products by oligonucleotide-EDTA•Fe(II) (**6-12**): **6** (lane 3); **7** (lane 4); **8** (lane 5); **9** (lane 6); **10** (lane 7); **11** (lane 8); **12** (lane 9).

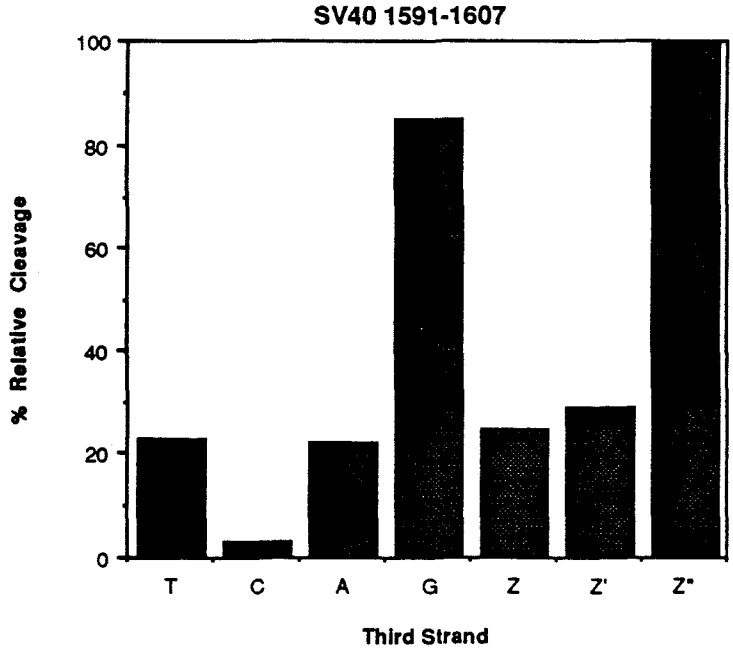
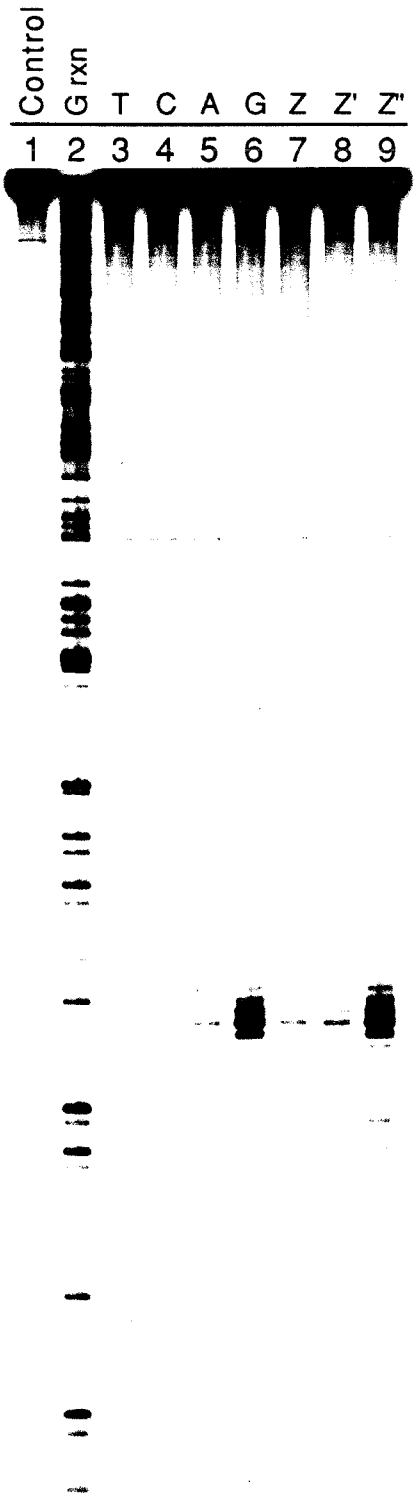


Figure 7C. Bar graph presenting the relative cleavage data from densitometric analysis of Fig. 7B. The data are reproducible within ± 10% of reported values.


Control
G rxn
T
C
A
G
Z
Z'
Z"
1 2 3 4 5 6 7 8 9

Binding-cleavage of SV40 XhoI-NdeII restriction fragment at a site containing a CG base pair. The cleavage of a 294 bp SV40 restriction fragment by the oligonucleotide-EDTA•Fe(II) series (**13-19**) (5'-T*CTTCTTCNTCTCCTTT-3') was examined next (Fig. 8A). This restriction fragment contains the 17-bp sequence d(AGAAGAAGCAGAGGAAA), which represents a purine-rich site containing one pyrimidine (C). The cleavage efficiencies of oligonucleotides **13-19**, differing at one base position opposite the TA Watson-Crick base pair, were examined under conditions sensitive to the stability of the base triplet at the CG site (pH 6.6, 37 °C). Oligonucleotide-EDTA **13-16,** and **18** showed a weak cleavage pattern at the target site (Fig. 8B). The most intense cleavage pattern was observed with oligonucleotide **19** containing the base Z". The relative cleavage data is presented in a bar graph in Figure 8C. Interestingly, oligo-

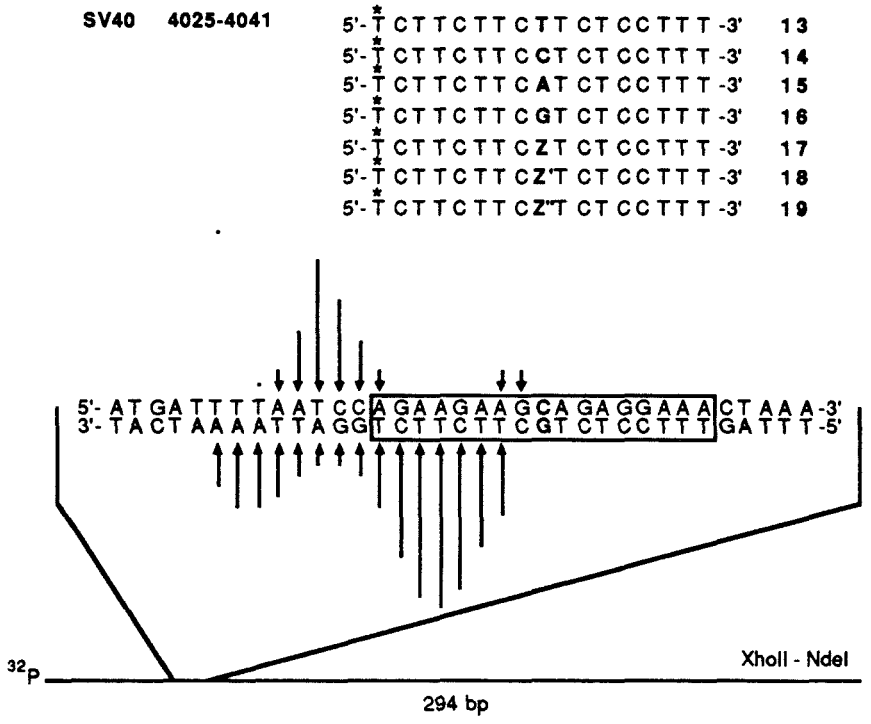


Figure 8A. (Above) Sequence of oligonucleotide-EDTA **13-19** where T* is the position of thymidine-EDTA. The oligonucleotides differ at one base position indicated by bold type. **(Below)** Histogram of the DNA cleavage patterns derived by densitometry of the autoradiogram in Fig. 8B (lane 6) and the 5' labeled strand (not shown) from the cleavage of the 294 bp SV40 restriction fragment with oligonucleotide-EDTA **19**.

nucleotide **17**, a weak binder, is not specific for binding the target site. It also binds a site three bases 3' of the target site where a Z•CG triplet is traded for a T•CG and Z•AT triplet (Fig. 8D). This result is consistent with oligonucleotide data (Fig. 5B) where stronger binding-cleavage is observed for the T•CG triplet relative to the Z•CG triplet. This tendency to bind to an alternative site is not observed in the case of oligonucleotide **18** containing Z'. At this DNA sequence the base Z lacking the 3-amino group displays different triple helix forming properties than the base Z'.

Figure 8B. Autoradiogram of the 8 percent denaturing polyacrylamide gel of cleavage reactions on the XhoI-NdeII restriction fragment of SV40 DNA. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled restriction fragment [~100 nM (bp) (~9,000±300 cpm)] in a solution of tris-acetate, pH 6.6 (50 mM), NaCl (50 mM), and calf thymus DNA [100 μM (bp)] (51). This was incubated at 37 °C for 1 hr. Cleavage reactions were initiated by addition of DTT (3 mM) and allowed to proceed for 3 hours at 37°C. The reactions were stopped by precipitation with ethanol and the cleavage products were analyzed by gel electrophoresis (1200-2000 V, BPB 38 cm). (Lanes 1-9) 3' end-labeled XhoII-NdeI restriction fragment of SV40. (Lane 1) Control containing no oligonucleotide-EDTA•Fe(II). (Lane 2) Maxam-Gilbert G sequencing reactions. (Lanes 3-9) DNA cleavage products by oligonucleotide-EDTA•Fe(II) (**13-19**): **13** (lane 3); **14** (lane 4); **15** (lane 5); **16** (lane 6); **17** (lane 7); **18** (lane 8); **19** (lane 9).

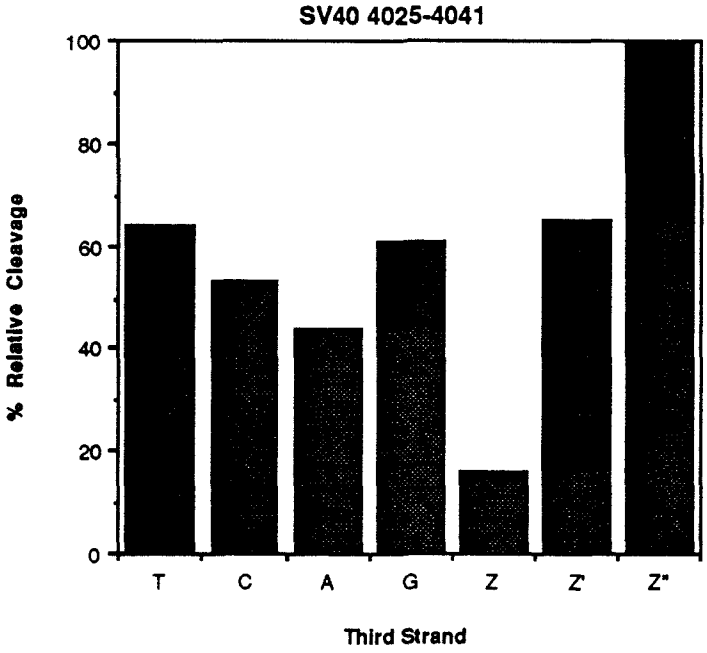
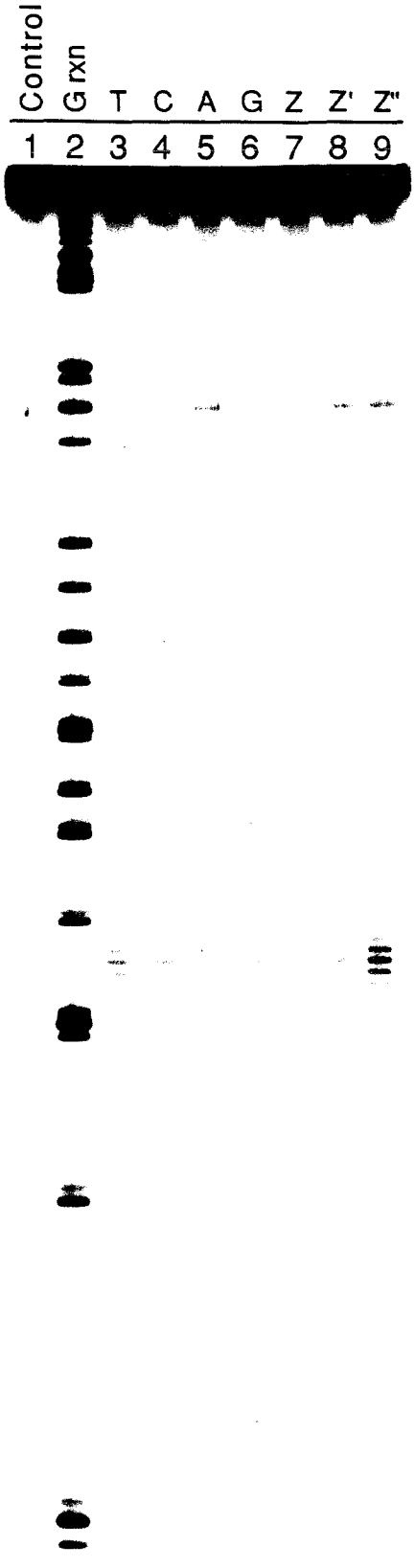


Figure 8C. Bar graph presenting the relative cleavage data from densitometric analysis of Fig. 8B. The data are reproducible within ± 10% of reported values.


Control
G rxn
T C A G Z Z' Z"
1 2 3 4 5 6 7 8 9

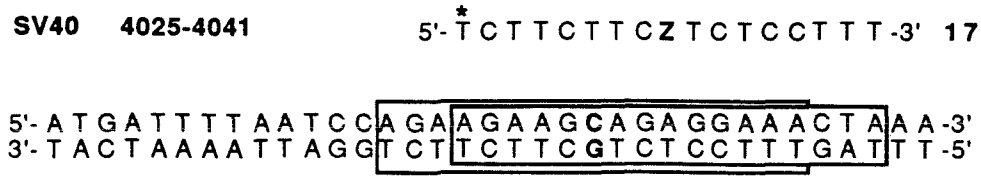


Figure 8D. Sequence of oligonucleotide-EDTA **17** where T* is the position of thymidine-EDTA. The box in back indicates the original target sequence. The box in front indicates the double-stranded sequence bound by oligonucleotide•EDTAFe(II) **17**.

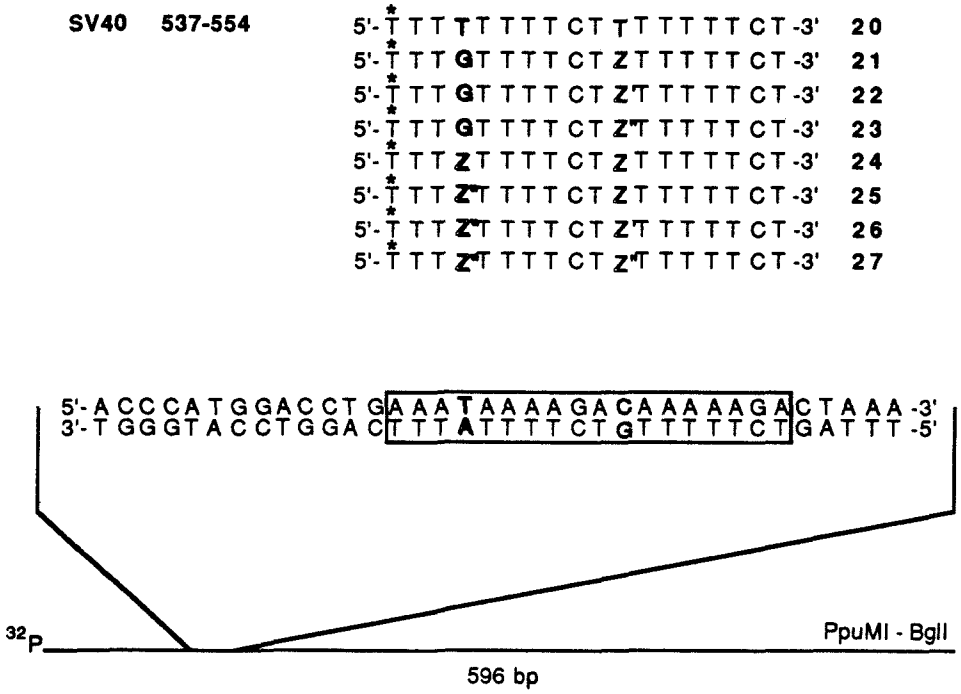
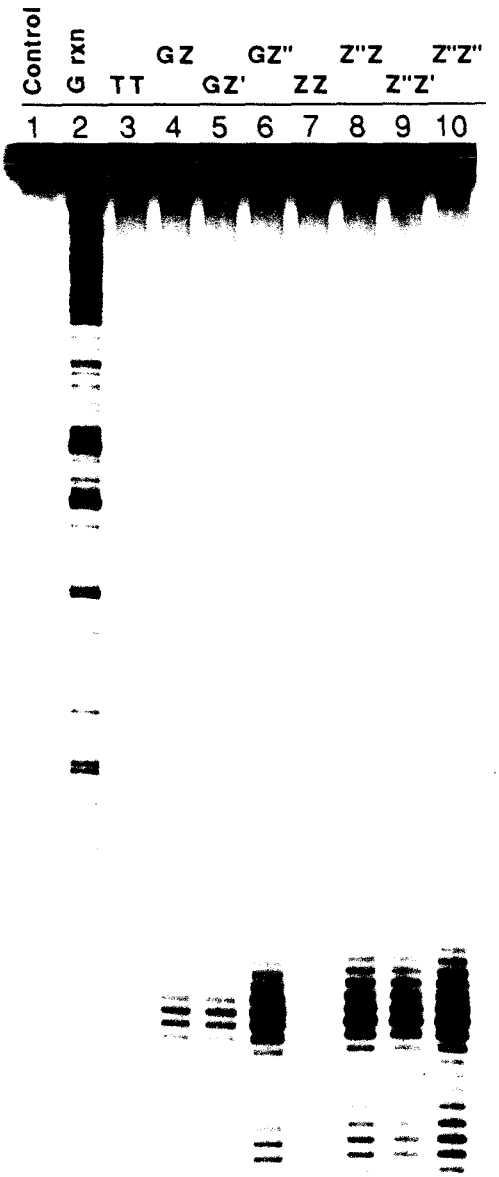


Figure 9A. (Above) Sequence of oligonucleotide-EDTA **20-27** where T* is the position of thymidine-EDTA. The oligonucleotides differ at one base position indicated by bold type. **(Below)** Histogram of the DNA cleavage patterns derived by densitometry of the autoradiogram in Figure 9B (lane 6) and the 5' labeled strand (not shown) from the cleavage of the 596 bp SV40 restriction fragment with oligonucleotide-EDTA **27**.

Binding-cleavage of SV40 PpuMI-BglI restriction fragment at a site containing a TA and CG base pair. The cleavage of a 596 bp SV40 restriction fragment by the oligonucleotide-EDTA•Fe(II) series (**20-27**) (5'-T*$T_2NT_4CTNT_5CT$-3') was the third site examined. This restriction fragment contains the 18-bp sequence d(AAATAAAAGACAAAAAGA), which represents a purine-rich site containing two pyrimidines (T and C) (Fig. 9A). Cleavage at this site demonstrates binding of a sequence containing all four base pairs. The cleavage efficiencies of oligonucleotides **20-27**, differing at two base positions opposite a TA and CG Watson-Crick base pair, were examined under conditions sensitive to the stability of the base triplets at the TA and CG sites (pH 7.4, 40 °C). Oligonucleotide-EDTA **23** (G•TA and Z"•CG) and **27** (Z"•TA and Z"•CG) produced the most significant site-specific cleavage on the 799-bp fragment (Fig. 9B). The relative cleavage data is presented in a bar graph in Figure 9C.

Oligonucleotides containing Z" opposite a TA base pair stabilize triple helix formation at this sequence to a greater extent than G. Oligonucleotide **20** does not bind its target site but instead a site located five bases 3' to the target site (Fig. 9D). A T•TA interaction has been substituted for the T•GC interaction. This finding is consistent with oligonucleotide data (Fig. 5B) where the cleavage intensity observed for the T•CG base triplet is greater than that for the T•TA base triplet.

Figure 9B. Autoradiogram of the 8 percent denaturing polyacrylamide gel of the cleavage reactions on a PpuMi-BglI restriction fragment of SV40 DNA. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled restriction fragment [~100 nM (bp) (~12,500±600 cpm)] in a solution of tris-acetate, pH 7.4 (50 mM), NaCl (50 mM), and calf thymus DNA [100 μM (bp)] (51). This was incubated at 40 °C for 1 hr. Cleavage reactions were initiated by addition of DTT (3 mM) and allowed to proceed for 3 hours at 40 °C. The reactions were stopped by precipitation with ethanol and the cleavage products were analyzed by gel electrophoresis (1200-2000 V, BPB 30 cm). (Lanes 1-9) 3' end-labeled PpuMI-BglI restriction fragment of SV40. (Lane 1) Control containing no oligonucleotide-EDTA•Fe(II). (Lane 2) Maxam-Gilbert G sequencing reactions. (Lanes 3-9) DNA cleavage products by oligonucleotide-EDTA•Fe(II) (**20-27**): **20** (Lane 3); **21** (lane 4); **22** (lane 5); **23** (lane 6); **24** (lane 7); **25** (lane 8); **26** (lane 9); **27** (lane 10).


Control
G rxn
TT
GZ
GZ'
GZ"
ZZ
Z"Z
Z"Z'
Z"Z"
1 2 3 4 5 6 7 8 9 10

The cleavage patterns for oligonucleotides **21-23,** which place G opposite the TA base pair, are slightly different than those for oligonucleotides **24-27** where Z" is opposite the TA base pair. This difference in pattern can be attributed to altered positioning of the T*-EDTA•Fe(II) cleaving functionality with respect to the DNA, indicating that oligonucleotides differing only at one base (G/Z") exert different structural effects on the Watson-Crick duplex upon binding by triple helix formation (36).

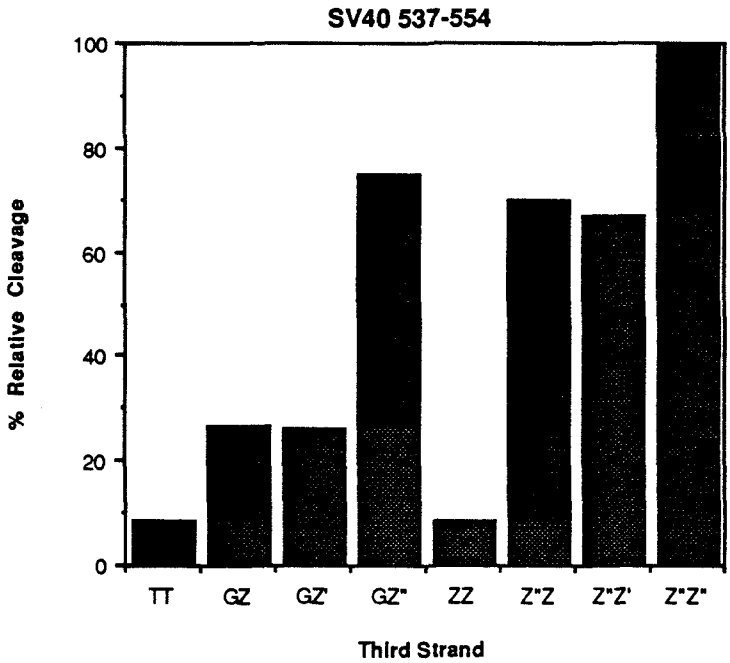


Figure 9C. Bar graph presenting the relative cleavage data from densitometric analysis of Figure 9B. The data are reproducible within ± 10% of reported values.

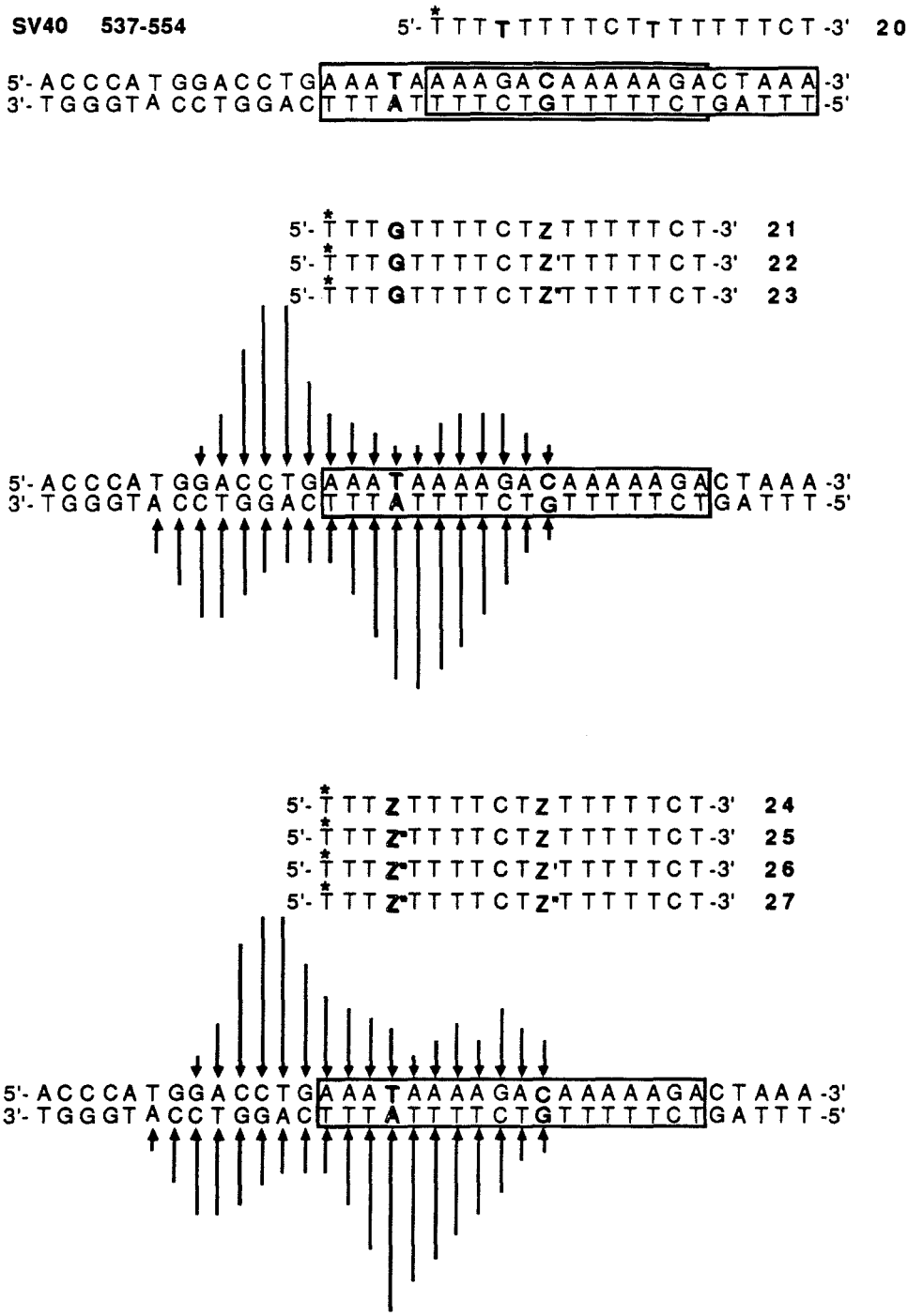


Figure 9D. Sequence of oligonucleotide-EDTA **20-27** where T* is the position of thymidine-EDTA. **(Above)** The box in front indicates the double-stranded sequence bound by oligonucleotide•EDTAFe(II) **20**. The box in back indicates the original target sequence. **(Below)** Histogram of the DNA cleavage patterns for oligonucleotides **21-27** derived by densitometry of the autoradiogram in Fig. 9B.

The effects of pH and temperature at this sequence for oligonucleotides **20**, **21**, **23**, **24**, **25**, and **27** are presented in a bar graph in Figure 10. Raising the temperature and pH has the effect of increasing the stringency of triple helix formation conditions. The binding-cleavage of oligonucleotides with less affinity for the target site falls off more rapidly when the temperature and pH are increased.

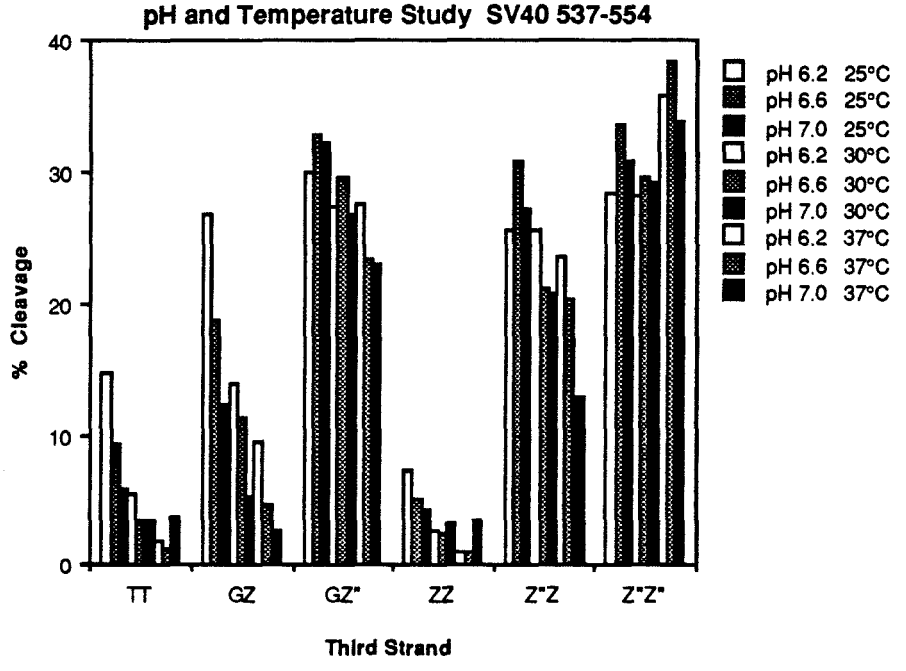


Figure 10. Bar graph presenting the relative cleavage data from densitometric analysis of pH and temperature studies of oligonucleotides **20**,**21**,**23**,**24**,**25**,and **27.** The data are reproducible within ± 10% of reported values.

Double strand cleavage of SV40 DNA. The ability of oligonucleotide-EDTA•Fe(II) **27** to cause site-specific double strand breaks in SV40 DNA is documented in Fig. 11. SV40 was digested with BglI to produce a 5.24-kbp fragment, which contained the sequence d(AAATAAAAGACAAAAAGA) located 2.21 kbp and 3.03 kbp from the ends. The ^{32}P end-labeled DNA was allowed to react with oligonucleotide-EDTA•Fe(II) in the presence of DTT (pH 7.0, 37 °C). Separation of the cleavage products by agarose gel electrophoresis revealed one major cleavage site producing two DNA fragments, 2.21 and 3.03 kb in size (Fig.11B).

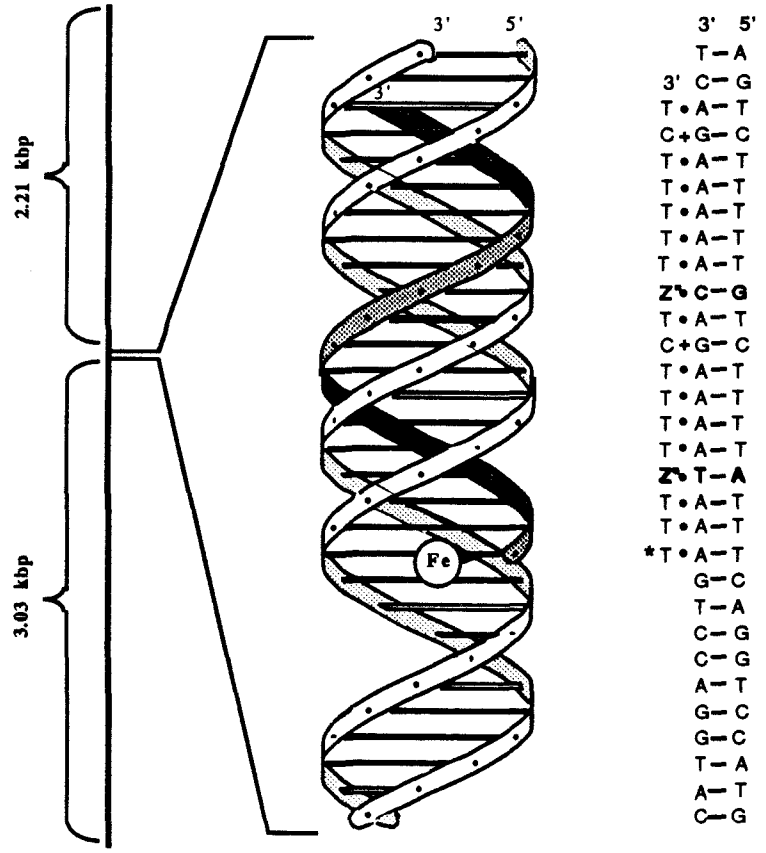


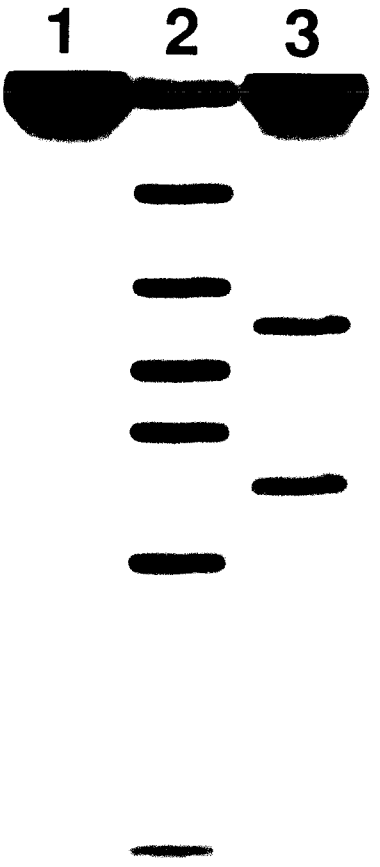
Figure 11A. (Left) The coarse resolution cleavage pattern from Fig. 11B. **(Right)** Simplified model of the triple helix complex between the bound oligonucleotide-EDTA•Fe(II) **27** and a single site within the 5.24 kb plasmid DNA.

Triple strand target site composition. Several oligonucleotides (**28-34**) containing 6-17 Z" bases were synthesized to bind the TAR site in HIV DNA and two sites in SV40 DNA (Fig. 12). No cleavage was observed for these oligonucleotides. One possibility for the failure of these oligonucleotides to bind their target sites may be that steric interactions preclude a Z" base followed by another Z" base within the triple helix motif. Another consideration is that it is not clear how far one can deviate from homopurine•homopyrimidine target sequences and still obtain triple helix formation. Normally, in oligonucleotide duplexes, cleavage by triple strand formation at homopurine target sites requires 20% ethanol. No cleavage is observed under the same conditions when one pyrimidine is introduced into the purine target site. In these cases, 35-40% ethanol is required for strong binding-cleavage to be observed. Ethanol may facilitate triple strand formation by lowering the relative humidity of the DNA, which induces a B to A transition (25). Consequently, for a mixed sequence, the B to A transition may not occur as readily.

In fact, all homopurine sites do not form triple helices with equal ease (46). The second site studied in this work d(AGAAGAAGCAGAGGAAA) does not form a triple helix as readily as the other two sites d(AAAAGGAGGAATAGAAG) and d(AAATAAAAGACAAAAAGA). Generally, triple helices appear to form more easily for sequences containing more AT base pairs than CG base pairs (46). This is in part due to

---

Figure 11B. Autoradiogram of double strand cleavage of SV40 length analyzed on a 1 percent agarose gel. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA **27** (8 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled linearized plasmid [~100 nM (bp) (~50,000 cpm)] in a solution of tris-acetate, pH 7.0 (50 mM), NaCl (50 mM), and calf thymus DNA [100 μM (bp)] and incubated 1 hour at 37 °C (51). Cleavage reactions were initiated by addition of DTT (3 mM) and allowed to proceed for 4 hours at 37 °C. The reactions were stopped by preciptitation with ethanol and the cleavage products were analyzed by gel electrophoresis (120 V, BPB to the bottom of the gel). (Lanes 1-3) SV40 linearized with BclI and 3' end-labeled at both ends. (Lane 1) Control containing no oligonucleotide-EDTA•Fe(II). (Lane 2) DNA size markers obtained by digestion of BclI linearized SV40 with AccI, HaeII, and BglI: 5243 (undigested DNA), 4101, 3305, 2778, 2465, 1938, 1142. (Lane 3) DNA cleavage products produced by oligonucleotide-EDTA•Fe(II) **27**.


1
2
3

the required protonation of C in C+GC base triplets. The pH sensitivity at these sequences can be decreased by substituting 5-methylcytosine for cytosine (26). Higher GC content may explain why the second sequence, with one pyrimidine, is a poorer triple strand site than the third sequence even though it contains two pyrimidines. However, the first and second site contain the same number of GC and AT base pairs and differ only in one TA/CG base pair. Thus, it is possible that not only content but also the ordering of the nucleotides affects ease of triple helix formation.

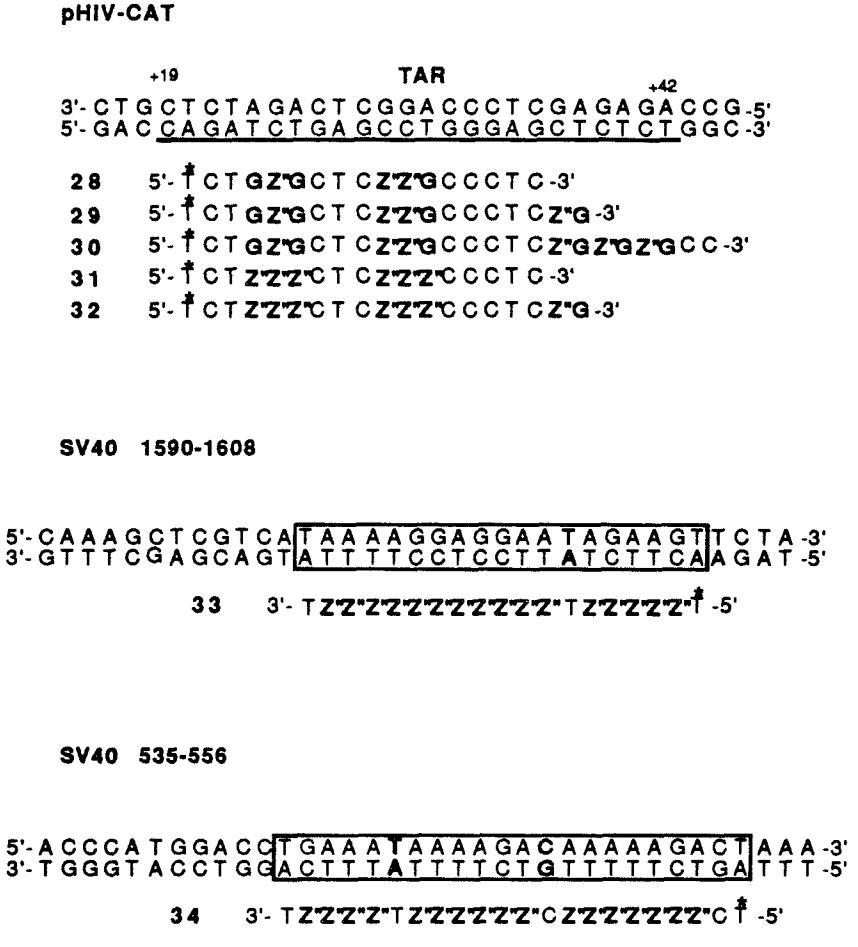


Figure 12. Sequence of oligonucleotide-EDTA **28-34** where T* is the position of thymidine-EDTA. **(Above)** Sequences synthesized to target the TAR site in HIV DNA. **(Below)** Sequences synthesized to target SV40 DNA sites.

Conclusion

The novel base 4-(3-benzamido)phenylimidazole enables the selective binding and cleavage of a site containing all four base pairs by triple helix formation. Although the Z"•TA and Z"•CG triplets within a pyrimidine oligonucleotide extend, in a formal sense, triple helix specificity to all four possible base pairs of double helical DNA, some limitations on sequence composition exist. Rather these results provide important structural leads for the design of deoxyribonucleosides (or their analogs) with nonnatural heterocycles directed toward a more general solution to sequence-specific recognition of duplex DNA at any sequence.

Experimental Section

Synthesis. ^{1}H NMR spectra were recorded at 400 MHz on a JEOL JNM-GX400. Chemical shifts are reported in parts-per-million downfield from tetramethylsilane. High resolution fast atom bombardment mass spectra (FAB MS) were obtained from the Midwest Center for Mass Spectrometry at the University of Nebraska, Lincoln. Flash chromatography was carried out under positive air pressure using EM Science Kieselgel 60 (230-400 Mesh). Corrected melting points were obtained on a Thomas Hoover capillary melting point apparatus. Elemental analyses were performed by the analytical laboratory, California Institute of Technology. Reagent grade chemicals were used as received, except for diisopropylethylamine, which was distilled from CaH_2. A Beckman System 1 Plus DNA Synthesizer was utilized for synthesis of oligonucleotides. Concentrations of oligonucleotides were determined by UV using a Perkin-Elmer Lambda 4c UV-vis spectrophotometer.

1-Chloro-2-deoxy-3,5-di-O-*p*-toluoyl-α-D-erythropentofuranose (37). To a stirring solution of 2-deoxy-D-ribose (13.41 g, 100 mmol) in methanol (240 ml) was added dichloromethylsilane (250 μL in 30 mL MeOH, 2.06 mmol). After stirring at room

temperature for 15 min the reaction was neutralized with silver carbonate (Aldrich, 5.0 g, 18.1 mmol). The mixture was filtered, rotoevaporated, and dried under vacuum to give a light yellow syrup. The syrup was further dried by addition of a small amount of dry pyridine and re-evaporating under vacuum. (TLC: CH_3CN/CH_2Cl_2, 2:1, v/v; R_f=.31; anisaldehyde stain). The syrup was then dissolved in dry pyridine (80 ml), placed in an ice bath, and cooled to 0°C. *p*-Toluoylchloride (29.1 mL, 220 mmol) was added dropwise to the stirring solution and a white precipitate was formed. After addition of the *p*-toluoylchloride was complete the ice bath was removed and the reaction stirred for 2 hours between 40-50 °C. Completion of the reaction was determined by TLC using an acidic p-anisaldehyde stain (TLC: hexane/ether, 1:1; v/v; R_f=.59 and .50). Additional portions of *p*-toluoylchloride (5 mL) were added until no starting material (R_f=.10) was observed. Upon completion, water (400 mL) and ether (300 mL) were added to the reaction. The water layer was removed and the ether layer washed with .2N H_2SO_4 (~900 mL) until the aqueous layer was slightly acidic. The ether layer was then washed with aqueous saturated $NaHCO_3$, dried over anhydrous Na_2SO_4, filtered and evaporated to a honey colored syrup. The syrup was dissolved in glacial acetic acid (50 mL) and added to glacial acetic acid (80 mL) saturated at 10 °C with HCl. The bubbling of HCl gas through the stirring solution was continued (~5 min) at 10 °C until a thick crystalline slurry was obtained. The white crystals were filtered and washed with ether (~400 mL) until the filtrate was colorless. The crystals were further dried and stored under vacuum (TLC: hexane/ether, 1:1, v/v, R_f=.22; m.p.=109 °C). The yield obtained over three steps was 24.5 g (63%). ^1H NMR ($CDCL_3$) δ 7.99 (d, 2H, J=8.06, toluoyl); δ 7.89 (d, 2H, J=8.06, toluoyl); δ 7.26 (d, 2H, J=8.06, toluoyl); δ 7.23 (d, 2H, J=8.06, toluoyl); δ 6.47 (d, 1H, J=5.13, $H_{1'}$); δ 5.56 (dxd, 1H, J=2.93, 7.32, $H_{3'}$); δ 4.86 (m, 1H, J=2.93, 4.03, 7.32, $H_{4'}$); δ 4.68 (dxd, 1H, J=2.93, 12.09, $H_{5'}$); δ 4.59 (dxd, 1H, J=4.03, 12.09, $H_{5'}$); δ 2.87 (dxdxd, 1H, J=5.13, 7.32, 15.02, $H_{2'}$); δ 2.74 (d, 1H, J=15.02, $H_{2'}$); δ 2.42 (s, 3H, CH_3 toluoyl); δ 2.41 (s, 3H, CH_3 toluoyl).

1-β-(4-Phenylimidazole)-2-deoxy-3,5-di-O-*p*-toluoyl-ribose 3 was prepared using the stereospecific sodium salt glycosylation procedure of Kazimierczuk *et al.* (38). Sodium hydride (60% dispersion in oil, 0.8 g, 22.0 mmole) was added to a suspension of 4-phenylimidazole (2.88 g, 20.0 mmol) in dry CH_3CN (180 mL). The mixture was stirred under argon at room temperature for 30 min. 1-Chloro-2-deoxy-3,5-di-O-p-toluoyl-α-D-erythropentofuranose (37) (8.50 g, 22 mmol) was added portionwise with stirring over a period of 30 min. After stirring under argon at room temperature for an additional hour, the reaction was filtered to remove a small amount of insoluble material and evaporated. Flash chromatography (hexane/CH_2Cl_2/CH_3CN, 3:2:1, v/v/v, TLC: R_f=.28) yielded 5.86 g of slightly impure product. Crystallization from ethanol and a small amount of CH_2Cl_2 gave 5.26 g (53%) of pure white needles (m.p. 160-161°C). ^1H NMR ($CDCL_3$) δ 7.96 (d, 2H, J=8.06, toluoyl); δ 7.95 (d, 2H, J=8.06, toluoyl); δ 7.70 (d, 1H, J=1.10, imidazole); δ 7.55-7.52 (m, 2H, phenyl); δ 7.31-7.19 (m, 8H, 4H toluoyl/ 3H phenyl/ 1H imidazole); δ 6.16 (t, 1H, J=6.96, $H_{1'}$); δ 5.73 (m, 1H, $H_{3'}$); δ 4.75 (m, 1H, $H_{4'}$); δ 4.63-4.57 (m, 2H, $H_{5'}$); δ 2.76-2.72 (m, 2H, $H_{2'}$); δ 2.44 (s, 3H, CH_3); δ 2.42 (s, 3H, CH_3). FAB MS, calc. for $C_{30}H_{29}N_2O_5$ ($M+H^+$): 497.2118. Found: 497.2076. Anal. calcd. for $C_{30}H_{28}N_2O_5$: C, 72.56; H, 5.68; N, 5.64. Found: C, 72.42; H, 5.67; N, 5.64. A major side product in the reaction (~15-20%, slower moving by TLC) results from the formation of the glycosyl bond at the imidazole N-3 instead of N-1.

1-β-(4-phenylimidazole)-5-dimethoxytrityl-2-deoxyribose 4. The toluoyl groups were removed from phenylimidazole sugar **3** by the method of Mashimo and Sato (39). Ribose **3** (1.99 g, 4.0 mmol) was stirred in 1% (by weight) NaOH/MeOH (80 mL) at room temperature until it had completely dissolved (~10 min). Silica gel (10 g) was added to the clear colorless solution. The silica gel/reaction mixture was concentrated under reduced pressure, further dried under vacuum, and then loaded on top of a silica gel column (90 g). Column chromatography (CH_2Cl_2/toluene/MeOH, 8:1:1, v/v/v, TLC: R_f=.31) yielded 1.0 g (96%) of a clear colorless oil, which crystallized upon standing. ^1H

NMR (CD_3OD) δ 7.93 (d, 1H, J=1.10, imidazole); δ 7.73 (m, 1H, phenyl); δ 7.71 (m, 1H, phenyl); δ 7.69 (d, 1H, J=1.10, imidazole); δ 7.37-7.32 (m, 2H, phenyl); δ 7.25-7.21 (txt, 1H, J=7.32, 1.10, phenyl); δ 6.11 (t, 1H, J=6.59, $H_{1'}$); δ 4.48 (dxt, 1H, J=6.23, 3.30, $H_{3'}$); δ 3.97 (dxd, 1H, 7.69, 4.03, $H_{4'}$); δ 3.75 (dxd, 1H, J=11.72, 4.03, $H_{5'}$); δ 3.69 (dxd, 1H, J=11.72, 4.40, $H_{5'}$); δ 2.52 (dxdxd, 1H, J=13.56, 6.96, 6.23, $H_{2'}$); δ 2.42 (dxdxd, 1H, J=13.56, 6.23, 3.66, $H_{2'}$). The product was further dried by azeotropic removal of water with pyridine. The 1-β-(4-phenylimidazole)-2-deoxyribose (0.94 g, 3.6 mmol) was dissolved in dry pyridine (6 mL). 4,4-Dimethoxytritylchloride (1.22 g, 3.6 mmol) was added to the solution, and this mixture was stirred at 4°C. Additional portions of DMTCl (60 mg, 0.18 mmol) were added until most of the starting material was converted to product as determined by TLC. The reaction was stopped by the addition of methanol (3 mL) and diluted in CH_2Cl_2 (50 mL). The organic layer was washed twice with water, dried over anhydrous Na_2SO_4, filtered and evaporated to an orange foam. Chromatography (EtOAc/hexane/MeOH, eluted initially with 20:10:0 then 20:10:1, v/v/v, TLC: R_f=.21) yielded 1.72 g (85%) of the product as a white foam. ^1H NMR ($CDCl_3$) δ 7.57 (d, 1H, J=0.74, imidazole); δ 7.48 (m, 2H, J=7.32, phenyl); δ 7.39 (m, 2H, J=7.32, phenyl); δ 7.34 (d, 1H, J=1.10, imidazole); δ 7.29 (d, 2H, J=8.80, methoxyphenyl); δ 7.27 (d, 2H, J=8.43, methoxyphenyl); δ 7.24 (d, 2H, J=8.43, phenyl); δ 7.20 (d, 2H, J=8.43, phenyl); δ 7.17-7.12 (m, 2H, phenyl); δ 6.75 (d, 4H, J=9.16, methoxyphenyl); δ 5.95 (t, 1H, J=5.87, $H_{1'}$); δ 4.58 (m, 1H, $H_{3'}$); δ 4.14 (m, 1H, $H_{4'}$); δ 3.64 (s, 6H, OCH3); δ 3.30 (m, 2H, $H_{5'}$); δ 2.45 (dxdxd, 1H, J=13.56, 5.86, $H_{2'}$); δ 2.34 (dxdxd, 1H, J=13.56, 5.86, 2.93, $H_{2'}$). FAB MS calcd. for $C_{35}H_{35}N_2O_5$ ($M+H^+$): 563.2596. Found: 563.2544 .

1-β-(4-phenylimidazole)-5-DMT-2-deoxyribose-*N,N*-diisopropyl-β-cyano-phosphoramidite 1 (Z) was prepared using the method of McBride and Caruthers (47). 2-Cyanoethyl-*N,N*-diisopropyl-chlorophosphoramidite (100 μL, 0.45 mmol) was added dropwise to a solution of alcohol **4** (169 mg, 0.30 mmol) and

diisopropylethylamine (156 μL, 0.90 mmol) in dry CH_2Cl_2 (1.2 mL, puurified by passing through basic aluminum oxide) under Argon. After 1 hour, the completed reaction was quenched with ethanol (0.5 mL), diluted with EtOAc (30 mL), and washed with satd. aq. $NaHCO_3$ (2 x 30 mL) and brine (2 x 30 mL). The organic layer was dried over anhydrous Na_2SO_4 and then concentrated. The phosphoramidite diastereomers were purified by flash chromatography on a small column (5 g silica gel) (EtOAc/hexane/MeOH/TEA, 64:32:2:2, v/v/v/v, TLC: R_f=.47 and .35). The resulting white foam was further dried by concentration from dry CH_2Cl_2 to give the product in a 94% yield. FAB MS calcd. for $C_{44}H_{52}N_4O_6P$ ($M+H^+$): 763.3699. Found: 763.3625.

4-(3-nitro)phenylimidazole 5 (40). 3-Nitroacetophenone (12.0 g, 72.7 mmole) was dissolved in chloroform (100 mL) and heated to 48° C. Bromine (11.68 g in 15 mL chloroform, 73.1 mmol) was added dropwise over 10 min to the stirring solution. Upon addition the orange Br_2 color vanished and gas (HBr) evolution was observed. The solution was stirred for an additional 10 min at room temperature and then evaporated. The residue was dissolved in CH_2Cl_2 and washed with satd. $NaHCO_3$ and dried over anhydrous Na_2SO_4, filtered, evaporated, and crystalized from 1:1 CH_2Cl_2/ether to give 3-nitrobromoacetophenone in 79% yield (TLC: CH_2Cl_2, R_f=.64; m.p. 93.5-94.5°C, lit. 96°C (Evans and Brock)). ^1H NMR (Me_2SO-d_6) δ 8.71 (t, 1H, J=1.83, H_2); δ 8.51 (dxdxd, 1H, J=8.06, 2.20, 1.10, aromatic); δ 8.42 (dxd, 1H, J=8.06, 1.10, aromatic); δ 7.87 (t, 1H, J=8.06, aromatic, H_5); δ 5.08 (s, 2H, CH_2). 3-Nitrobromoacetophenone (7.32 g, 30 mmol) was dissolved in formamide (36.6 mL) and heated to 185 °C. Between 50-70 °C the starting material dissolved to give a bright yellow solution, which on further heating turned red and finally dark brown. The reaction was heated at 180-185 °C for 2 hours and then cooled to room temperature. The dark slurry was poured over a cation exchange column (Biorad AG 50 W X5 50-100-mesh hydrogen form) and eluted first with water, then with 10% ammonia, and finally with methanol. The methanol fractions gave the desired crude material as a brown solid. Recrystalization from boiling water (1 L) gave the pure product

5 as a yellow powder in 51% yield (2.69 g, 15.3 mmol) (m.p. 223-224 °C, lit. 224 °C; TLC CH_2Cl_2/isopropanol/NH_4OH, 49:49:2, v/v/v, R_f=.35). ^1H NMR (Me_2SO-d_6) δ 8.59 (t, 1H, J=1.83, H_2); δ 8.21 (d, 1H, J=8.06); δ 8.02 (dxt, 1H, J=8.06, 1.10); δ 7.87 (s, 1H, $H_{2'}$); δ 7.80 (s, 2H, $H_{5'}$); δ 7.64 (t, 1H, J=8.06, H_5).

1-β-(4-(3-nitro)phenylimidazole)-2-deoxy-3,5,-di-O-*p*-toluoyl-ribose 6. The same procedure for the synthesis of **3** was followed with 4-(3-nitro)phenylimidazole (38). The product (R_f=.48) and the 3,5-diprotected sugar (37) (R_f=.46) can be distinguished by TLC using an acidic anisaldyhyde stain, which on heating stains the product blue and the diprotected sugar purple. Crystallization from EtOH and a small amount of CH_2Cl_2 gave (53%) off white needles (m.p. 144-145 °C). ^1H NMR ($CDCL_3$) δ 8.46 (t, 1H, J=1.94, phenyl); δ 8.06 (dxdxd, 1H, J=8.06, 2.20, 0.98, phenyl); δ 7.96 (d, 2H, J=8.30, toluoyl); δ 7.94 (d, 2H, J=8.30, toluoyl); δ 7.85 (dxdxd, 1H, J=7.81, 1.71, 0.98, phenyl); δ 7.75 (d, 1H, J=1.46, imidazole); δ 7.45 (t, 1H, J=7.81, phenyl); δ 7.45 (s, 1H, imidazole); δ 7.29 (d, 2H, J=8.08, toluoyl); δ 7.24 (d, 2H, J=8.08, toluoyl); δ 6.16 (dxd, 1H, J=7.81, 6.23, $H_{1'}$); δ 5.74-5.71 (m, 1H, $H_{3'}$); δ 4.76 (dxd, 1H, J=12.18, 4.63, $H_{5'}$); δ 4.64-4.60 (m, 2H, $H_{4'}$, $H_{5'}$); δ 2.79–2.71(m, 2H, $H_{2'}$); δ 2.45 (s, 3H, CH_3); δ 2.39 (s, 3H, CH_3). FAB MS, calc. for $C_{30}H_{28}N_{23}O_7$ ($M+H^+$): 542.1967. Found: 542.1927. Anal. calcd. for $C_{30}H_{27}N_3O_7$: C, 66.54; H, 5.03; N, 7.76. Found: C, 66.74; H, 5.05; N, 7.44.

1-β-(4-(3-nitro)phenylimidazole)-5-dimethoxytrityl-2-deoxyribose 7a and 7b. The toluoyl groups were removed from the sugar **6** by the same procedure described for nucleoside **3** (39). Column chromatography (CH_2Cl_2/toluene/MeOH, 76:12:12, v/v/v, TLC: R_f=.29) yielded 1-β-(4-(3-nitro)phenylimidazole)-2-deoxyribose in 96%. ^1H NMR (Me_2SO-d_6) δ 8.57 (t, 1H, J=2.20, 1.83, phenyl); δ 8.20 (dxt, 1H, J=8.06, phenyl); δ 8.12 (d, 1H, J=1.10, imidazole); δ 8.05 (dxdxd, 1H, J=8.06, 2.20, 1.10, phenyl); δ 8.00 (d, 1H, J=1.10, imidazole); δ 7.66 (t, 1H, J=8.06, phenyl); δ 6.08 (t, 1H, J=6.22, $H_{1'}$); δ 5.31 (d, 1H, J=4.03, 3'OH); δ 4.95 (t, 1H, J=5.50, 5'OH); δ

4.34 (m, 1H, $H_{3'}$); δ 3.82 (m, 1H, $H_{4'}$); δ 3.59-3.49 (m, 2H, $H_{5'}$); δ 2.44 (dxdxd, 1H, J=13.56, 7.32, 5.86, $H_{2'}$); δ 2.27 (dxdxd, 1H, J=13.56, 6.23, 3.30, $H_{2'}$). 1-β-(4-(3-nitro)phenylimidazole-2-deoxyribose was converted to 1-β-(4-(3-amino)phenylimidazole-2-deoxyribose by hydrogenation. A solution of the nitro compound in MeOH (60 mL) was shaken with 10% Pd/C (80 mg) under H_2 [50 psi] for 2 hours , using a Parr apparatus. Filtration and evaporation afforded analytically pure compound (TLC: CH_2Cl_2/toluene/MeOH, 8:1:1, v/v/v, R_f=.10) in 99% yield. ^1H NMR (Me_2SO-d_6) δ 7.85 (s, 1H, aromatic); δ 7.62 (s, 1H, aromatic); δ 7.02 (s, 1H, aromatic); δ 6.89 (d, 1H, J=2.57, aromatic); δ 6.41 (dxt, 1H, J=7.81, 1.10, aromatic); δ 6.03 (t, 1H, J=6.60, $H_{1'}$); δ 5.26 (d, 1H, J=3.90, 3'OH); δ 4.99 (bs, 2H, NH_2); δ 4.91 (t, 1H, J=5.37, 5'OH); δ 4.31 (m, 1H, $H_{3'}$); δ 4.09 (m, 1H, $H_{4'}$); δ 3.56-3.46 (m, 2H, $H_{5'}$); δ 2.40 (m, 1H, $H_{2'}$); δ 2.23 (dxdxd, 1H, J=13.43, 6.86, 2.93, $H_{2'}$).

The amino group was protected by acylation using a transient protection method (42). 1-β-(4-(3-amino)phenylimidazole)-2-deoxyribose (.743 g, 2.7 mmol) was suspended in dry pyridine (25 ml) in a 100 ml round bottom flask equipped with a drying tube and cooled in an ice bath. Trimethylchlorosilane (5 eq., 1.17 mL, 13.5 mmol) was added dropwise with stirring. After 30 min trifluoroacetic anhydride (5 eq., 1.91 mL, 13.5 mmol) was added dropwise to the stirring solution. The reaction was allowed to warm to room temperature and stirred an additional hour. Cold water (5 mL) was added to the reaction, which was then rotoevaporated several times with toluene and MeOH. The residue was dissolved in ethylacetate, washed with satd. $NaHCO_3$, and dried over anhydrous Na_2SO_4. Chromatography (CH_2Cl_2/MeOH, 9:1, v/v, TLC: R_f=.25) yielded 0.86 g (86%) of the trifluoroacetamide protected nucleoside. ^1H NMR (Me_2SO-d_6) δ 11.26 (s, 1H, NHCO); δ 8.08 (s, 1H, aromatic); δ 7.93 (s, 1H, aromatic); δ 7.83 (s, 1H, aromatic); δ 7.59 (d, 1H, J= 7.69, phenyl); δ 7.51 (d, 2H, J=7.69, phenyl); δ 7.38 (t, 1H, J=7.92, phenyl); δ 6.06 (t, 1H, J=6.65, $H_{1'}$); δ 5.28 (d, 1H, J=3.85, 3'OH); δ 4.93 (t, 1H, J=5.56, 5'OH); δ 4.33 (m, 1H, $H_{3'}$); δ 3.82 (m, 1H, $H_{4'}$); δ 3.58-3.46 (m, 2H, $H_{5'}$);

δ 2.42 (m, 1H, $H_{2'}$); δ 2.25 (dxdxd, 1H, J=13.25, 5.98, 3.42, $H_{2'}$). The same procedure for the synthesis of the *N*-benzoyl protected nucleoside was followed using benzoyl chloride as the acylating agent instead of trifluoroacetic anhydride. Chromatography (CH_2Cl_2/toluene/MeOH, 8:1:1, v/v/v, TLC: R_f=.22) yielded 91% of the benzamide protected nucleoside. ^{1}H NMR (CD_3OD) δ 7.99-7.94 (m, 3H, aromatic); δ 7.73 (d, 1H, J=1.10, aromatic); δ 7.65 (dxt, 1H, J=8.06, 1.10, aromatic); δ 7.61-7.40 (m, 3H, aromatic); δ 7.37 (t, 1H, J=8.06, aromatic); δ 7.23-7.08 (m, 2H, aromatic); δ 6.13 (t, 1H, J=6.60, $H_{1'}$); δ 4.49 (m, 1H, $H_{3'}$); δ 3.98 (m, 1H, $H_{4'}$); δ 3.76 (dxd, 1H, J=12.09, 4.03, $H_{5'}$); δ 3.70 (dxd, 1H, J=11.73, 4.40, $H_{5'}$); δ 2.53 (m, 1H, $H_{2'}$); δ 2.42 (dxdxd, 1H, J=13.56, 6.23, 3.66, $H_{2'}$). The same procedure for the synthesis of **4** was followed for the 5'OH protection of the *N*-acylated-phenylimidazole nucleosides (42). Chromatography (EtOAc/hexane/MeOH, 65:32:3, v/v/v, TLC: R_f=.26) yielded 86% of the N-trifluoroacetamide 5'-DMT protected nucleoside **7a**. ^{1}H NMR (Me_2SO-d_6) δ 11.25 (s, 1H, NHCO); δ 8.11 (s, 1H, aromatic); δ 7.92 (s, 1H, aromatic); δ 7.75 (s, 1H, aromatic); δ 7.49 (d, 1H, J= 7.27, phenyl); δ 7.38-7.16 (m, 12H, aromatic); δ 6.84 (d, 2H, J=7.27, aromatic); δ 6.81 (d, 2H, J=7.27, aromatic); δ 6.11 (t, 1H, J=6.20, $H_{1'}$); δ 5.36 (d, 1H, J=4.27, 3'OH); δ 4.36 (m, 1H, $H_{3'}$); δ 3.94 (m, 1H, $H_{4'}$); δ 3.68 (s, 3H, OCH_3); δ 3.66 (s, 3H, OCH_3); δ 3.16 (dxd, 1H, J=9.40, 5.56, $H_{5'}$); δ 3.10 (dxd, 1H, J=10.26, 2.99, $H_{5'}$); δ 2.54 (m, 1H, $H_{2'}$); δ 2.33 (m, 1H, $H_{2'}$). FAB MS calcd. for $C_{37}H_{34}N_3F_3O_6Li$ ($M+Li^+$): 680.2610. Found: 680.2560. Chromatography (EtOAc/hexane/MeOH, 64:30:6, v/v/v, TLC: R_f=.29) yielded 88% of the benzamide-5'-DMT protected nucleoside **7b**. ^{1}H NMR ($CDCl_3$) δ 7.87-7.81 (m, 4H, aromatic; δ 7.67 (s, 1H, aromatic); δ 7.58-7.18 (m, 15H, aromatic); δ 6.81 (d, 2H, J=8.80 , aromatic); δ 6.79 (d, 2H, J=8.80, aromatic); δ 6.06 (t, 1H, J=6.60, $H_{1'}$); δ 4.61 (m, 1H, $H_{3'}$); δ 4.10 (m, 1H, $H_{4'}$); δ 3.72 (s, 6H, OCH_3); δ 3.39 (dxd, 1H, J=10.26, 4.03, $H_{5'}$); δ 3.35 (dxd, 1H, J=10.26, 4.40, $H_{5'}$); δ 2.58 (m, 1H, $H_{2'}$); δ 2.47 (m, 1H, $H_{2'}$). FAB MS calcd. for $C_{42}H_{40}N_3O_6$ ($M+H^+$): 682.975. Found: 682.2917 .

1-β-(4-(3-trifluoroacetamide or 3-benzamide)phenylimidazole)-5-DMT-2-deoxyribose-*N,N*-diisopropyl-β-cyano-phosphoramidite **2a** and **2b** were prepared by the same procedure describe in the synthesis of nucleoside **1** (47). Chromatography of the phosphoramidite diastereomers (EtOAc/hexane/TEA, 66:32:2, v/v/v, TLC: R_f=.29 and .19) gave nucleoside **2a** in a 96% yield. FAB MS calcd. for $C_{46}H_{51}N_5F_3O_7PLi$ ($M+Li^+$): 880.3712. Found: 880.3632. Chromatography of the phosphoramidite diastereomers (EtOAc/hexane/MeOH/TEA, 64:32:2:2, v/v/v/v, TLC: R_f=.27 and .18) gave nucleoside **2b** in a 90% yield. FAB MS calcd. for $C_{51}H_{57}N_5O_7P$ ($M+H^+$): 882.4078. Found: 882.3996.

Oligonucleotides. Oligonucleotides **1-34** were machine synthesized using β-cyano phosphoramidite chemistry (43). Modified nucleosides gave a ≥97% coupling efficiency. The oligonucleotides were removed from the support and deprotected by treatment with 0.1N NaOH (1.5 mL/1 umol synthesis, 55 °C, 24 hr), neutralized with glacial acetic acid (6-7 μl/1.5 ml .1N NaOH), applied to a column of Sephadex G-10-120 (Sigma), and eluted with H_2O. The crude oligonucleotides were lyophilized and purified by electrophoresis (~550 V, 20 hr) on a 2-mm-thick 20% polyacrylamide gel (Maxam-Gilbert). The major UV-absorbing bands were cut out and eluted (0.2N NaCl, 1 mM EDTA, 37 °C, 24 hr), then passed through Sephadex G-10-120, and extensively dialyzed. Oligonucleotide 4-mers **T-T-Z-T**, **T-T-Z'-T**, and **T-T-Z"-T** for NMR studies were synthesized on a 10 μmol scale in a similar manner. The crude deprotected desalted 4-mers were filtered (.45 μ) and lyophilized. ^1H NMR (Me_2SO-d_6) **T-T-Z-T**: δ 7.96 (s, 1H, imidazole); δ 7.91 (s, 1H, T); δ 7.82 (s, 1H, imidazole); δ 7.79 (d, 2H, J= 2.20, phenyl); δ 7.78 (s, 1H, T); δ 7.75 (s, 1H, T); δ 7.31 (t, 2H, J=7.57, phenyl); δ 7.16 (t, 1H, J=7.32, phenyl); δ 6.21–6.02 (4xt, 4H, $H_{1'}$). **T-T-Z'-T**: <5% amide present. ^1H NMR (Me_2SO-d_6) **T-T-Z"-T**: d 10.56 (s, 1H, NHCO); δ 8.18 (s, 1H, imidazole); δ 8.06 (s, 1H, J=1.46, benzoyl); δ 8.04 (s, 1H, aromatic); δ 7.92 (s, 1H, T); δ 7.84 (d, 1H, J=9.16,

phenyl); δ 7.83 (s, 1H, aromatic); δ 7.77 (s, 1H, T); δ 7.57-7.48 (m, 4H, aromatic); δ 7.29 (t, 1H, J=8.06, phenyl); δ 6.20-6.04 (4xt,4H, $H_{1'}$).

DNA Manipulations. Distilled, deionized water was used for all aqueous reactions and dilutions. Enzymes were purchased from either Boehringer-Mannheim or New England Biolabs and were stored and used under conditions of temperature, buffer, etc. recommended by the manufacturer. Deoxynucleoside triphosphates were purchased from Pharmacia as 100 mM solutions. 5'-[α-^{32}P]dNTPs (3000 Ci/mmol) and 5'-[γ-^{32}P]ATP (>5000 Ci/mmol) were obtained from Amersham. SV40 DNA was purchased from Besthesda Research Laboratories. DNA reactions were in general carried out in 1.5 mL polypropylene tubes obtained from Tekmar Company. Calf thymus DNA was purchased from Sigma, sonicated, and then extracted with 3 x 0.2 volumes (vol.) water-saturated phenol, 0.2 vol. 24:1 $CHCl_3$/isoamyl alcohol, and 0.2 vol. $CHCl_3$. After extensive dialysis against H_2O, the DNA was passed through a 0.45 μ Centrex filter (Schleicher and Schuell) and assayed for concentration by UV (assuming ε_{260} 11,800 $Lmol^{-1}bp^{-1}cm^{-1}$. DNA was precipitated from either 1:3 0.3 M NaOAc (pH 5.2)/EtOH or 1:2 2.5 M NH_4OAc/EtOH. DNA pellets were dried in a Savant Speed Vac. Agarose gel electrophoresis was carried out using 40 mM tris-acetate, 5 mM NaOAc, 1 mM EDTA, pH 7.9 buffer. Ten-fold concentrated (10X) loading buffer for agarose gels was 25% (w/v) ficoll solution, which contained 0.2% (w/v) bromophenol blue (BPB) and xylene cyanol (XC) tracking dyes. Polyacrylamide gel electrophoresis was carried out using 100 mM tris-borate, 1 mM EDTA, pH 8.3 buffer. Loading buffer for denaturing polyacrylamide gels was 80% formamide in 100 mM tris-borate, 1 mM EDTA, pH 8.3 buffer. 3' and 5' labeled restriction fragments of SV40 or pHIV-CAT DNA (49) were prepared by standard procedures (50). Fragments were separated by gel electrophoresis and isolated by elution form the gel. Specific radioactivity was measured with a Beckman LS 3801 scintillation counter. Gels were dried (after transferring them to Whatman 3 MM paper) on a Bio-Rad Model 483 slab drier at 80 °C. Autoradiography was carried out using Kodak X-Omat

film. In most cases the photographic emulsion on the side of the film away from the gel was removed with bleach in order to obtain sharper images and higher contrast. Optical densitometry was performed using an LKB Bromma Ultroscan XL Laser Densitometer operating at 633 nm. Relative peak area for each cleavage band or locus was equated to the relative cleavage efficiency at that site.

Procedure for the Preparation of 5'-^{32}P End-Labeled 30-mer duplexes. 50 picomoles of each single-stranded oligonucleotide $A_5T_7YT_7G_{10}$ (Y=T, C, G, A) were dissolved in 28 μL H_2O, 5 μL 10X kinase buffer (10X = 700 mM Tris·HCl, 100 mM $MgCl_2$, 1 mM spermidine, 1 mM EDTA, pH 7.6), and 5 μL 10 mg/mL DTT. This solution was treated with 10 μL 5'-[α-^{32}P]ATP and 2 μL T4 Polynucleotide Kinase (20 u). The 5' end-labeling was allowed to proceed 45 min at 37 °C. The end-labeling reactions were precipitated from NH_4OAc/EtOH, washed with 70% EtOH, and dried. The pellets were dissolved in 100 mM NaCl, 50 mM tris-acetate, pH 7.4, and 50 pmol of the complementary 30-mer in a total volume of 30 μl. To effect hybridization of the 30-mers the oligonucleotide mixture was heated to 90 °C for 5 min and then allowed to cool to room temperature. The oligonucleotide mixtures were treated with 5 μL 25% ficoll loading buffer, loaded onto a 15 cm long x 2 mm thick 1:20 cross-linked 15% nondenaturing polyacrylamide gel and electrophoresed at 240 V until the BPB tracking dye was near the bottom of the gel. The four bands (Y=T, C, G, A) of end-labeled DNA were visualized by autoradiography, excised from the gel, crushed, transferred to 1.5 mL polypropylene tubes, treated with 1 mL 0.2N NaCl and eluted at 37 °C for 24 hours. The end-labeled DNA 30-mer duplexes were recovered by passing the eluents through a 0.45 μ Centrex filter and extensively dialyzing against H_2O.

References and Notes

1. Dervan, P. B. *Science* **1986**, *232*, 464.
2. Dervan, P. B. in *Nucleic Acids and Molecular Biology*, Eckstein, F. and Lilley, D. M. J. Ed. (Springer-Verlag, Heidelberg, **1988**) *vol. 2*, pp. 49-64.
3. Barton, J. K. *Science* **1986**, *233*, 727.
4. Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023.
5. Hoogsteen, K. *Acta Crystallogr.* **1959**, *12*, 822.
6. Michaelson, A. M.; Massoulie, J.; Guschlbauer, W. *Prog. Nucleic Acids Res. Mol. Biol.* **1967**, *6*, 83.
7. Felsenfeld, G. and Miles, H. T. *Annu. Rev. Biochem.* **1967**, *36*, 407.
8. Lipsett, M. N. *Biochem. Biophys. Res. Commun.* **1963**, *11*, 224.
9. Lipsett, M. N. *J. Biol. Chem.* **1964**, *239*, 1256.
10. Howard, F. B.; Frazier J.; Lipsett, M. N.; Miles, H. T. *Biochem. Biophys. Res. Commun.* **1964**, *17*, 93.
11. Miller, J. H. and Sobell, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 1201.
12. Morgan, A. R. and Wells, R. D. *J. Mol. Biol.* **1968**, *37*, 63.
13. Lee, J. S.; Johnson, D. A.; Morgan, A. R. *Nucleic Acids Res.* **1979**, *6*, 3073.
14. Mirkin, S. M. *et al. Nature* **1987**, *330*, 495.
15. Htun, H. and Dahlberg, J. E. *Science* **1988**, *241*, 1791.
16. Kohwi, Y. and Kohwi-Shigematsu, T. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3781.
17. Hanvey, J. C. *et al. ibid.* **1988**, *85*, 6292.
18. Hanvey, J. C. *et al. J. Biol. Chem. 1988*, **263**, 7386.
19. Rajagopal, P. and Feigon, J. *Nature* **1989**, *239*, 637.
20. Moser, H. E. and Dervan, P. B. *Science* **238**, 645 (1987).

21. Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7927.

22. Arnott, S. and Bond, P. J. *Nature (London) New Biol.* **1973** *244*, 99.

23. Arnott, S. and Selsing, E. *J. Mol. Biol.* **1974**, *88*, 509.

24. Arnott, S.; Bond, P. J.; Selsing, E.; Smith, P. J. C. *Nucleic Acids Res.* **1976**, *3*, 2459.

25. Saenger, W. *Principles of Nucleic Acid Structure*, C. R. Cantor, Ed. (Springer-Verlag, New York, **1984**).

26. Povsic, T. J. and Dervan, P. B. *J. Am. Chem. Soc.*. **1989**, *111*, 3059 .

27. Hertzberg, R. P. and Dervan P. B. *Biochemistry* **1984**, *23*, 3934.

28. Povsic, T. and Dervan, P. B. unpublished observations.

29. Maher, L. J.; Wold, B. J.; Dervan, P. B. *Science* in press (**1989**).

30. Marck, C.and Thiele, D. *Nucleic Acids Res.* **1987**, *5*, 1017.

31. Broitman, S. L.; Im, D. D.; Fresco, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 5120.

32. Letai, A. J. *et al.*, *Biochemistry* **1988**, *27*, 9108.

33. Cooney, M. *et al.*, *Science* **1988**, *241*, 456.

34. Doan, T. L. *et al. Nucleic Acids Res.* **1987**, *15*, 7749.

35. Praseuth, D. *et al. Proc. Natl. Acad. Sci. U.S.A.* 1988, **85**, 1349.

36. Griffin, L. C. and Dervan, P. B. *Science* in press (**1989**).

37. Hoffer, M. *Chem. Ber.* **1960**, *93*, 2777.

38. Kazimierczuk, Z.; Cottam, H. B.; Revanker, G. R.; Robbins, R. K. *J. Am. Chem. Soc.* **1984**, *106*, 6379-6382.

39. Mashimo, K. and Sato, Y. *Tetrahedron* **1970**, *26*, 803.

40. Bredereck, H. and Theilig, G. *Chem. Ber.* **1953**, *86*, 88-96.

41. Kool, E.; Marsh, R. E.; Henling, L. M.; Schaefer, W. P.; Dervan, P. B. unpublished data.

42. Ti, G. S.; Gaffney, B. L.;Jones, R. A. *J. Am. Chem. Soc.* **1982**, *104*, 1316.

43. Sinha, N. D.; Biernat, J.; McManus, J.; Koster, H. *Nucleic Acids Res.* **1984**, *12*, 4539.

44. Dreyer, G. B. and Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968.

45. We thank S. Singleton for valuable assistance in model building.

46. Dervan *et al.* unpublished observations.

47. McBride, L. J. and Caruthers, M. H. *Tetrahedron Lett.* **1983**, *24*, 245.

48. Evans, W. L. and Brock, B. T. *J. Am. Chem. Soc.* **1908**, *30*, 404-412.

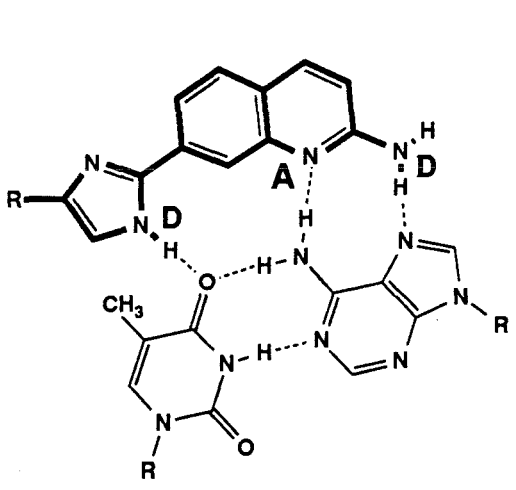
49. We thank Dr. Alan Frankel for the gift of pHIV-CAT and Tom Povsic for large scale preparation.

50. Maniatis, T.; Fritsch, E. F.; Sambrook, J. Molecular Cloning, a laboratory manual, Cold Spring Harbor Laboratory (1982).

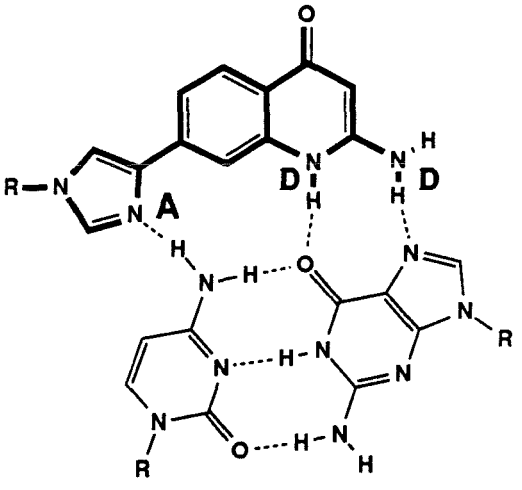
51. The pH values are not corrected for temperature or the presence of ethanol and are given for the tenfold concentrated buffer solutions at 25 °C.

52. We are grateful to the National Institutes of Health, the Office of Naval Research and the Parsons Foundation for generous support. We thank Dr. Heinz Moser for valuable discussions.

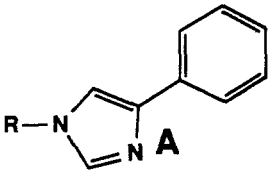
Appendix
Chapter 3



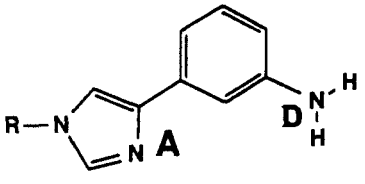
?•TA



?•CG



(Z)



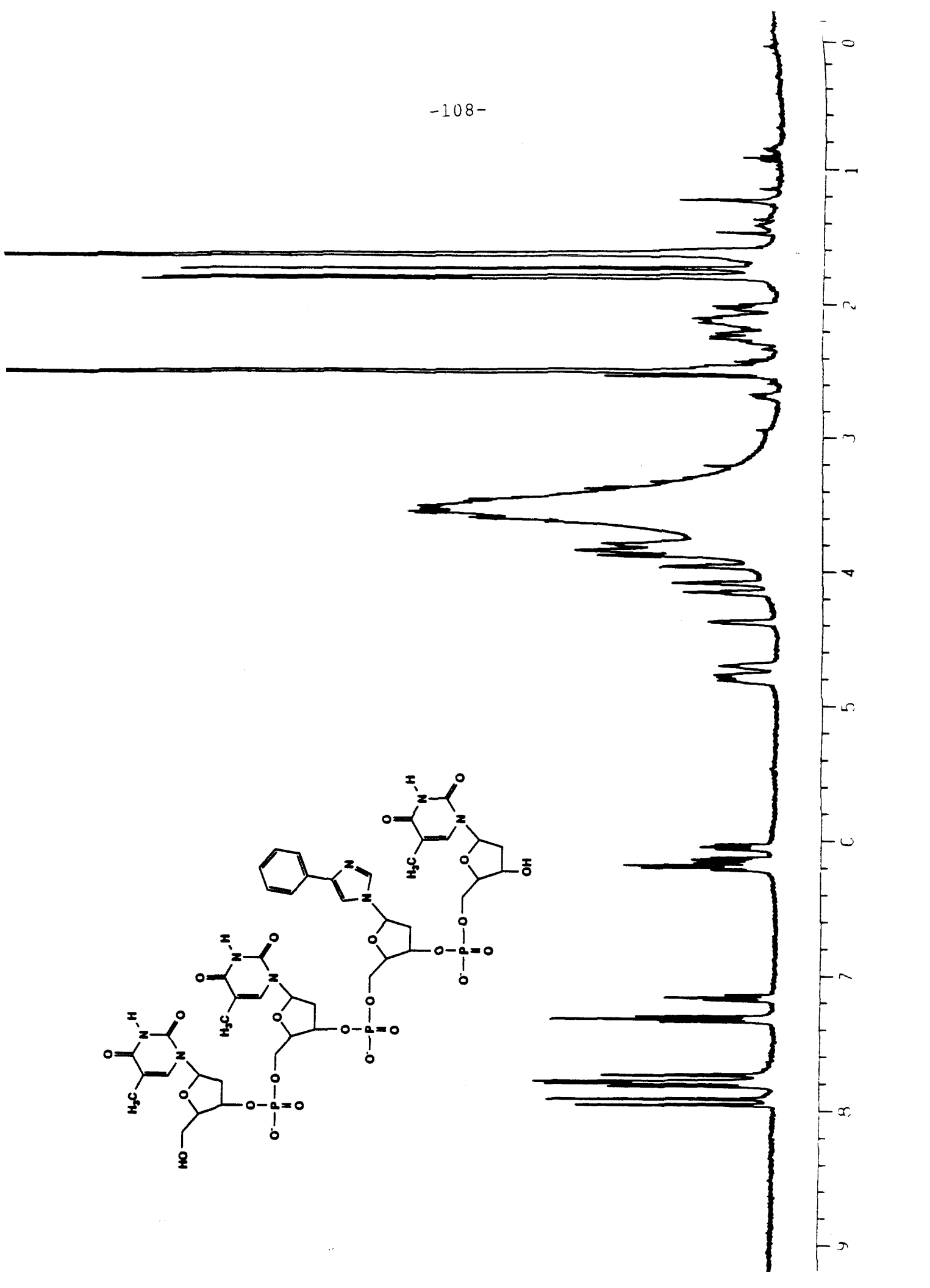
(Z')

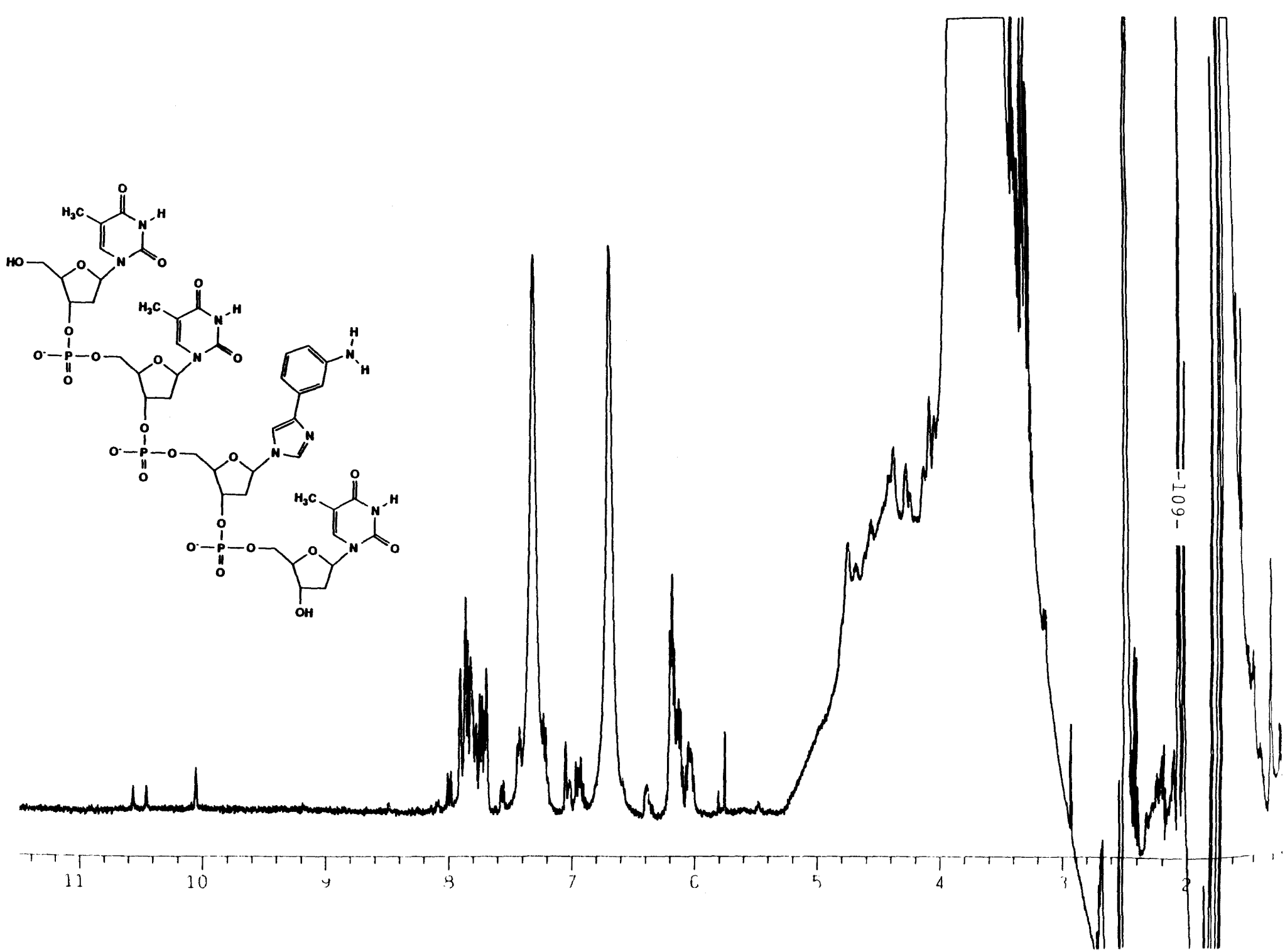
Figure 1 (Above) Possible new compounds (in bold) for base triplets ?•TA and ?•CG that place hydrogen bond donors (**D**) and acceptors (**A**) in position to form hydrogen bonds to the desired Watson-Crick base pair. (**Below**) Two initial targets **Z** and **Z'** for CG recognition.

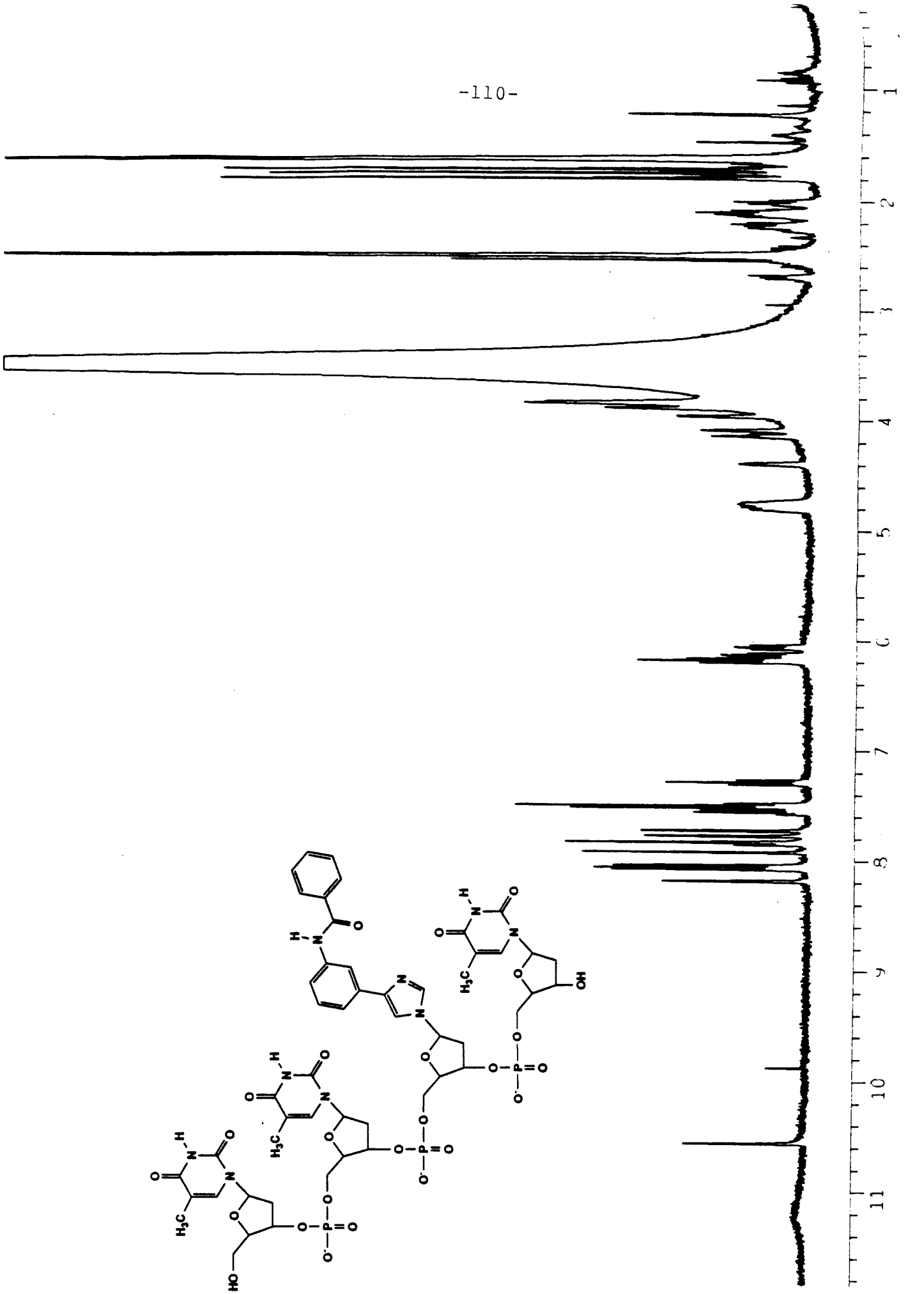
Figure 2A. ^{1}H NMR (Me_2SO-d_6) of the oligonucleotide tetramer 5'-T-T-Z-T-3'.

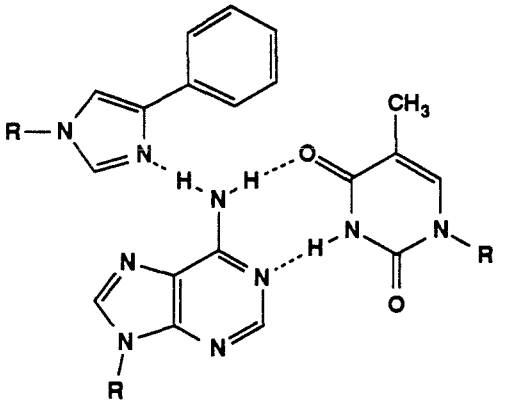
Figure 2B. ^{1}H NMR (Me_2SO-d_6) of the oligonucleotide tetramer 5'-T-T-Z'-T-3'.

Figure 2C. ^{1}H NMR (Me_2SO-d_6) of the oligonucleotide tetramer 5'-T-T-Z"-T-3'.

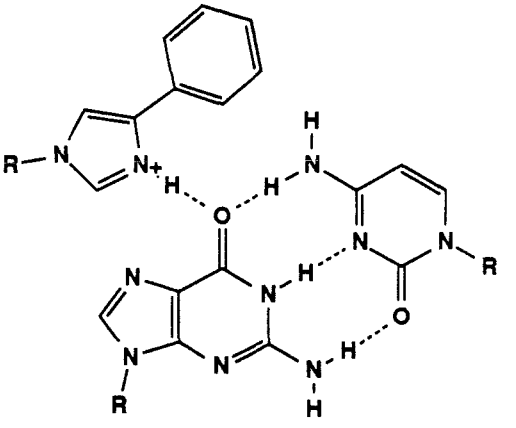

0
1
2
3
4
5
6
7
8
9


11
10
9
8
7
6
5
4
3
2
-109-

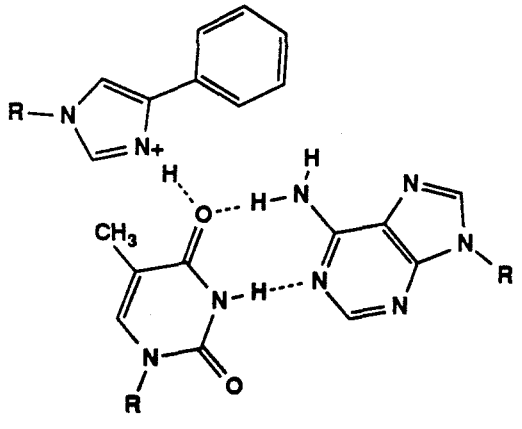

1
2
3
4
5
6
7
8
9
10
11



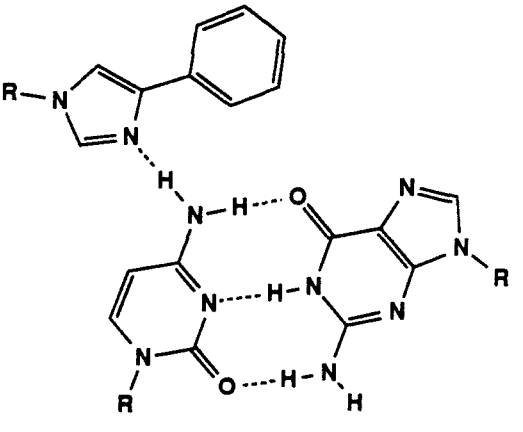
Z•AT



Z•GC

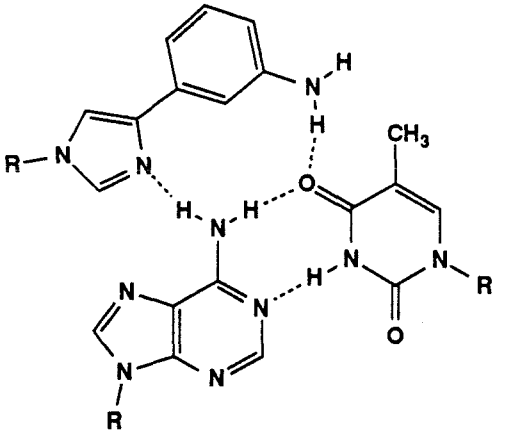


Z•TA

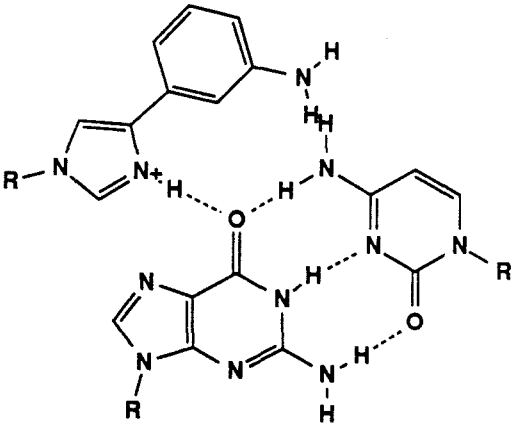


Z•CG

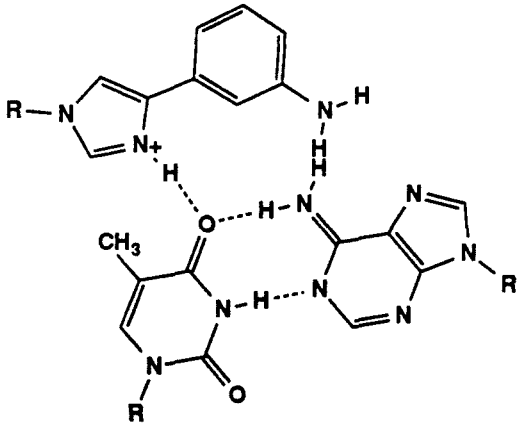
Figure 3A. Possible base triplets **Z**•AT, **Z**•GC, **Z**•CT, **Z**•TA. For each base triplet the positioning of the third base with respect to the Watson-Crick base pair is based upon forming possible hydrogen bonds, an *anti* conformation at the glycosidic link, and a backbone alignment close to the T•AT and C+GC base triplets.



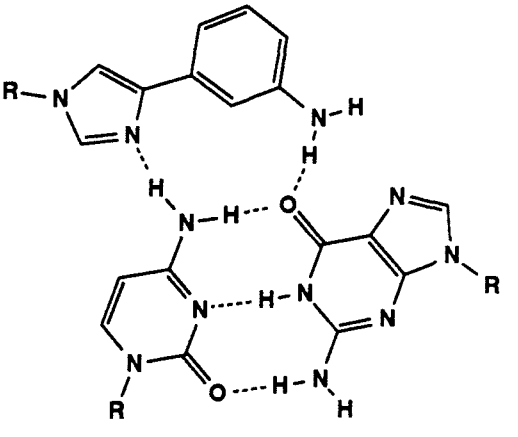
Z'•AT



Z'•GC

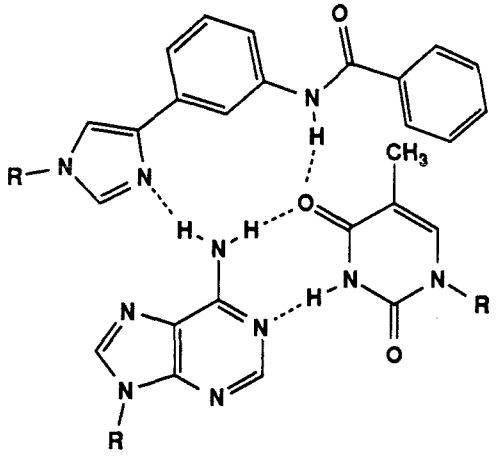


Z'•TA

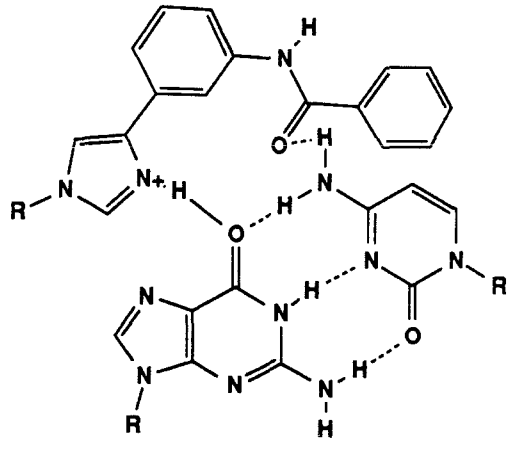


Z'•CG

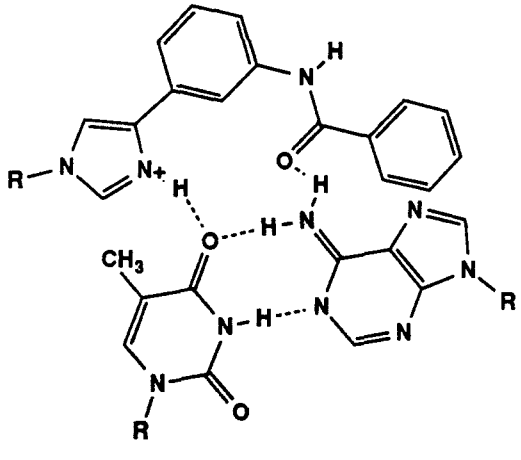
Figure 3B. Possible base triplets Z'•AT, Z'•GC, Z'•CT, Z'•TA. For each base triplet the positioning of the third base with respect to the Watson-Crick base pair is based upon forming possible hydrogen bonds, an *anti* conformation at the glycosidic link, and a backbone alignment close to the T•AT and C+GC base triplets.



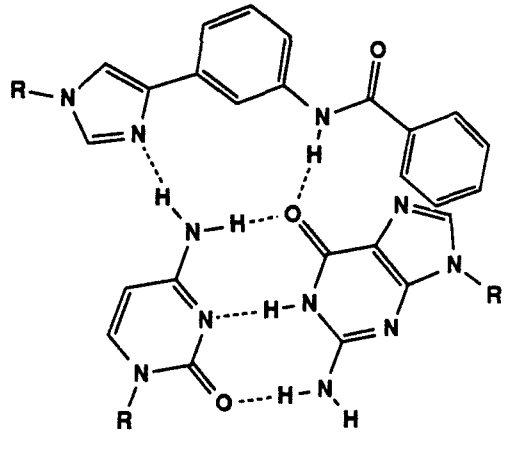
Z"•AT



Z"•GC



Z"•TA

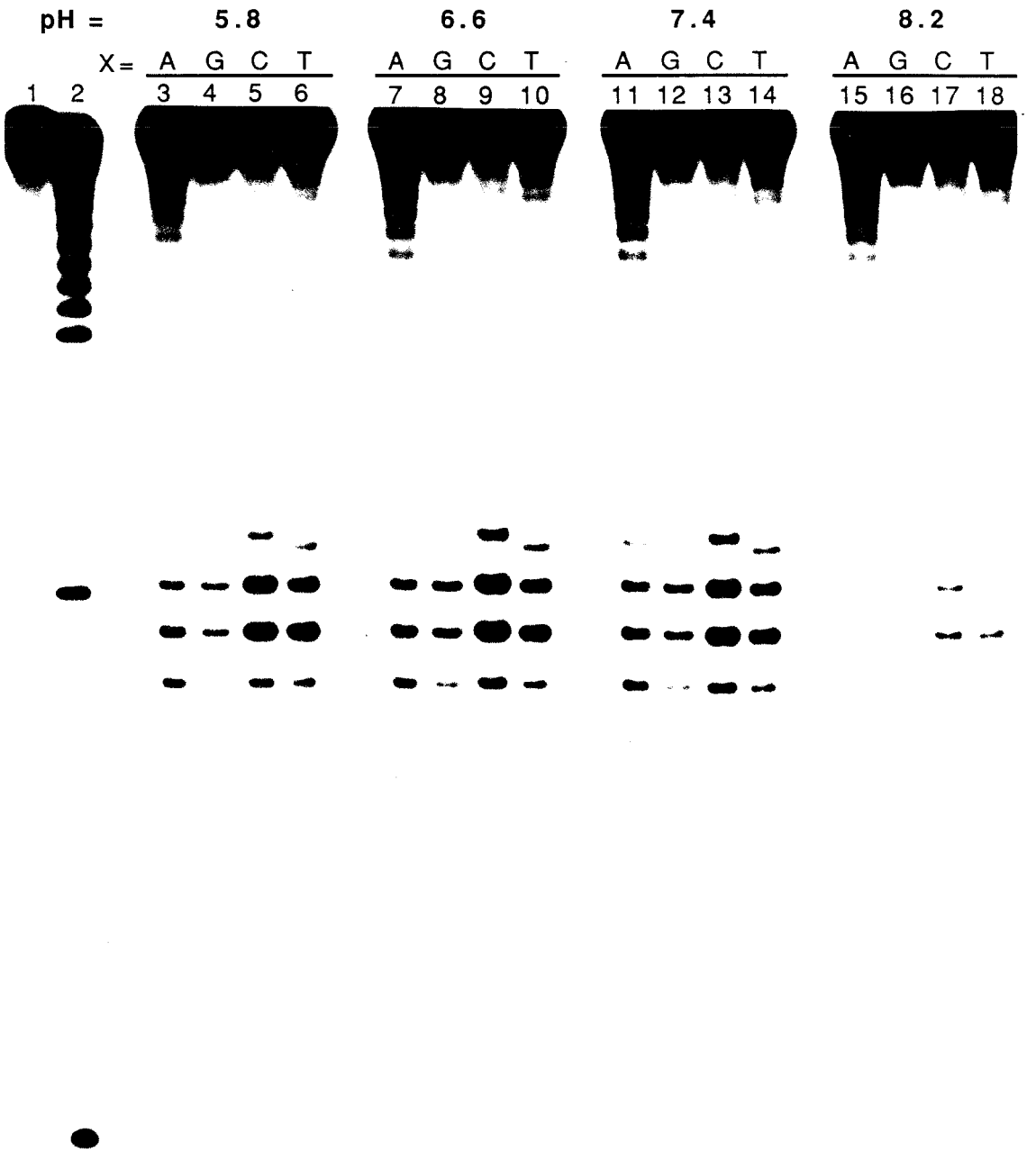


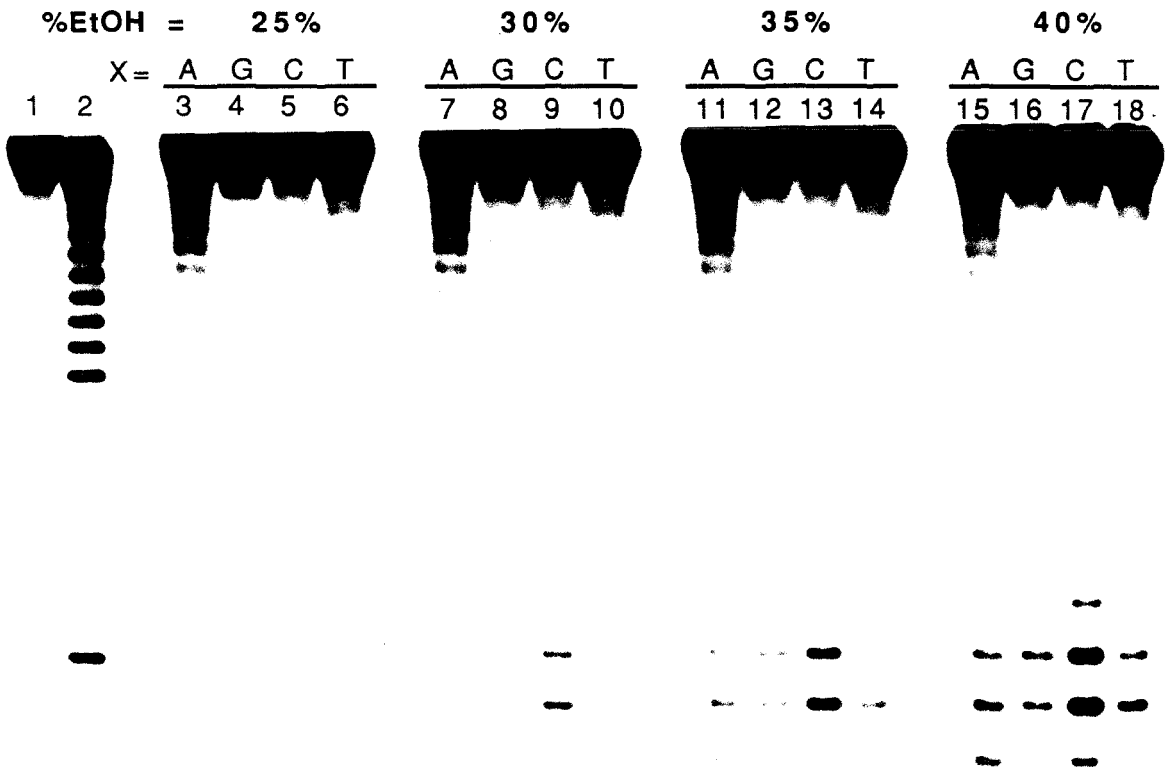
Z"•CG

Figure 3C. Possible base triplets Z"•AT, Z"•GC, Z"•CT, Z"•TA. For each base triplet the positioning of the third base with respect to the Watson-Crick base pair is based upon forming possible hydrogen bonds, an *anti* conformation at the glycosidic link, and a backbone alignment close to the T•AT and C+GC base triplets.

Figure 4A. pH study (**Z'**). Autoradiogram of the 20 percent denaturing polyacrylamide gel. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled 30-mer duplex [~0.5 μM (bp) (~14,100±600 cpm)] in a solution of tris-acetate (25 mM), NaCl (50 mM), calf thymus DNA [100 μM (bp)], and 40% ethanol and incubated at 30°C for 1 hr. Cleavage reactions were initiated by addition of DTT (3mM) and allowed to proceed for 6 hours at 30°C. The reactions were stopped by freezing and lyophilization and the cleavage products were analyzed by gel electrophoresis (1200-2000 V, BPB 23 cm). (Lanes 1-18) Duplexes containing 5' end-labeled d($A_5T_7YT_7G_{10}$). (Lane 1) Control showing intact 5' labeled 30 bp DNA standard obtained after treatment according to the cleavage reactions in the absence of oligonucleotide-EDTA. (Lane 2) Products of Maxam-Gilbert G+A sequencing reaction. (Lanes 3-18) DNA cleavage products produced by oligonucleotide-EDTA•Fe(II) **23**. pH 5.8 (lanes 3-6); pH 6.6 (lanes 7-10); pH 7.4 (lanes 11-14); pH 8.2 (lanes 15-18). **XY=AT** (Lanes 3,7,11,and 15); **XY=GC** (lanes 4,8,12, and 16); **XY=CG** (lanes 5,9,13, and 17); **XY=TA** (lanes 6,10,14, and 18).

Figure 4B. EtOH study (**Z'**). Autoradiogram of the 20 percent denaturing polyacrylamide gel. The cleavage reactions were carried out by combining a mixture of oligonucleotide-EDTA (2 μM), spermine (1 mM), and Fe(II) (25 μM) with the ^{32}P labeled 30-mer duplex [~0.5 μM (bp) (~14,200±400 cpm)] in a solution of tris-acetate, pH 7.0 (25 mM), NaCl (50 mM), calf thymus DNA [100 μM (bp)], and ethanol and incubated at 30°C for 1 hr. Cleavage reactions were initiated by addition of DTT (3mM) and allowed to proceed for 6 hours at 30°C. The reactions were stopped by freezing and lyophilization and the cleavage products were analyzed by gel electrophoresis (1200-2000 V, BPB 23 cm). (Lanes 1-18) Duplexes containing 5' end-labeled d($A_5T_7YT_7G_{10}$). (Lane 1) Control showing intact 5' labeled 30 bp DNA standard obtained after treatment according to the cleavage reactions in the absence of oligonucleotide-EDTA. (Lane 2) Products of Maxam-Gilbert G+A sequencing reaction. (Lanes 3-18) DNA cleavage products produced by oligonucleotide-EDTA•Fe(II) **23**. 25% EtOH (lanes 3-6); 30% EtOH (lanes 7-10); 35% EtOH (lanes 11-14); 40% EtOH (lanes 15-18). **XY=AT** (Lanes 3,7,11,and 15); **XY=GC** (lanes 4,8,12, and 16); **XY=CG** (lanes 5,9,13, and 17); **XY=TA** (lanes 6,10,14, and 18).


pH =
5.8
6.6
7.4
8.2
X=
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


%EtOH = 25%
30%
35%
40%
X= 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